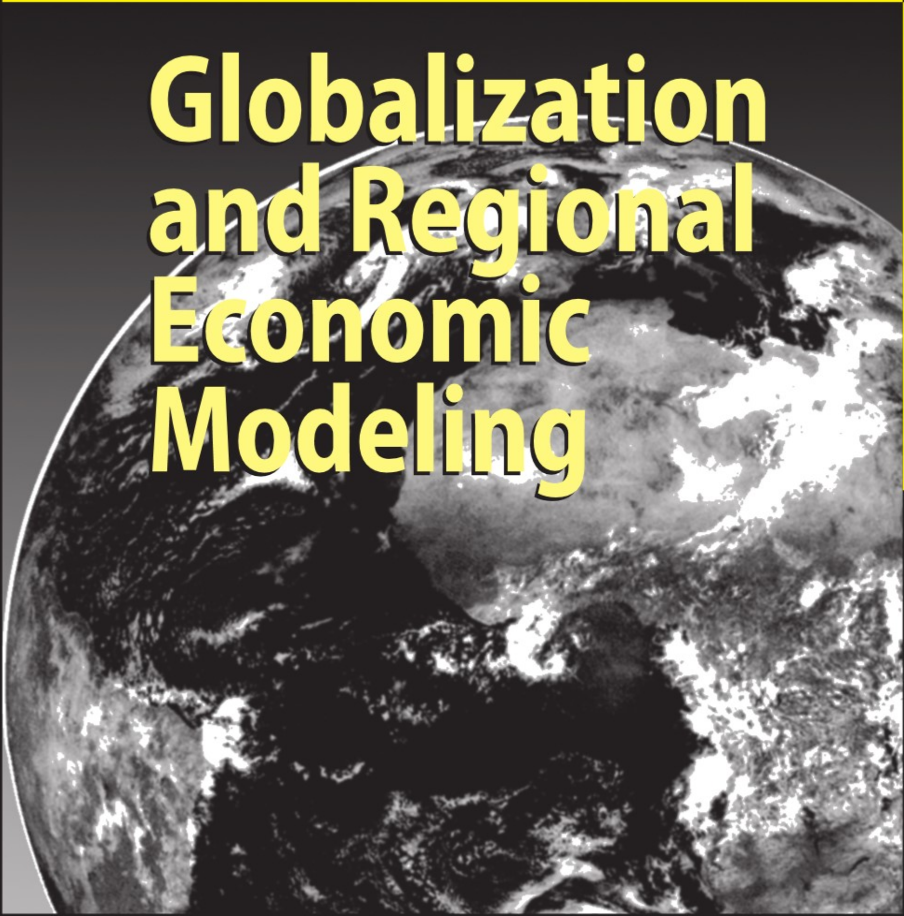
R. Cooper · K. Donaghy
G. Hewings
Editors

Globalization and Regional Economic Modeling



ADVANCES IN
SPATIAL SCIENCE

Springer

Advances in Spatial Science

Titles in the Series

M. M. Fischer, J. Revilla Diez and *F. Snickars*
Metropolitan Innovation Systems
VIII, 270 pages. 2001. ISBN 978-3-540-41967-9

L. Lundqvist and *L.-G. Mattsson* (Eds.)
National Transport Models
VIII, 202 pages. 2002. ISBN 978-3-540-42426-0

J. R. Cuadrado-Roura and *M. Parellada* (Eds.)
Regional Convergence in the European Union
VIII, 368 pages. 2002. ISBN 978-3-540-43242-5

G. J. D. Hewings, M. Sonis and *D. Boyce* (Eds.)
Trade, Networks and Hierarchies
XI, 467 pages. 2002. ISBN 978-3-540-43087-2

G. Atalik and *M. M. Fischer* (Eds.)
Regional Development Reconsidered
X, 220 pages. 2002. ISBN 978-3-540-43610-2

Z. J. Acs, H. L. F. de Groot and *P. Nijkamp* (Eds.)
The Emergence of the Knowledge Economy
VII, 388 pages. 2002. ISBN 978-3-540-43722-2

R. J. Stimson, R. R. Stough and *B. H. Roberts*
Regional Economic Development
X, 397 pages. 2002. ISBN 978-3-540-43731-4

S. Geertman and *J. Stillwell* (Eds.)
Planning Support Systems in Practice
XII, 578 pages. 2003. ISBN 978-3-540-43719-2

B. Fingleton (Ed.)
European Regional Growth
VIII, 435 pages. 2003. ISBN 978-3-540-00366-3

T. Puu
Mathematical Location and Land Use Theory,
2nd Edition
X, 362 pages. 2003. ISBN 978-3-540-00931-3

J. Bröcker, D. Dohse and *R. Soltwedel* (Eds.)
Innovation Clusters
and Interregional Competition
VIII, 409 pages. 2003. ISBN 978-3-540-00999-3

D. A. Griffith
Spatial Autocorrelation and Spatial Filtering
XIV, 247 pages. 2003. ISBN 978-3-540-00932-0

J. R. Roy
Spatial Interaction Modelling
X, 239 pages. 2004. ISBN 978-3-540-20528-9

M. Beuthe, V. Himanen, A. Reggiani
and *L. Zamparini* (Eds.)
Transport Developments
and Innovations in an Evolving World
XIV, 346 pages. 2004. ISBN 978-3-540-00961-0

Y. Okuyama and *S. E. Chang* (Eds.)
Modeling Spatial
and Economic Impacts of Disasters
X, 323 pages. 2004. ISBN 978-3-540-21449-6

L. Anselin, R.J.G.M. Florax and *S. J. Rey*
Advances in Spatial Econometrics
XXII, 513 pages. 2004. ISBN 978-3-540-43729-1

R.J.G.M. Florax and *D. A. Plane* (Eds.)
Fifty Years of Regional Science
VIII, 400 pages. 2004. ISBN 978-3-540-22361-0

D. Felsenstein and *B.A. Portnov* (Eds.)
Regional Disparities in Small Countries
VI, 333 pages. 2005. ISBN 978-3-540-24303-8

A. Reggiani and *L.A. Schintler* (Eds.)
Methods and Models in Transport
and Telecommunications
XIII, 364 pages. 2005. ISBN 978-3-540-25859-9

H.W. Richardson and C.-H.C. Bae (Eds.)
Globalization and Urban Development
X, 321 pages. 2005. ISBN 978-3-540-22362-7

G. Arbia
Spatial Econometrics
XVII, 207 pages. 2006. ISBN 978-3-540-32304-4

B. Johansson, C. Karlsson,
R. Stough (Eds.)
The Emerging Digital Economy
X, 352 pages. 2006. ISBN 978-3-540-34487-2

H. Westlund
Social Capital in the Knowledge Economy
X, 212 pages. 2006. ISBN 978-3-540-35364-5

A.E. Anderssson, L. Pettersson,
U. Strömquist (Eds.)
European Metropolitan Housing Markets
VI, 363 pages. 2007.
ISBN 978-3-540-69891-3

A.T. Murray, T.H. Grubesic (Eds.)
Critical Infrastructure
VIII, 311 pages. 2007.
ISBN 978-3-540-68055-0

R. Cooper, K. Donaghy, G. Hewings (Eds.)
Globalization and Regional
Economic Modeling
XIII, 475 pages, 2007. ISBN 978-3-540-72443-8

R. Capello, R. Camagni, B. Chizzolini, U. Fratesi
(Eds.)
Modelling Regional Scenarios for the Enlarged
Europe: European Competiveness and
Global Strategies
approx. 300 pages, 2008.
ISBN 978-3-540-74736-9

Russel Cooper · Kieran Donaghy
Geoffrey Hewings
(Editors)

Globalization and Regional Economic Modeling

With 38 Figures
and 67 Tables

Springer

Professor Russel Cooper
Department of Economics
Macquarie University
Bldg. E4A, Room 410
2109, Sydney NSW
Australia
rcooper@efs.mq.edu.au

Professor Kieran Donaghy
Department of City and Regional Planning
Cornell University
106 West Sibley Hall
Ithaca, NY 14853
USA
kpd23@cornell.edu

Professor Geoffrey Hewings
Regional Economics Applications Laboratory
607 S. Matthews, # 318
Urbana, IL 61801-3671
USA
hewings@uiuc.edu

Library of Congress Control Number: 2007932627

ISSN 1430-9602

ISBN 978-3-540-72443-8 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover-design: WMX Design GmbH, Heidelberg

SPIN 12062973 88/3180YL - 5 4 3 2 1 0 Printed on acid-free paper

Preface

Contemporary philosopher John Searle has observed that while economics is a systematic and formalized science, it is not independent of context or free of history. When the social practices that economics seeks to explain change, economic theory and the models on which its explanations are based must change also. This holds for regional economics and regional economic modeling *a fortiori*. The many large and small changes that have contributed to the relatively recent acceleration of *globalization*—or the increasing integration and interdependence of the world's economies and societies—have conspired to challenge, if not completely moot, the explanatory adequacy of staple theories and models of regional economies. And while regional economists and regional scientists, more widely, have been aware of the need to, say, sharpen conceptual distinctions, modify underlying assumptions, explicitly account for new institutional arrangements, collect different data, and adapt existing or adopt new methodological approaches to modeling in response to these developments, there has been little explicit discussion of the challenges globalization poses to their field or how best to respond to them.

This state of affairs is not surprising. The process by which we advance our understanding of the workings of regional economies, the impacts on them of structural changes, and potential outcomes of policy interventions is a conservative one, framed in the 'small picture,' as befits all scientific inquiry. But complexity theorists who have begun to study regional economies also worry about the possibility of a so-called *red-queen effect*—a situation in which change in systems behavior outpaces our ability to track it, discern enduring from transitory structure, or advance our understanding of the dynamics of systems in question enough to cope well within them.

It is therefore good for those working in the field of regional economic modeling, broadly construed, to exploit available opportunities to discuss explicitly just what is transpiring in regional economies as globalization proceeds apace, how to make sense of these developments in terms of economic principles, to evaluate the adequacy of existing modeling frameworks and data sets to support study of these developments, and suggest or demonstrate methodological innovations that will facilitate modeling for purposes of testing theoretical propositions, forecasting, and conducting policy simulations and impact analyses.

The World Congress of the Regional Science Association International in Port Elizabeth, South Africa in April 2004 provided just such an opportunity. Most of the chapters contained within this volume began as papers presented at the 2004 World Congress in a track of sessions on the challenges of globalization to regional economic modeling. Chapters not originally presented as Congress papers were also invited from leading authorities on issues or modeling approaches otherwise not represented, in order to provide a more complete discussion of the volume's overarching theme. As befits contemporary work in the spatial sciences and a book concerning globalization, the contributors themselves hail from North America, Europe, Asia, Australia, and New Zealand and from the diverse disciplinary fields of economics, econometrics, environmental science, geography, logistics, operations research and management science, and regional planning.

The editors wish to thank the Editorial Board of the series of *Advances in Spatial Sciences* for their enthusiastic response to the suggestion to publish a volume on globalization and regional economic modeling. We also wish to express our gratitude to Katherina Wetzel-Vandai and Christiane Beisel from Springer for their valuable assistance in bringing the book to fruition and to several anonymous referees for their constructive criticisms and suggestions, which have improved the volume overall.

Russel Cooper, Kieran Donaghy, and Geoffrey Hewings
Sydney (Australia), Ithaca (USA), and Urbana (USA) September 2007

Contents

Part A: Advances in the Analysis of the Effects of Globalization on Regional Economies

Part B: Methodological Advances—Models of Networks

Part C: Methodological Advances—General Equilibrium Models

Part D: Methodological Advances—Econometric Models

1 Globalization and Regional Economic Modeling: Analytical and Methodological Challenges

Kieran P. Donaghy

Department of City and Regional Planning, Cornell University, Ithaca, New York, USA

1.1 Globalization and Regional Economies[1]

For the philosopher of history, G.W.F. Hegel, the fundamental challenge for any student of societal evolution is to apprehend in thought the spirit of the age (or the *zeitgeist*)—i.e., to understand the motive force of change *while it is still at work* (Lauer, 1974). Catching the *zeitgeist* 'in the act,' so to speak, is a matter of practical importance; for gaining such an understanding would seem to be a necessary, if not sufficient, condition for successfully shaping 'for the better' any future state of affairs. Hegel does not give us much cause for optimism here. He famously observed that the owl of Minerva, Roman goddess of wisdom, only spreads her wings at dusk.

It is not for want of trying that spirits of ages past have gone unapprehended (or misapprehended) by their contemporaries, as the collective works of such eminent economic historians as Adam Smith, Karl Marx, Max Weber, Emile Durkheim, Joseph Schumpeter, Arnold Toynbee, Walter Rostow, and John Kenneth Galbraith and growth theorists of more recent vintage attest. And there is no shortage of scholarship on processes of transformation now underway across the world, which may be summed up in the word 'globalization.'[2]

We may characterize globalization operationally in terms of a number of trends: the closer integration of the countries and economies of the world, greater international specialization and increased intra-industry trade, increasing trade in services, increasing integration of emerging markets into the world economy, and consolidation of production systems. Factors identified as contributing to globalization include reductions in costs of transportation and communications and bar-

---

[1] Parts of this introductory section are similar to the introductory section of Donaghy (2006).

[2] See, *inter alia*, Stiglitz (2002), Wolf (2004), and Friedmann (2005).

riers to trade. While, as many commentators have observed, globalization—*qua* global economic integration—is not new, the increasing speed of the movement of goods and services, people, capital and technology around the world, and opportunities and challenges this *velocitization of flows* presents, are.

All changes, no matter how widely distributed, are encountered locally and as globalization has progressed, regional economies have become radically transformed along a number of lines. In regions of developed economies, declining transportation and communication costs, and the fragmentation of production activities thereby enabled, have allowed firms to exploit *economies of scale* that specialization allows. Consolidation of production systems has also permitted semi-finished goods and services to be used in the intermediate stages of production across different product lines, thereby permitting the exploitation of *economies of scope*. A consequence of these developments is that production processes have become increasingly transport-intensive and metropolitan areas have become nodes in transport networks through which raw materials and goods in various stages of production pass (Castells, 2000). With increasing amounts of trade (through outsourcing) occurring *within* the same industries and most of the interactions between establishments now occurring over greater distances, multiplier effects in local and regional economies have diminished and industrial bases have been 'hollowed out.' Now, when expansion or contraction occurs at a branch plant, the largest impacts are likely to be experienced at a more centrally placed node in the network of production and distribution (Lakshmanan and Chatterjee, 2005). While, as Wolf (2004) has documented, the standard of living for the average person in many developing countries has improved from the integration of their economies into the world system, economic integration has also often contributed to the stagnation of indigenous industries and local sourcing networks in these countries, as they become sites of unskilled assembly operations (Hunter-Wade, 2005). Unless localities in such countries can move quickly up the value chain of production and establish backward linkages to domestic industries, they stand to lose newly gained branch-plant operations to other developing economies along with their indigenous industries.

Globalization in its most recent guise has challenged communities and regions to understand what is happening to them and to respond and adapt to promote sustainable development. To assist them in making sense of structural changes and to anticipate potential impacts of undertaking alternative responses, regional policy makers have in the past turned to regional economic modelers.[3] In such situations, regional economic modelers have risen to the occasion, often developing new frameworks for explaining phenomena, developing new data bases, testing theo-

---

[3] Regional economic modeling has been undertaken to test theories (causal explanations) of regional economic behavior, forecast the levels of key economic variables, analyze actual and potential impacts of events and structural changes, support planning exercises (such as determining if resources are adequate to support various undertakings), support integrated assessments of phenomena (such as climate change) that cut across social, economic, and physical systems, characterize evolutionary dynamics of regional systems, and promote understanding of linkages and interdependencies. (See, *inter alia*, Richardson (1969), Isard et al. (1998), and Schaeffer and Bukenya (2001).)

ries, forecasting outcomes, and conducting impact analyses, which have embodied new behavioral theories or implemented a new solution algorithm, and supporting difficult policy decisions. (See Isard 2003). The present volume represents a collective response by widely dispersed regional economists and regional scientists from Europe, North America, Asia and Australia to some of the analytical and methodological challenges globalization now presents.

1.2 Meeting the Analytical and Methodological Challenges

To grasp better what is transpiring we need both analytical (or conceptual) and methodological contributions. We need new schemes to organize our thinking, to direct our attention, and help us to conduct appropriate thought experiments on the basis of which guidance may be offered. And we need methodological innovations that enable us to carry out studies and thought experiments at levels of spatial and temporal resolution and formal complexity adequate to capture and account for the phenomena that characterize globalization. The chapters this volume comprises represent contributions of both types. Most of the chapters were papers invited for presentation in a track of sessions organized around the theme of 'the challenges of globalization to regional economic modelers' at the World Congress of the Regional Science Association International in Port Elizabeth, South Africa in April, 2004. The balance of this chapter introduces the chapters to follow and discusses briefly the contributions they make. While all the chapters make both analytical and methodological contributions to regional modeling, where globalization is concerned, a distinction can be made between them on the basis of where the greater contribution lies. Among those chapters whose contribution is primarily methodological, we can distinguish between those which pertain to modeling networks, those which pertain to application of general equilibrium models, and those which pertain to econometric modeling. Note that in the ensuing discussion and throughout the book, the term 'region' is used fairly elastically to refer to areas that may comprise groups of countries, sub-country states, or counties—as by convention in the field of regional science.

1.2.1 Analytical Advances

With the development and implementation of information and communications technologies, and the new commerce that they enable, some have proclaimed the death of distance. To the contrary, Philip McCann discusses in the second chapter the re-emergence of interest in geography in accounting for trade and economic growth and in particular the role of transaction costs of various sorts in accounting for changes in economic geography. Surveying both empirical developments and theoretical explanations, he discusses ways in which distance and location still matter and what regional development policies must promote to be successful. In so doing, he also identifies what regional economic models must capture to support regional economic policy making.

In the subsequent chapter, Roger Vickerman offers observations on the changing sense of what a region is, the changing nature of interregional linkages, and the implications of growing intra-industrial trade. He comments on the fluidity of developments and the overlapping nature of regional structures and contrasts different approaches to modeling regions so construed. Vickerman argues that a more dynamic notion of transportation costs is needed to appreciate the relationship between the growth of interregional trade and the cost of supporting that trade. He discusses ways in which globalization is contributing to unevenness in regional development and imbalances it is contributing to segmented labor markets within regions, hence the effects it is having on commuting patterns and migration. Vickerman observes (as does Edward Feser in Chapter 7) that policy will continue to be implemented at the regional level but raises the question of how best to adapt existing models to suit present needs.

As noted above, one of the *dei ex machina* of many accounts of globalization is information and communication technology (ICT). Cooper and Madden address in Chapter 4 the question of whether or not the benefits of ICT are being spread, or are capable of being spread, to developing countries to ameliorate the substantial economic North-South imbalance.[4] The model they develop provides a conceptual basis for addressing this question empirically and untangling various influences. Focusing on network effects that operate through increasing returns to scale in production and consumption, they develop indicators of network sophistication to determine whether or not less-developed countries (LDCs) are getting the same value from ICT innovations as developed countries (DCs). Lacking direct information on the true degree of ICT permeation, Cooper and Madden construct indirect evidence to make inferences about the existence and possible growth of a network stock, which they posit to be a source of positive consumption and production externalities. Cooper and Madden point out that the exact role of ICT is uncertain, even as there is ICT *capital deepening* (increased investment per worker). To get to their conclusions, Cooper and Madden posit rationality and consistency of static and intertemporal optimization on the part of actors involved. The challenge posed by globalization in this context is to sort out quasi-empirically what difference ICT makes in LDCs and DCs and why. Employing a 55-country data set, the authors address the question of what kind of analytical structure we need to impose on a set of constructed data in order to make meaningful inferences about various matters of interest.

While aging of the labor force in many countries isn't a feature of globalization, *per se*, it is anticipated to become a global phenomenon and an important factor that makes globalization difficult to deal with. How increasingly costly retirement and medical benefits can be supported in a time of declining industrial bases and labor forces is a quandary faced by most developed countries. In Chapter 5 Ronald McQuaid discusses some of the issues concerning the aging of the labor force that increasingly need to be considered in regional policies and model-

---

[4] The notion of a region that is operative in their analysis is a multi-national one.

ing. Focusing on OECD countries plus Brazil, Russia, India and China, he asks how aging will affect labor supply in terms of size, participation rate, and productivity. In addressing this question he observes that "Regional models should seek to fully incorporate age structures, participation rates, and related differential productivity rates, among regional and sub-regional factors, if we are to more fully understand the implications of aging upon labor markets across the globe." Turning his attention to policy implications, McQuaid frames some of the difficult choices that aging societies must face and in so doing shifts our gaze from the near term to the longer term.

One of the most prominent characteristics of globalization, commented on by several chapters in this volume, is the growth of intra-industry trade. Darla Munroe, Geoffrey Hewings, and Dong Guo examine in Chapter 6 the role of intra-industry trade in interregional trade in the Midwest of the United States. A question they raise is that if increased international trade has significant impacts on economic growth and welfare concerns, what impacts does increased trade have within countries? And how are changes in international and intra-national trade related. The authors make an important observation that trade between states is much more voluminous than foreign trade by states and introduce the notion of 'returns to trade' to support the type of interregional analysis they deem necessary. They observe that the same factors contributing to intra-industry trade internationally are also at work interregionally. A question this chapter raises is 'which developments in international trade theory best help us to understand what is transpiring and what are the implications for regional and interregional economic modeling?'

In his chapter on 'Globalization, Regional Economic Policy, and Research,' which concludes the first part of the book, Edward Feser frames the question: "In a globalizing world of scarce public sector resources, what is the appropriate subnational economic policy response?" Feser distinguishes between regional impacts of globalization and regional-level policy responses and considers two more specific questions:

1. Are there any unique implications of growing public economic integration for development planning and policy making at the regional level?
2. What kinds of spatial empirical research and model building would be most valuable to regional policy makers?

Feser discusses how economic integration today differs from in the past—e.g., services are now traded. He notes that while globalization is altering the set of locational determinants in specific sectors, the absence of empirical detail at the industrial level makes it difficult to get a fix on what is transpiring. The missing data represent a critical deficiency not only for modelers. Feser observes that

> "… the growing extent to which subnational economies are linked to one another and the global economy … will never be appreciated by policy makers until it can be demonstrated for their own regions."

Feser infers that to support sub-national policy analysis, interregional models need to be developed. But more than this, he adds

> " … adjusting out regional economic development strategies to the new world economic order will require more robust partnerships between those who would develop new tools and generate facts, those who collect the information necessary to properly understand trends and drive planning tools, and those who would put planning tools and facts to good use."

1.2.2 Methodological Advances

Models of networks. To make sense of evolving economic institutional configurations, regional economic modelers are turning increasingly to network—and *supernetwork*—constructs. (See Nagurney and Dong, 2002). The three chapters contributing advances in modeling of networks are complementary in the sense that the first focuses on aspects of optimal network design at several scales, the second on the evolution of networks of value-added flows pertaining to ICT-based industrial clusters, and the third on the co-evolution of financial and social networks.

Intermodal transport plays a key role in promoting international integration. Any policy or plan addressing such transport and associated issues results from interaction by networked public and private sector actors. In their contribution to this volume, van Duijn and van Wee address the methodological question of how to plan regional freight terminals—as central nodes of transport networks—so as to meet simultaneously policy considerations at continental, national, and subnational regional scales, and accommodate anticipated increases in traffic flows. The authors demonstrate how solutions of several constrained optimization models can provide the information needed for facility planning and how this information can be integrated in a decision support system. An implication of their analysis is that in a world of greater economic interdependence infrastructural and logistical decision-making have become interdependent and that tools to support such decision-making can be developed and implemented in more than just a heuristic fashion.

Because of the increase in specialization, there is a need to focus on the connection between industrial clustering and fragmentation of economic value chains. In their chapter on the evolution of ICT inter-cluster networks, Brian Wixted and Russel Cooper argue that

> "Continuing specialization in the division of labor would suggest that if clusters matter, then production is likely to be specifically spatially structured and connected. A study of the evolution of connections between places should therefore provide us with insights into the strengthening or weakening positions of various cluster relationships across time."

In particular, they seek to determine how supply relationships between major economies have changed with the rise of the ICT sector and Southeast Asian economies. Wixted and Cooper draw our attention to the need to revise and expand our notion of clustering and characterize clusters as nodes in production networks. A methodological contribution of their chapter is to capture the evolution of production networks—involving flows of ICT products—connecting clusters. The research reported in this chapter represents the first use of Cooper's (2000) approach to calculating net value added input-output flows from empirical data sets. The authors show that between-country linkages were universally stronger in 2000 than in 1970 and that there was an increase in the number of linkages and less dependency on a small number of linkages over the thirty-year period. This methodological contribution is important because it enables us to visualize changes in the spatial structure of interdependence—i.e., the very process of globalization. And it enables us to capture increasing component modularity and product complexity within input-output analytical frameworks.

One of the most significant developments characteristic of globalization is the growth of trade in financial services, which has both been made possible by and given rise to the evolution and emergence of international financial networks. Anna Nagurney, Jose Cruz, and Tina Wakolbinger argue that this development must be understood in the context of social networks with which they have co-evolved and are co-integrated. In their chapter, these authors adopt a network perspective for the "theoretical modeling, analysis and computation of solutions to international financial network with intermediation in which [they] explicitly integrate the social network component." The microbehavioral foundations and dynamic cast of the supernetwork model that they introduce is well suited to characterizing the evolution of service markets affecting regional economies, since it permits the co-evolution of the international financial flows, product prices, and relationship levels to be tracked over space and time. In addition to elaborating the theoretical multi-level network construct, which should prove most helpful in relating higher-level (international) to lower-level (regional) developments, the authors characterize the model's disequilibrium dynamics, discuss its solution algorithm, and demonstrate the model's applicability in several numerical examples.

General Equilibrium Models. General equilibrium models (GEMs), especially computable general equilibrium models (or CGEs), have become a stock in trade for regional economic modelers who wish to exploit well-developed theories of market behavior to impose structure in thought experiments, in which numerical answers are sought but empirical data are not plentiful, to evaluate the impacts of exogenous developments or policy interventions. Aspects of globalization that may be studied with GEMs/CGEs include increased direct foreign investment, technological spillovers, effects of climate change, movement of peoples (migration) and the emergence of the 'new (or knowledge) economy.' Four chapters in this volume illustrate how CGEs can be exploited to assess the impacts of globalization upon regional economies or examine how regional economies respond to alternative policy interventions. Whereas static frameworks tend to be the rule in the use of GEMs/CGEs, we note that all the frameworks employed by the researchers here reporting are dynamic and are exceptional in this regard.

Employing a multi-regional dynamic CGE model of Australia, John Giesecke and John Madden focus on regional adjustment to globalization and, more specifically, regional adjustment of labor markets. They demonstrate how the features of such a model are suited to characterizing various aspects of globalization and they note in particular that two possible negative impacts of globalization, increased inequality and short-term disruptions in labor markets, lend themselves to study with their model. Giesecke and Madden devote a considerable amount of attention to the details of model construction, validation, and use in alternative types of simulation. An issue that they flag for further discussion is the degree to which the assumptions of increasing returns to scale and imperfect competition—which have featured prominently in many explanations of globalization—affect outcomes of simulations with CGEs.

In their chapter, Ian Sue Wing and William Anderson take up issues of modeling *small-area* economic change in the context of a multiregional CGE model. Their work is particularly relevant because globalization has teased out economic relations—and functional regions—that do not coincide with state borders and many significant transformations cannot be adequately addressed at the state level of resolution. Wing and Anderson specify a comprehensive and rigorous framework of sub-state economic analysis that is useful for doing areal analysis of impacts of globalization. Their model overcomes limitations of demand-driven input-output models—because it includes supply-side and market dynamics—and employs state-level variable values to impose boundary conditions at the county level. For comparison with models discussed later, we note that representative firms in this model are assumed to be finitely lived and myopic, investing in fixed capital recursively, instead of being infinitely lived representative agents. The authors remark that the model may be extended to produce estimates of transportation activity levels and emissions from mobile sources and the model's demand structure can facilitate investigations of regional growth on interstate freight transportation and associated emissions.

One of the difficult questions confronting the international community is whether or not, at a time when increased development through increased international trade seems to be possible only with increased energy use, development with a small environmental footprint is possible. In their contribution to this volume, Shunli Wang and Peter Nijkamp demonstrate how a clean development mechanism (CDM) approach to environmental policy can be used to promote development while reducing emissions in a world of increasingly integrated economies. Their chapter presents analyses of CDM policies in an applied general spatial equilibrium context conducted with a variant of the Global Trade Analysis Project (GTAP) model and evaluate economic impacts of CDM policies for different regions of the world. The authors also introduce an alternative notation to characterize the stylized behaviors of actors in the model.

Certainly one of the developments most associated with globalization has been the increased industrialization of the economies of China and India. China in particular is now the world's second largest producer and consumer of energy and much of its energy supply is coal. Within the context of a multiple-period GEM, Bin Li and Yoshiro Higano, in their contribution to this volume, examine a number of scenarios in which optimal carbon taxes are applied to industrial sectors on a sector-by-sector basis to maximize GNP subject to satisfying an aggregate mitigation target. The authors find that solution taxes are 3 to 17 times current taxes on oil and coal and would be, in Li and Higano's opinion, difficult to implement. Hence market mechanisms alone would not seem to be the appropriate policy response. The authors therefore suggest that other institutional approaches to emissions reduction be pursued.

Econometric Models. Each of the last three chapters of this volume illustrates a different type of contribution that applied econometrics can make to our understanding of the implications of increasing interdependence of regional economies.

The 25-member-state European Union (EU) in its striving to attain an 'ever closer union' of European countries and regions—and the free movement of goods, services, capital and people between the same—has consciously promoted globalization. The EU is also a leading proponent of greenhouse gas emissions reduction. A relevant question to ask is what the impacts of increased trade, induced by closer integration, between newer and older EU member states will be. To investigate this question, Nazmiye Balta, Kieran Donaghy, and Clifford Wymer develop in Chapter 15 a two-bloc macro-econometric model based on the assumption of interdependent intertemporally optimizing representative agents.[5] The blocs comprise the 15 member states that constituted the EU prior to May 1, 2004 (EU15) and the 13 so-called new accession countries (NAC13) that either recently joined the EU or are candidate countries.[6] In this model growth and technological change are endogenously determined and are affected by trade positions. The continuous-time model is estimated from discrete-time observations with Wymer's nonlinear quasi-full-information-maximum-likelihood continuous-time estimator.[7] Simulations with the estimated model indicate a range of possible trade-related emission effects which depend on the direction of trade and whether changes in preferences for the other bloc's goods are unilateral or bilateral.

One of the more celebrated papers on how globalization can result in greater inequality between nations is that of Krugman and Venables (1995). In Chapter 16, Bernard Fingleton discusses the empirics of globalization and how they compare with the stylized facts discussed by Krugman and Venables. Fingleton demonstrates how global spillovers can be modeled using a quadratic reduced-form model, which is capable of manifesting the range of behaviors to which Krugman and Venables's theoretical model gives rise. Fingleton further demonstrates how, by taking into account spatial dependence of data in the estimation of the model, one may obtain a model from which much can be learned about the patterns of income convergence between national and regional economies and how shocks to individual economies will spread to others. Hence, Fingleton illustrates some of

---

[5] In essence, the model is that of a non-cooperative dynamic game.

[6] Time-series data corresponding to these two blocs, employed to estimate the model, were developed expressly for this exercise.

[7] This is the first publication of results of a macro-econometric model of intertemporally optimizing representative agents so estimated.

the potential that spatial econometrics holds for helping us understand regional implications of globalization.

Among the benefits claimed for liberalization of capital markets is that, by increasing the number of outlets for direct foreign investment (DFI), over-all investment risk will be lowered and more stable growth will be promoted.[8] There is in the empirical growth literature, however, substantial disagreement over just *how* growth and volatility are related and *what* relative weights policy makers should give to growth *vis a vis* volatility. Findings tend to be dependent upon assumptions made about underlying economic behavior (i.e., whether or not there is a 'representative consumer' maximizing expected utility) and the generality of functional forms employed in empirical work (i.e., whether or not the intertemporal elasticity of substitution and measure of relative risk aversion can be evaluated separately). In the concluding chapter of this volume, Russel Cooper and Kieran Donaghy develop the theoretical background necessary to examine the relationship between growth and volatility and its effect on DFI from the perspective of an intertemporally optimizing representative agent and then elaborate a methodology for modeling this relationship empirically, employing results in duality theory to derive estimating forms that are consistent with microeconomic foundations. They undertake a preliminary investigation using South African and OECD data and interpret their findings in terms of the theory developed.

1.3 Concluding Remarks

In concluding this introductory chapter, and on behalf of my fellow editors, I should like to express our deep gratitude to all the contributors to this volume for participating in a scholarly conversation of much practical import, advancing our theoretical and empirical understanding of the regional impacts of globalization … and helping to apprehend the *zeitgeist.*

References

Cooper RJ (2000) An extension of the block spatial path approach to analysis of the influence of intra and interregional trade on multiplier effects in general multiregional input-output models. In: Reggiani A (ed) Spatial economic science.. Springer-Verlag, Berlin, pp 303-327

Donaghy KP (2006) Market Integration, growth and inequality. In: Stewart C, Super J (eds) Unflat Global Futures. International Programs and Studies, University of Illinois, Champaign

Friedmann T (2005) The World is Flat. Farrar, Straus, and Giroux, New York

Hunter-Wade R (2005) Why free trade has costs for developing countries. Financial Times, August 11

---

[8] Other forms of risk—environmental and economic—are dealt with in other chapters of the volume.

Isard W (2003) History of regional science and the Regional Science Association International. Springer-Verlag, Berlin

Isard W, Aziz IJ, Drennan MP, Miller RE, Saltzman S, Thorbecke E (1998) Methods of interregional and regional analysis. Ashgate, Aldershot

Krugman P, Venables A (1995) Globalization and the inequality of nations. Quarterly Journal of Economics 110:857-880

Lakshmanan TR, Chatterjee LR (2005) Economic Consequences of Transport Improvements. Access 26:28-33

Lauer Q (1974) Hegel's idea of philosophy. Fordham University Press, New York

Nagurney A, Dong J (2002) Supernetworks: Decision-making for the information age. England, Cheltenham

Richardson H (1969) Elements of regional economics. Penguin Books, Baltimore

Shaeffer P, Bukenya JO (2001) What Distinguishes Regional Economics from Macroeconomics? Paper presented at the North American Meetings of the Regional Science Association, Charleston, South Carolina

Stiglitz J (2002) Globalization and its discontents. Penguin, London

Wolf M (2004) Why globalization works. Yale University Press, New Haven

Part A: Advances in the Analysis of the Effects of Globalization on Regional Economies

2 Technology, Information and the Geography of Global and Regional Trade

Philip McCann

Department of Economics, University of Reading UK and Waikato Management School, University of Waikato

2.1 Introduction to Geography and Trade

Over the last two decades there have been various analytical breakthroughs within the fields of economic growth, trade and economic geography which have forced analysts to reconsider how these phenomena are related. Within the growth literature, the work of Romer (1986, 1987) and Lucas (1988) has re-focused our attention on the role which externalities in learning-by-doing (Arrow 1962) and human capital acquisition can play in augmenting factor inputs, thereby allowing for differential shifts in the relative long-term equilibrium growth rates of different economies. Between any two economies with equivalent factor stocks, the economy which benefits from such learning externalities will tend to exhibit a relatively higher equilibrium growth rate than the economy without such learning effects. This is because the learning externalities are perceived to provide an endogenous feedback to the factor inputs which relaxes the constraints imposed by the productivity properties associated with variable factor proportions. However, the argument implicit in these models is rather more subtle than this. This is because many technological changes are seen to embody certain endogenous characteristics (Arthur 1988) which may (Lewin 2002) have implications for not only the levels, but also the patterns and characteristics, of long-term investment. There may be differences in the extent to which such endogenous process take place even between advanced OECD economies, and understanding the reasons for these differences across countries brings us to the question of the relationships between growth, trade and geography. What is it about geography and geographical trade patterns which determine the extent to which endogenous growth processes take place locally?

Since the late 1980s there has been a widespread revival of both academic and public policy interest in the links between geography, trade and economic growth. This interest is not confined to any particular part of the world, although the major emphasis of these discussions has tended to take place among OECD countries. There are several reasons for the recent renewed interest in the role which geography plays in determining economic growth; one reason is technological, a second reason is institutional, and a third reason is analytical.

The primary technological development which has contributed to the renewed interest in the economic impacts of geography, has been the rapid improvement in information, communications and transportation technologies. These technological advances have improved the ability of corporate and government decision-makers to coordinate either market or organizational activities across progressively larger geographical areas. This is because the new technologies provide for the better planning and control of activities across multiple locations, resulting in an improved ability to exploit intra-marginal differences in international and inter-regional rates of return. It is not clear, however, whether these developments will alter the spatial distribution of economic benefits on a global basis, in comparison with the existing patterns determined by previous technological regimes. Yet, where any such changes do actually occur, these changes will be generated by changes in the geographical patterns of trade and growth.

At the same time as these technological changes have taken place, there have also been widespread institutional changes within the global and regional trade frameworks. The movements towards free-trade and integrated market areas such as EU, NAFTA, ASEAN and MERCOSUR, have meant that the tariff structures associated with national borders may be becoming progressively less important in terms of their effects in shaping a nation's economic performance. In particular, reduced trade barriers may lead to both quantitative and qualitative changes in the spatial patterns of investment both within and between countries. Any such changes may lead to differential growth impacts between different geographical areas, and once again, such issues require us to ask questions about the relationship between geography, trade and growth.

The combination of these technological changes and institutional changes has encouraged widespread discussions about the supposed economic and social impacts of globalization on the gap between rich and poor countries. Of particular interest to regional scientists and economic geographers, however, is whether any changes in the spatial patterns of trade and growth, associated with either changes in communications technology or trade barriers, will tend to favor geographically central or peripheral countries, irrespective of their levels of development.

In these discussions, the question of whether or not there are any adverse consequences associated with geographical peripherality, depends primarily on whether or not economic integration is seen as a universal equilibrating growth mechanism. Evidence from common trade areas suggests that lower tariffs, trade barriers and communication tend to favor low wage peripheral economies, primarily via inflows of capital. The results of this equilibrating process imply a convergence in incomes across spatially differentiated markets (Barro and Sala-i-Martin 1992). Yet the robustness of these convergence observations appears to be very dependent both on the time scales of analysis (Fingleton and McCombie 1999; Armstrong 1997) and also on the individual spatial units chosen for the analysis (Cheshire and Carbonaro 1995). On the other hand, there is much analytical evidence to suggest that the growth effects of continuing economic integration may be quite different between different areas. In particular the work of two key commentators, Krugman (1991) and Porter (1990), has opened up discussions of the role which geography plays in economics and business matters, to a much wider academic and policy-making audience than was previously the case. The work of Krugman (1991) has lead to the development of the 'new economic geography' literature, which argues that the uneven distribution of industrial activities across space is a natural result of market processes. Meanwhile the work of Porter (1990) has fostered the literature promoting the importance of industrial 'clusters' in promoting industrial innovation. The primary lessons from these two literatures are that geography really does matter in determining economic performance, and geographic peripherality can indeed have adverse consequences. In particular, there are strong reasons to expect systematic growth advantages accruing to central areas over geographically peripheral regions (Krugman and Venables 1990) in which there are concentrations or 'clusters' of industrial activity (Porter 1990).

Such new economic geography and industrial clustering arguments may be of real concern to geographically peripheral regions or countries, because they imply that many of the low costs comparative advantages they may previously have exhibited will tend to become continuously eroded relative to other areas. However, the validity of these various arguments and conclusions depends largely on the specific assumptions we make concerning the characteristics which are ascribed to geographical transactions costs. Information communications costs, transportation costs and institutional tariff barriers, can all be considered to be just different forms of market transactions costs. Yet, each of these various types of transactions costs is explicitly geographical both in nature and impact. Any changes in the levels or structure of these spatial transactions in any particular geographical region, will have profound impacts for the patterns of international and interregional trade in that region, and also between that region and any other region. Therefore, in order to understand the possible economic growth impacts of possible changes in the international and interregional transactions costs faced by firms in more geographically peripheral regions, relative to more centrally located firms, it is necessary to consider both the nature of these transactions costs and also the nature of the changes in these transactions costs over recent years.

2.2 Spatial Transactions Costs

Apart from institutional tariff barriers, the spatial transactions costs faced by firms are primarily of two types: transportation costs and information transmission costs. Both of these costs will also be incorporated in retail, distribution costs, and current evidence suggest that for international trade costs at least, these costs are still very significant indeed (Anderson and Van Wincoop 2004). However, there are also arguments which suggest that trade costs have also changed over recent years due to technological changes. In this section we will therefore review the developments and changes in each of these two types of transactions costs in order to understand the transactions costs' environment faced by firms which are competing in regional and international markets.

2.2.1 Information Costs

Since the 1980s we have seen dramatic improvements in the ability of decision-makers and planners to coordinate activities across space. The primary reasons for these improvements have been the enormous technological developments in information technology, and also the advent of widespread usage of these technologies. These developments have meant that complex operations can now be managed both more efficiently and effectively than was previously possible. There are two aspects to these developments.

First, the new information technologies have reduced the real costs of communicating across distance, allowing us to more efficiently control existing spatial arrangements of activities (*The Economist* 1999b). This is a common observation in industrial sectors and activities where physical commodities are being moved across large distances, such as in the management of international importing and exporting supply chains (*Financial Times* 1999) or the coordination of multinational manufacturing activities (*The Economist* 1999a). Analogous arguments also exist for the case of the service sectors, in situations where information rather than physical goods is being transferred across space. In many situations, information technologies employing satellite and fiber-optical technology allow for greater quantities of information to be transmitted at much lower costs than was previously possible.

Secondly, the existence of these new information technologies also allows decision-makers to undertake the coordination of spatial arrangements of activities which were previously not possible. This is evident in examples such as international accounting, where New York banks transfer their book-keeping requirements overnight to firms in Dublin, in order to have them updated in time for the opening of the money markets the next day. Other examples include Silicon Valley firms which subcontract software development activities to firms in Bangalore India, while still maintaining daily contact and control of the Indian software development process from California. Meanwhile, for service industries such as finance and marketing, the new possibilities provided by information technologies for the supply of information-based services across global space appear almost unlimited (*The Economist* 1999a).

The reductions in the real costs of transmitting information across space, which are associated with these new information and communications technologies, would suggest that geographical peripherality is becoming relatively less of a handicap to accessing international markets. On the other hand, however, there are some other arguments which suggest that over time the development of these information technologies is actually leading to increases in the costs of transmitting information across space, thereby increasing the relative importance of geographical centrality. The argument here is that an increase in the quantity, variety and complexity of information produced, itself increases the costs associated with transmitting this information across space. This is because much of the information will be of a non-standardized tacit nature, and the transmission of this type of information essentially requires face-to-face contact (Gaspar and Glaeser 1997). The opportunity costs involved in not having face-to-face contact will consequently increase with the quantity, variety and complexity of the information produced. The effects of this will be to increase the costs of doing business across large geographical distances. As such, these arguments would suggest that geographical peripherality may become progressively more of a handicap to business growth due to the increased relative costs of distance.

2.2.2 Transportation Costs

As we suggested at the beginning of this chapter, transportation technologies have improved dramatically over recent years. Obvious examples of this include the growth in roll-on roll-off trucking, containerization, rapid-turnaround shipping, and the increased efficiency and frequency of airline services. Moreover, to the extent that modern transportation technology is able to achieve ever more significant economies of scale and distance[1], as with information transmission, the reductions in the real costs of transporting goods associated with these new technologies would suggest that geographical peripherality is becoming relatively less of a handicap to accessing international markets. Once again, such observations appear to imply that any adverse competitive effects associated with geographical peripherality will have fallen over time.

On the other hand, the quantity, variety and complexity of market information generated in the modern economy are increasing. This also implies that in many industries which involve the production or shipping of goods across space, the variety and complexity of the logistics operations being undertaken will also increase. The reason for this is that as modern consumer demand requirements become more sophisticated, there is an increasing preference for goods shipments characterized by speed, reliability and timeliness. In other words, the consumer's opportunity costs of time have also increased for goods shipments. Modern household and industrial consumers now require a level of service customization and delivery speed, which previously was not considered either so important or even possible. As the demand for delivery speed increases, the associated opportunity costs of lead-times also increase, and the average inventory levels maintained will fall. The effects of this on distance costs can be explained by adopting a similar argument to that employed above. For any two agents at a given distance apart, the optimized delivery frequency increases as the opportunity costs of time

---

[1] Transportation costs generally exhibit both economies of scale and economies of distance simultaneously. The reason for this is that these costs structures are the natural results of the frequency-cost trade offs faced by all hauliers, the (envelope) results of which turns out to be a non-linear square root function of all cost variables (McCann 1993, 1998, 2000)

increase.[2] Analytically, the effect of this is to increase the transactions costs associated with shipping goods over any given distance. The spatial outcome of this argument is that, ceteris paribus, agents will move closer to each other as the variety and complexity of information increases, in order to reduce the resulting opportunity costs of distance.

The most extreme example of this trend towards more frequent shipments, is the application of Just-In-Time (JIT) manufacturing and distribution techniques, the influence of which has pervaded all areas of modern production, distribution and retailing. New information technologies allow firms to coordinate logistics activities across huge geographical areas in a very sophisticated and timely manner. In the new JIT production and distribution arrangements (Nishiguchi 1994; Schonberger 1996), it is necessary to control the flows of goods between firms to a very high degree, in order to ensure the timeliness of deliveries. The ability to track and monitor the speed of movements of goods therefore becomes essential, particularly if the goods are being shipped over significant distances. Similar arguments also hold for the case of customized high-speed mail services. Yet, these technological developments have also lead to a change in consumer behavior. Both household and industrial consumers now expect goods to be delivered JIT. As such, the nature of demand for transactions across space has changed dramatically. Customers require much shorter lead-times than was previously possible. As such, there is a direct parallel with the argument regarding information costs, only in this case, the opportunity time-costs of goods shipments are tied up in the levels of inventory being held, rather than the opportunity time-costs of not having face-to-face contact.

There is a range of empirical evidence which suggests that the spatial transaction costs involved in shipping of goods have indeed increased over the last two decades, because of this demand for more frequent deliveries. First, the average inventory levels for almost all manufacturing and distribution sectors in the developed world have fallen dramatically since the 1980s, relative to the value of output (Shonberger 1996; *Financial Times* 1998). This implies that the average lead times of goods-shipments have fallen over recent years, with a concomitant increase in goods-shipment frequencies. Secondly, although the level of transport costs as a proportion of global GDP has fallen over time (Glaeser 1998), by carefully disentangling the various components of transport costs it becomes clear that the proportion of global output which is accounted for by the combination of logistics and transportation activities in the economy has not fallen over recent decades for all sectors (Hummels 1999; *Financial Times* 1997). While the transportation cost component of bulk materials has indeed generally fallen, in the case of manufactured goods, there is evidence that this proportion has actually increased over the recent decades, in spite of the improvement in transportation and logistics technologies (Hummels 1999). Thirdly, industries which are very dependent on JIT shipments have tended to reorganize their trade patterns in favor of geographi-

---

[2] As with the case of transportation costs discussed above, the (envelope) result turns out to be a non-linear square root function of all the cost variables (McCann 1993, 1998, 2000).

cally close suppliers and customers (Reid 1995; McCann 1998). Moreover, this behavior is even evident in industries in which the product value-weight ratios are extremely high (McCann and Fingleton 1996). In other words, such localization behavior is present in the very industries which traditional Ricardian trade theories would have ruled out.

2.2.3 Changes in International Transactions Costs

The preceding sections provide a range of arguments and evidence which suggest that the real costs involved in transacting information and goods across space have both decreased and increased over recent decades. However, these apparently conflicting conclusions can be reconciled in that the different types of changes in transactions costs described above have tended to take place in different types of sectors and activities. On taking a broad view of the issues, it appears that most of the evidence points to falling international and geographical transactions costs for existing types of activities. The sectors in which spatial transactions costs have indeed fallen significantly over recent decades, are generally the sectors in which the nature of the spatial transactions undertaken have not changed fundamentally over time, in terms of the required frequency of interaction. This is typically the case in many raw material, agricultural or extraction industries, and in industries producing manufactured products at a mature stage within their product cycles (Vernon 1966). This is also the case in service sector industries in which the nature of the information being transacted is rather standardized, such as retail banking. In other words, if we apply a ceteris paribus criterion to the case where the nature and characteristics of the transactions have not changed, then international transport and transactions costs can be unambiguously assumed to have fallen steadily over time. In these cases, geographical peripherality would appear to be less of a disadvantage than it might have been previously.

On the other hand, in production sectors in which the demand lead-times have fallen dramatically, or in industries in which the variety and complexity of information generated has increased significantly, spatial transactions costs would appear not to have fallen over recent decades, and in some cases will actually have increased. Where such costs may have risen over time, it appears that this is a result of the fact that the nature and characteristics of such transactions have changed, thereby violating the ceteris paribus criterion. In these cases, the requirement for geographic proximity would appear to have increased, and the potential disadvantages of geographical peripherality would appear to have increased.

2.3 International Geographical Peripherality and Competitive Advantage

From the perspectives of both trade and growth, the arguments outlined above implying falling international and spatial transactions costs provide encouraging possibilities for geographically peripheral areas. The reason for this is that falling international transactions costs reduce the wedge between origin and delivered prices, thereby allowing geographically peripheral economies more efficient access to international markets, both in terms of production and consumption. Firms located in peripheral regions will be better able to compete internationally because lower transactions costs will allow them to benefit to a greater extent from the comparative advantage provided by the relatively low domestic factor prices.

In terms of economic growth, the effects of any reductions in international transactions costs depend First, on whether such reductions are stepwise or continuous, and secondly on the existence of economies of scale. If any reductions in international transactions costs are simply a stepwise, once-and-for-all phenomenon, geographically peripheral economies will not be expected to experience growth effects which are consistently different from more centrally located areas. On the other hand, if reductions in international transactions costs are broadly a continuous phenomenon, as would be expected with steady technological progress, geographically peripheral areas would be expected to consistently generate economic growth levels which above those of geographically central economies. This resulting strong growth performance would then encourage the inflow of production factors into geographically peripheral areas seeking higher factor rewards, which itself will encourage further growth. This argument is the basis of the Borts and Stein (1964) and Barro and Sala-i-Martin (1992) convergence models, which were initially applied to the processes of economic integration across the large geographical areas of the USA and EU, respectively. However, orthodox neo-classical models of economic growth, factor allocation and trade, assume that these convergence processes are more generally applicable to an ever-increasingly integrated world. Consequently, these arguments imply that geographically peripheral economies will experience a relatively high growth performance as we face steady reductions in international transaction costs. Obviously, the converse arguments also hold in situations where we face steadily increasing international transactions costs. However, given that most evidence tends to suggest that international transactions costs are falling steadily over time, these arguments provide geographically peripheral regions and countries with many reasons to be optimistic.

These generally optimistic observations associated with falls in international transactions costs hold as long as the aggregate production functions of geographically peripheral regions and areas experience largely constant returns to scale. However, the new international trade (Helpman and Krugman 1985; Krugman and Venables 1990) and new economic geography literature (Fujita et al. 1999) suggests that the spatial patterns of economic growth will be quite different, depending on the extent to which varying levels of economies of scale are operative in different locations. The new international trade and new economic geography literature suggests that if individual economies experience economies of scale, falling international transactions costs will tend to benefit the larger and more centrally-located economies, at the expense of the geographically peripheral economies. The reason for this is that these models assume that market size and centrality provides for significant economies scale in both consumption and production, due to the presence of greater product and input variety as well as greater industry diversity. Under these conditions, high international transactions costs act in a manner which is analogous to that of high trade tariffs, in which the peripheral economies are protected from the external competitive pressures of the larger more central economies. In such protected situations, domestic producers are allowed to continue in business, because the high transactions costs and trade barriers rule out the competitive advantages of the larger or more centrally-located external producers. On this type of argument, generally falling international transactions costs will not be advantageous to geographically peripheral regions and countries, because firms located in these regions will become progressively more open to competition from competitors in other regions (Krugman 1996). As such, although the global trading system as a whole will benefit from such falls, the relative distribution of such benefits will not favor peripheral regions, unless these regions can themselves generate significant economies of scale. These arguments provide geographically peripheral regions and countries with many reasons to be rather pessimistic.

In the face of generally falling international transactions costs, the key question therefore raised by these new international trade and new economic geography arguments is, are geographically peripheral regions and countries more or less likely to exhibit economies of scale. Alternatively, in situations where international transactions costs have increased, are the geographically peripheral regions and countries in a position to take advantage of the natural geographical concentration effects of such cost increases? To answer these questions it is necessary to discuss economic growth behavior at much smaller geographical scales and dimensions than are implied by the international trading system; namely that of the scale of the individual country and more particularly, at the scale of the individual city-region economy within the individual country.

2.4 Agglomeration Economies and Economic Growth

To what extent are geographically peripheral regions and countries able to generate sufficient domestic economies of scale in order to compensate for the reduced domestic trade protection effects associated with falling international transactions costs? In attempting to answer this question, it is first necessary to consider the underlying factors which determine the not only the generation of economies of scale, but also the uneven spatial distribution of such scale economies.

The current thinking on these issues generally revolves around the notion of industrial clustering and the associated potential benefits of external agglomeration economies. The existence of domestic agglomeration economies within a country are perceived to allow for a more rapid economic growth on the part of the country as a whole. Here the arguments tend to focus on the role which geographical proximity can play in the fostering, facilitating and nurturing of flows of inter-firm information which then allow for the local generation of mutually beneficial information externalities. This kind of logic underlies each of Alfred Marshall's (1920) three explanations for the existence of positive agglomeration externalities in situations of urban industrial clustering.

Marshall's first observation concerned the existence of 'informal' information spillovers, where informal refers to the fact that they are non-traded information spillovers between agents, primarily of a tacit nature. Such informal and tacit information spillovers can take place between geographically proximate agents, in cases where all the agents are firms, or where some of the agents are units of labor. Marshall's assumption is that information spillovers operate specifically at the level of the individual urban area, and it is over this spatial extent that transactions costs are assumed to become critical. In other words, from the point of view of information transactions, it is the geographical scale of the individual urban area which is critical in terms of determining economic performance. This is also the particular spatial logic which has been adopted by the 'new economic geography' models of Krugman (1991) and Fujita et al. (1999).

Marshall's second explanation for local external economies arises due to the presence of non-traded specialist local input providers, who find the investment in such input provisions profitable in situation where they are servicing locations of clustered producers of a similar sector. Once again, the validity of this argument depends on the availability of local information allowing for not only the provision, but also the efficient consumption of these specialist inputs.

The third argument of Marshall in favor of the existence of local external economies is based on the fact industrial clustering permits the rise of specialist pools of skilled labor. Here, geographical proximity allows not only for a more efficient search and matching process within the labor market, but also an easier adjustment to adverse shocks within the local labor market, as long as the shocks are not correlated across sectors (Mills 1972). As such, Marshall's observations suggest that industrial clustering better allows both firms and workers to reduce the downside risk costs associated with investment in any particular capital technology, whether physical or human. Both net returns and profit growth will be maximized because the industrial clustering itself provides a mechanism for circumventing many features of market failure which are endemic in real-world markets. This appears to be particularly so for complex inter-firm production arrangements involving many small firms.

The Marshall arguments outlined here provide possible explanations for the scale economy and efficiency benefits of industrial clustering. However, it is still not entirely clear why we should be concerned by these arguments. Just because there has been a recent increase in the perceived importance of these agglomeration phenomena as potential determinants of economic growth does not necessarily mean that there is any substantive change to the competitive conditions faced by geographically peripheral regions. As we have already seen, there have been widespread technological and institutional changes which appear to have largely reduced many aspects of spatial transactions costs, thereby potentially benefiting peripheral economies. Similarly, large cities and industrial clusters have been a longstanding feature of our economic system, so why should there be a recent focus of interest on these questions?

In response to these arguments, Glaeser (1998) argues that if we consider the changes in the transactions costs of goods-shipments alone, then the rationale for industrial clustering and the existence of modern cities disappears. On the other hand, he argues that the transportation costs involved in ensuring that people have regular face-to-face contact, is the crucial driving force behind the generation of modern cities and industrial clusters. In other words, the overcoming of increased modern information transactions costs appears to be the primary theoretical rationale underlying the existence of modern cities. Identifying this empirically, however, and in particular identifying the critical spatial extent which defines whether a location is advantageous or not is very difficult (Glaeser et al. 1992; Henderson et al. 1995). Indirect methods have to be employed, such as observing the spatial patterns of patent citations (Jaffe et al. 1993; Acs 2002). These empirical techniques tend to confirm the argument that many aspects information spillovers are constrained primarily within the individual urban area, thereby implying that the urban area is the critical geographical range of advantage for localized economies of scale.

In addition, there are two other sources of evidence which support the argument that spatial information transactions costs have increased over recent decades, thereby increasing the importance of the urban area as the potential source of economies of scale. The first source of evidence comes from observations of telephone usage patterns (Gaspar and Glaeser 1998). Using data from Japan and the US they observe the relationship between the density and frequency of telephone usage and the location of the users. First, they find that users who are geographically closer together, and for whom greater face-to-face contact is therefore easier, spend more time talking to each other on the telephone, than do users who are at greater distances from each other. Secondly, the same result also holds for urban size, in that users in larger urban areas talk to each other relatively more frequently than users in smaller urban centers. Thirdly, the frequency of airline business travel has increased more or less in line with the growth in telecommunications usage.

The second source of evidence suggesting that the individual urban area has become progressively more important as a source of economies of scale involves an assessment of the rates of global urbanization. Over the last three decades, the proportion of people living in urban areas has increased in all parts of both the developed and developing world (United Nations 1997). While the reasons for this are complex, and particularly in relation to the out-migration of labor from rural areas in developing economies, the ubiquitous urbanization phenomenon in the developed parts of the world where information technologies are mostly applied, also suggests that the geographical proximity of firms and people within individual urban areas is becoming relatively more important over time.

The implication of these empirical observations is that the individual urban industrial area is, if anything, becoming even more important nowadays as a determinant of domestic scale economies of than it was previously. The reason for this is that while international transactions costs are generally decreasing, the (opportunity) costs of the spatial transactions contained within individual countries are actually increasing. This is because information and communications technologies and face-to-face contact, are not necessarily substitutes for each other, but are often complements for each other. In other words, a general increased usage of information and communications technologies often leads to an increase in the quantity, variety and complexity of goods and services produced, which itself leads to an increase in spatial information transactions costs, and an associated increased need for spatial proximity to facilitate face-to-face contact. At the same time, an increase in the levels of spatial proximity encourages a greater usage of information and communications technologies, and the production of more varied and complex information, such that the process becomes cumulative. Glaeser's arguments (Glaeser 1998; Gaspar and Glaeser 1998) therefore suggest that in the modern world, the Marshallian foundations of agglomeration externalities are becoming an ever-more significant determinant of domestic economies of scale. The agglomeration arguments of Glaeser and Krugman therefore provide grounds for serious concern on the part of geographically peripheral regions and countries, because current changes in spatial transactions costs appear to lead to a process which tends to favor concentration and centralization of many activities within key city-regions.

2.5 Alternative Models of Industrial Clusters

While the agglomeration arguments of Glaeser and Krugman imply that there are increasingly inherent disadvantages associated with geographical peripherality, there are other models of industrial clustering and growth, which are rather more circumspect in terms of their perception of the critical spatial extent of information transactions, externalities and growth. While the new economic geography models of Krugman (1991) and the urban agglomeration models of Glaeser (1998) are based on the assumption that the individual urban area is the critical spatial extent which defines geographic advantage or disadvantage in growth performances, two other types of clustering-interaction models suggest that growth mechanisms may take place over rather different spatial and population scales, thereby providing some opportunities for optimism on the part of geographically peripheral regions. These two other models are the '*industrial complex model*' and the '*social network model*', and they suggest that simple observations of the scale of urban population levels and industrial clustering will not necessarily be instructive as to the nature of localized growth mechanisms. In order to understand how the insights of these two additional models of clustering may be interpreted, we will first explain their particular foundations and transactions-costs characteristics in direct comparison to the agglomeration model outlined above.

In order to do this, we can adopt a transactions costs approach to present three stylized sets of geography-firm-industry organizational relationships (McCann and

Gordon 2000; McCann 2001b; McCann and Sheppard 2003, McCann and Shefer 2004; Simmie and Sennet 1999). The three stylized characterizations of industrial clusters are distinguished in terms of the nature of firms in the clusters, the nature of their relations, and transactions undertaken within the clusters. These three distinct types of industrial clusters can be termed the *pure agglomeration,* the *industrial complex*, and the *social network*. In reality, all spatial clusters or industrial concentrations will contain characteristics of one or more of these ideal types, although one type will tend to be dominant in each cluster. The characteristics of each of the cluster types are listed in table 2.1, and as we see, the three ideal types of clusters are all quite different.

Table 2.1 Industrial Clusters

Characteristics	pure agglomeration	industrial complex	social network
firm size	atomistic	some firms are large	variable
characteristics of relations	non-identifiable fragmented unstable frequent trading	identifiable stable and frequent trading	trust loyalty joint lobbying joint ventures non-opportunistic
membership	open	closed	partially open
access to cluster	rental payments location necessary	internal investment location necessary	history experience location necessary but not sufficient
space outcomes	rent appreciation	no effect on rents	partial rental capitalization
example of cluster	competitive urban economy	steel or chemicals production complex	new industrial areas
analytical approaches	models of pure agglomeration	location-production theory input-output analysis	social network theory (Granovetter)
notion of space	urban	local or regional but not urban	local or regional but not urban

First, in the model of pure agglomeration, inter-firm relations are inherently transient. Firms are essentially atomistic, in the sense of having no market power, and they will continuously change their relations with other firms and customers in response to market arbitrage opportunities, thereby leading to intense local competition. As such, there is no loyalty between firms, nor are any particular relations long-term. The external benefits of clustering accrue to all local firms simply by reason of their local presence. The cost of membership of this cluster is simply the local real estate market rent. There are no free riders, access to the cluster is open, and consequently it is the growth in the local real estate rents which is the indicator of the cluster's performance. This idealized type is best represented by the Marshall (1920) model of agglomeration, as adopted by the new economic geography models (Krugman 1991; Fujita *et al.* 1999). The notion of space in these models is essentially urban space, in that this type of clustering only exists within individual cities.

Secondly, the industrial complex is characterized primarily by long-term stable and predictable relations between the firms in the cluster, involving frequent transactions. This type of cluster is most commonly observed in industries such a steel and chemicals, and is the type of spatial cluster typically discussed by classical (Weber 1909) and neo-classical (Moses 1958) location-production models, representing a fusion of locational analysis with input-output analysis (Isard and Kuenne 1953). Component firms within the spatial grouping each undertake significant long term investments, particularly in terms of physical capital and local real estate, in order to become part of the grouping. Access to the group is therefore severely restricted both by high entry and exit costs, and the rationale for spatial clustering in these types of industries is that proximity is required primarily in order to minimize inter-firm transport transactions costs. Rental appreciation is not a feature of the cluster, because the land which has already been purchased by the firms is not for sale. The notion of space in the industrial complex is local, but not necessarily urban, and may extend across a sub-national regional level. In other words, these types of complexes can exist either within or far beyond the boundaries of an individual city, and depend crucially on transportation costs.

The third type of spatial industrial cluster is the social network model. This is associated primarily with the work of Granovetter (1973), and is a response to the hierarchies model of Williamson (1975). The social network model argues that mutual trust relations between key decision making agents in different organizations may be at least as important as decision-making hierarchies within individual organizations. These trust relations will be manifested by a variety of features, such as joint lobbying, joint ventures, informal alliances, and reciprocal arrangements regarding trading relationships. However, the central feature of such trust relations is an absence of opportunism, in that individual firms will not fear reprisals after any reorganization of inter-firm relations. Trust relations between key decision-makers in different firms are assumed to reduce inter-firm transactions costs, because when such trust-based relations exist, firms do not face the problems of opportunism. As such, these trust relations circumvent many of the information issues raised by the markets and hierarchies dichotomy (Williamson 1975). Where such relations exist, the predictability associated with mutual non-opportunistic trust relations, can therefore partially substitute for the disadvantages associated geographic peripherality. Inter-firm cooperative relations may therefore differ significantly from the organizational boundaries associated with individual firms, and these relations may be continually reconstituted. All of these behavioral features rely on a common culture of mutual trust, the development of which depends largely on a shared history and experience of the decision-making agents.

This social network model is essentially aspatial, but from the point of view of geography, it can be argued that spatial proximity will tend to foster such trust relations over a long time-period, thereby leading to a local business environment of confidence, risk-taking and cooperation. Spatial proximity is thus necessary, but not sufficient to acquire access to the network. As such, membership of the network is only partially open, in that local rental payments will not guarantee access, although they will improve the chances of access. In this social network model space is therefore once again local, as with the complex, but not necessarily urban, and often extends over a sub-national regional level. Once again, in this case, both information transactions costs and transportation costs may play a role in determining the importance of geographical peripherality.

The major geographical manifestation of the social network is the so-called ‘new industrial areas’ model (Scott 1988), which has been used to describe the characteristics and long-term growth performance of areas such as the Emilia-Romagna region of Italy (Piore and Sabel 1984; Scott 1988). This region has large networks of primarily small firms which are tied together by close personal ties. The trust networks evident between the firms allow the firms to arrange co-operative syndicates for certain types of activities, such that longer-term and more comprehensive investment programs can be undertaken by the small firms than would be the case in an orthodox market mechanism. The result has been a continuous upgrading in the technology of the firms from traditional craft-based leather-goods activities to currently very high levels of technological inputs.

Meanwhile, the clustering model of Porter (1990, 1998) can also be argued to fit into this social network category. Although Porter assumes that the dominant competitive effects of clustering are mediated by information flows between firms and individuals within the urban sphere, the primary effect of which is to stimulate local competition by increasing the transparency associated with competitive improvements, he also acknowledges that such information flows may also extend well beyond the urban scale in situations where trust exists.

Both the industrial clustering model of Porter (1990, 1998) and the ‘new industrial areas’ model of Scott (1988), are therefore much less specific than the urban agglomeration about the particular spatial dimension which is critical in terms of information transactions costs. In cases where there are small-firm industrial structures, the spatial extent over which such trust relations operate will tend to be over small sub-national regional scales (Scott 1988; Porter 1990). On the other hand, in industrial structures characterized by large vertically-integrated firms, such trust relations may operate over much larger regional spatial scales, and in the case of contiguous small-area nations, these regional scales may extend beyond the individual country boundaries (Casson and McCann 1999). Where industrial structures are characterized by both small and large firm networks, such long-term trust relations can exist over national spatial scales.

There is some empirical evidence which supports these various arguments. Observations of the formal inter-firm outcomes of informal information exchanges (Arita and McCann 2000; Audresch and Feldman 1996; Suarez-Villa and Walrod 1997), technology spillovers (Cantwell and Iammarino 2000) or the spatial patterns of joint-lobbying activities (Bennett 1998), suggest that the spatial extent of such long-term inter-firm networks may be much greater than that of a single city, and may extend across whole national or sub-national regional areas. These various arguments suggest that the critical spatial areas which define geographic growth advantage or disadvantage, may be far larger than any of the Marshall, Glaeser or Krugman arguments imply.

2.6 Economic Geography and Public Policy

The orthodox agglomeration arguments (Marshall 1920; Krugman 1991; Glaeser 1998) outlined above appear to raise serious points of concern for the long-term growth performance of geographically peripheral areas. In the new interregional and international environment it appears that the types of industries which are currently the most innovative appear to be those which require the most face-to-face contact. This favors the centrally located countries. In addition, geographically central regions and countries tend to be characterized by relatively large populations and population densities, and this density of activities points towards spatial clustering characteristics typical of the pure agglomeration model. As such, market outcomes lead to both spatial concentration as well as local employment and production variety, and these areas of concentration also appear to be the areas generating the most innovations (Acs 2002). On the other hand, peripheral regions and countries appear to be automatically at a disadvantage in terms of the local promotion of such clustering possibilities. As such, where local economic development policies are employed in these regions, the focus of these policies must necessarily be on the promotion of industrial clusters, either of a type consistent with the industrial complex model or with the social network model.

In peripheral regions, traditional regional development policies which employ location incentives for primarily manufacturing firms, such as local rental or local tax reductions, have tended to foster the promotion of local industrial complexes (Isard *et al.* 1959). From the arguments above, for these policies to be successful in the long term, it is necessary that the re-located firms develop mutual long-term, stable and predictable local relations with the other local firms. Otherwise, without the development of such local relations, once the location incentive has expired, the firms will have little reason to remain in the locality. The reason for this is that the fact that an incentive was required in the first place suggests that the location is not optimal for the firm. Alternatively, following social network argument, even if local purchasing and input-output relations are not developed by the firms according to the industrial complex model, such re-location policies may still be successful if the re-located firms tend to develop long-term trust based relationships with other firms, organizations and institutions in the local area. The development of either long-term stable and predictable trading relationships, or alternatively local mutual trust-based relationships may increase the extent to which such firms become ‘embedded’ in the local economy. If such a phenomenon takes place, then the policies will have been successful. However, linking industrial re-location policies to the development of trust-based relations appears to be extremely difficult, because by definition, trust cannot be legislated for or enforced. For this reason, following the apparent success of some of the Italian industrial districts (Becattini 2004), an alternative approach which is now increasingly popular is for regional development agencies to try to promote the development of local trust based relations (Johansson et al 2002) among indigenous firms, through increased inter-firm networking and collaboration. Moreover, the emphasis here is often also on the promotion of high technology and information intensive sectors. However, these appear to be the very sectors for which geographic centrality and face-to-face contact are the most important tissues. Therefore, whether such new approaches to regional policies are successful remains to be seen, as this line of thinking is relatively recent.

2.7 Conclusions

In the post-war Bretton-Woods era, characterized by largely closed economies, restricted factor flows, and bilaterally-regulated trading arrangements, geography was regarded as being largely irrelevant by many economists, because Ricardian theories of comparative advantage appeared quite sufficient to explain observed trade behavior. However, over the last two or three decades the international situation appears to have changed dramatically. Changes in technology and institutional structures mean that geography is becoming ever-more crucial as a determinant of economic growth.

Glaeser (1998) argues that taking a broad view of all the empirical evidence indicates that the aggregate share of total output accounted for by transportation costs has fallen markedly over time. Therefore, if we follow the straightforward technological arguments outlined above which suggest that both information transmission costs and international transportation costs have fallen over time, we could also conclude that geographical peripherality is becoming much less of a competitive disadvantage for geographically peripheral regions and countries accessing international markets than it might have been previously. This is because the supply of activities, goods or services will become progressively cheaper and easier over greater spatial scales, due to better management and delivery possibilities provided by the new transport and communications technologies. These reduced costs of doing business over large geographical distances also imply that the range of activities supplied across all spatial areas will tend to converge. The reason for this is that a general reduction in spatial transactions costs will reduce any missing markets associated with transactions costs inefficiencies. As such, the advent of these new transport and communication technologies suggests that international differences in geographical location would appear to become successively less important over time in determining the range of products and activities available to any particular country. Some authors have even assumed that eventually this would lead to the death of geography as an issue in its own right (Toffler 1980; Naisbitt 1995).

On the other hand, however, as we have seen here, there are also arguments which suggest that the development of the information technologies themselves is actually leading to increases in the costs of doing business across space. Information technology alone obviously reduces the costs associated with transmitting particular quantities and types of information across space. However, an additional aspect of these technologies is that they also lead tend to lead to an increase in the quantity, the variety, and the complexity of the information and goods being transported across space. This tends to militate against the development of geographically peripheral regions and areas, because the most innovative and information-intensive sectors of the economy tend to require the most intense face-to-face contact. Such intense face-to-face contact is afforded by geographical centrality in the most densely populated regions. These complexities pose enormous challenges for regional policies designed to promote the economies of peripheral regions and countries.

References

Acs Z (2002) Innovation and the Growth of Cities. Edward Elgar, Cheltenham

Anderson JE, Van Wincoop E (2004) Trade Costs. Journal of Economic Literature 42:691-751

Arita T, McCann P (2000) Industrial alliances and firm location behavior: some evidence from the US semiconductor industry. Applied Economics 32:1391-1403

Armstrong H (1995) Convergence among regions of the European Union. Papers in Regional Science 74:153-152

Arrow KJ (1962) The economic implications of learning by doing, Review of Economic Studies 29:153-173

Arthur WB (1989) Competing technologies, increasing returns and lock-in by historical events. Economic Journal 99:116-131

Audrestch D, Feldman MP (1996) R&D spillovers and the geography of innovation and production. American Economic Review 86(3):630-640

Barro R, Sala-i-Martin X (1992) Convergence. Journal of Political Economy 100:223-251

Becattini G (2004) Industrial districts: A new approach to industrial change, Edward Elgar, Cheltenham

Bennett RJ (1998) Business associations and their potential to contribute to economic development: re-exploring an interface between the state and the market. Environment and Planning A 30:1367-1387

Borts GH, Stein JL (1964) Economic growth in a free market. Columbia University Press, New York

Cantwell JA, Iammarino S (2000) Multinational corporations and the location of technological innovation in the UK regions. Regional Studies 34:317-332

Casson M, McCann P (1999) Globalization, Competition and the Corporation: The UK Experience. In: Whitman M (ed) The evolving corporation: global imperatives and national responses. Group of Thirty, New York, pp 127-172

Cheshire PC, Carbonaro G (1995) Convergence-divergence in regional growth rates: an empty black box? In: Armstrong HW, Vickerman R (eds) Convergence and divergence among European regions. Pion, London, pp 89-111

Clement NC, del Castillo Vera G, Gerder J, Kerr WA, MacFadyen AJ, Shed S, Zepeda E, Alercon D (1999) North American economic integration: theory and practice. Edward Elgar, Cheltenham

Cohen B (1998) The Geography of Money. Cornell University Press, Ithaca NY

Financial Times (1997). Survey: Logistics. 7 October.

Financial Times (1998). Survey: Supply Chain Logistics. 1 December.

Financial Times (1999). Survey: Supply Chain Logistics. 17 June

Fingleton B, McCombie JSL (1998) Increasing returns and economic growth: some evidence for manufacturing from the European Union. Oxford Economic Papers 51:574-580

Fujita M, Krugman P, VenablesAJ (1999) The spatial economy: cities, regions and international trade. Cambridge, Mass, MIT Press

Gaspar J, Glaeser, EL (1998) Information Technology and the Future of Cities. Journal of Urban Economics 43:136-156

Glaeser EL (1998) Are Cities Dying? Journal of Economic Perspectives 12:139-160

Glaeser, E., Kallal HD, Scheinkman JA, Shleifer, A (1992) Growth in Cities. Journal of Political Economy 100:1126-1152

Granovetter M (1973) The strength of weak ties. American Journal of Sociology 78:1360-1389

Gordon, IR (2000) Distance deterrence and commodity values. Environment and Planning A 37:889-900

Gordon IR, McCann P (2000) Industrial clusters, complexes, agglomeration and/or social networks? Urban Studies 37:513-532

Hart P, Oulton N (1999) Galtonian regression, company age and job generation. Scottish Journal of Political Economy 48, 82-98.

Henderson JV, Kuncoro A, Turner M (1995) Industrial Development in Cities, Journal of Political Economy 103:1067-1085

Helpman E, Krugman P (1985) Market structure and foreign trade. MIT Press, Cambridge, Mass

Hummels D (1999) Have International Transportation Costs Declined? Department of Economics Working Paper, Purdue University.

Isard W, Kuenne, RE (1953) The Impact of Steel upon the Greater New York-Philadelphia Industrial Region. Review of Economics and Statistics 35:289-301

Isard W, Schooler EU, Vietorisz, TS (1959) Industrial Complex Analysis and Regional Development. John Wiley and Sons, New York

Jaffe A, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidence by patent citations. Quarterly Journal of Economics 108, 577-598.

Johansson B, Karlsson C, Stough R (2002) Regional Policies and Comparative Advantage. Edward Elgar, Cheltenham

Johnson HT, Kaplan, RS (1987) Relevance Lost: The Rise and Fall of Management Accounting. Harvard University Press, Cambridge, Mass

Krugman P (1991) Geography and Trade. MIT Press, Cambridge, Mass

Krugman P (1996) Urban concentration: the role of increasing returns and transport costs. International Regional Science Review 19:5-30

Krugman P, Venables AJ (1990) Integration and the competitiveness of peripheral industry. In: Bliss C, de Macedo, KB (1990) *Unity* with Diversity in the European Economy. Cambridge University Press, Cambridge

Lewin P (ed) (2002) The economics of QWERTY. Palgrave, Basingstoke, UK

Linbeck A, Snower DJ (1989) The insider-outsider theory of employment and unemployment. MIT Press, Cambridge Mass

Lucas RE (1988) On the mechanics of economic growth. Journal of Monetary Economics 22:3-42

Marshall A (1920) Principles of economics 8th edition. Macmillan, London

McCann P (1993) The logistics-costs location-production problem, Journal of Regional Science 33, 503-516

McCann P (1995) Rethinking the economics of location and agglomeration. Urban Studies 32, 563-577

McCann P (1998) The economics of industrial location: A logistics-costs approach. Springer, Heidelberg

McCann P (2001a) Urban and Regional Economics. Oxford University Press, Oxford.

McCann P (2001b) A Proof of the Relationship between Optimal Vehicle Size, Haulage Length and the Structure of Distance-Transport Costs. Transportation Research A 35:671-693.

McCann P, Fingleton, B (1996) The regional agglomeration impacts of just-in-time input linkages: Evidence from the scottish electronics industry. Scottish Journal of Political Economy 43:493-518

Mills ES (1970) Urban Economics. Scott, Foresman and Co., Glenview, Illinois

Moses L (1958) Location and the Theory of Production. Quarterly Journal of Economics 78:259-272

Naisbitt R (1995) The Global Paradox. Avon Books, New York

Nishiguchi T (1994) Strategic Industrial Sourcing: The Japanese Advantage. Oxford University Press, Oxford

Overman HG, Redding S, Venables AJ (2001) The Economic Geography of Trade Production and Income: A Survey of Empirics, CEPR discussion paper 2978, London

Piore MJ, and Sabel CF (1984) The second industrial divide: possibilities for prosperity. Basic Books, New York

Porter M (1990) The competitive advantage of nations. Free Press, New York

Porter M (1998) Clusters and the new economics of competition. Harvard Business Review November-December.

Rao PK (2003) The economics of transactions costs. Palgrave, Basingstoke

Reid N (1995) Just-in-time inventory control and the economic integration of Japanese-owned manufacturing plants with the county, state and national economies of the united states. Regional Studies 29:345-355

Romer PM (1986) Increasing returns and long run growth. Journal of Political Economy 94:1002-1037

Romer PM (1987) Growth based on increasing returns due to specialization. American Economic Review 77:56-62

Saxenian A (1994) Regional Advantage. Harvard University Press, Cambridge, Mass

Schonberger RJ (1996) World class manufacturing: The next decade. Free Press, New York

Scott AJ (1988) New Industrial Spaces. Pion, London

Simmie J, Sennett J (1999) Innovative clusters: Global or local linkages? National Institute Economic Review 170:87-98

Stiglitz JE (2000) The contributions of the economics of information to the twentieth century economics. Quarterly Journal of Economics 115:1441-1478

Suarez-Villa L, Walrod W (1997) Operational strategy, R&D and intrametropolitan clustering in a polycentric structure: The advanced electronics industries of the Los Angeles basis. Urban Studies 34:1343-1380

Toffler A (1980) The third wave. Bantam Books, New York

The Economist (1999a) The net imperative: A survey of business and the Internet. 26 June.

The Economist (1999b) The world in your pocket: A survey of telecommunications. 9 October.

Tyler P, Kitson M (1987) Geographical variations in transport costs of manufacturing firms in Great Britain. Urban Studies 24:67-73

United Nations Centre for Human Settlements (1997) Human Settlements Basic Statistics. Nairobi

Vernon R (1966) International investment and international trade in the product cycle. Quarterly Journal of Economics 80:190-207

Weber A (1909) Uber den Standort der Industrien. Translated by Friedrich CJ (1929) Alfred Weber's Theory of the Location of Industries. University of Chicago Press, Chicago

Williamson, OE (1975) Markets and Hierarchies. Free Press, New York

Yeung MT (1999) Regional trading blocs in the global economy: the EU and ASEAN. Edward Elgar, Cheltenham

3 Transport, Globalization and the Changing Concept of the Region*

Roger Vickerman

Department of Economics, University of Kent, Canterbury, Kent, UK

3.1 Introduction

Globalization is changing the way regions relate to one another, but it is also changing the concept of the region as a geographically defined entity. This has profound implications for the way we model regional systems. The traditional model, based primarily on regional markets which clear internally, but with some specific external linkages, mainly trade-related, is now much less appropriate. The initial response to the globalization question has been to attempt to make the trade relationships more realistic, for example by specifying them in terms of complete supply-chain linkages rather than just simple flows of goods from one region to another, but this does not go far enough in capturing the dimensionality of the changing regional situation.

The impact of globalization is to shift intra-regional linkages in all markets to extra-regional linkages, for example through the wider geographical sourcing of inputs, including labor. This has enabled a process of increasing specialization in regional economies. The key indicator of this is that inter-regional trade, previously dominated by inter-industry trade, is now predominantly intra-industry in character. But it also has ramifications for labor markets and labor mobility, for housing markets and how they relate to the distribution of employment, and for transport markets.

---

[*] Revised version of a paper presented in the 'Globalization and Regional Economic Modeling' Theme at the World Congress of the RSAI, Port Elizabeth, South Africa, April 2004

3.2 Inter- and Intra-regional Linkages

Regional modeling has followed two distinct paths of development which can be described as the economics tradition and the regional science tradition. The two are consistent with each other, but have different emphases. The former bases its models on theoretical consistency, the latter is more concerned with workable empirical models which can be used for regional planning purposes.

The economics tradition has itself gone though significant change in recent years. Indeed it can be argued that it was the deficiencies in the traditional economics approach which led to the development of the regional science approach. Traditional regional economic models were essentially regional adaptations of Keynesian macroeconomic models. The emphasis was on the determination of regional macroeconomic aggregates in which the focus was on the possibility of under-employment equilibrium and hence on aggregate demand deficiency. The absence of location and space and a lack of concern with the microeconomic foundations made this approach less suitable for dealing with a period of dynamic change in the way regions relate to one another. Spatial linkages are those between regions; regions trade with one another rather like countries and there is no specific modeling of the costs of that trade except in the way they reflect trade barriers or currency exchange rates. The regional economy itself is aspatial.

The ‘new economic geography’ (or ‘geographical economics’ or ‘spatial economics’), starts with a much more microeconomic foundation for all sectors and factors and thus is able to deal more effectively with the spatial structures both between and within regions. Two features stand out. One is the explicit recognition of increasing returns to scale and imperfect competition, which are an inherent feature of the spatial world. Thus local market size and the ability of firms to set prices, plus the resulting endogeneity of the number of firms in the market become significant features of the model. Secondly, trade and transport costs are recognized as a critical variable which is traded-off with market size and scale economies. This recognition of the substitutability of transport and location both with other factors of production and with market structure is a critical advantage of this approach over traditional regional science models. These tended to use classic location theory approaches to determine the optimal location of a price-taking firm for which size and production technology were given. However, although the model has become much more spatial there is still a tendency to view the model as an inter-regional model (i.e. handled identically to an international trade model) in which each region is aspatial.

Nevertheless it has proved difficult to develop tractable models which can incorporate both spatial dimensions, intra- and inter-regional relationships, as well as the basic requirements of a globalizing economy in terms of imperfect competition and the dynamic response of economic agents. Therefore models have continued to be developed in parallel and we need to take an eclectic approach to using these to try and understand some of the processes involved.

The typical regional model in the land-use transport interaction tradition includes a production sector, a labor market, and a series of external linkages determining trade and migration. A good example of a fairly advanced form of this is the SASI model (Wegener and Bökeman, 1998). This is particularly useful as it is set up to model the changes in regional economies resulting from changes in accessibility derived from changes in the transport network (see Figure 1 for the structure).

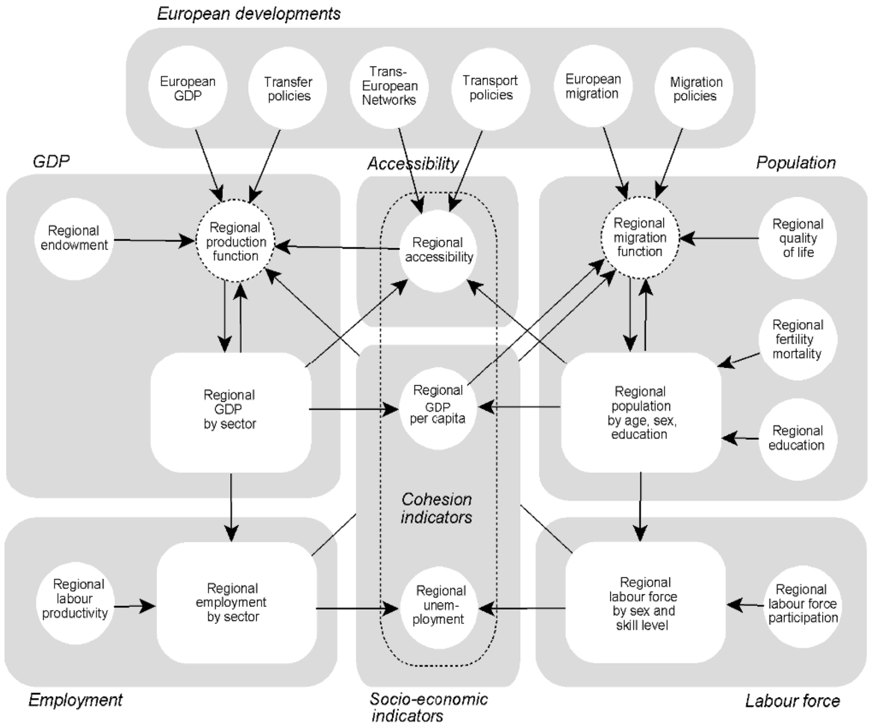


Fig. 3.1. SASI model structure.
Source: Bröcker et al (2004)

The model has been used subsequently in a series of EU research projects such as IASON and ESPON. The advantage of a model such as SASI is that it has explicit demand and supply side effects. For example, the labor market equilibrium depends on regional population change (including migration) and the regional demand for labor by sector. Similarly regional output is determined by a production function and regional income determines regional demand. In both of these markets there can be a local lack of market clearing such that excess demand for goods leads to imports and excess supply of labor leads at least in the short-run) to unemployment, but may ultimately lead to some out-migration.

There are two areas in which this approach may be unduly limiting. First, the trade element will depend on regional deficient or excess supplies by sector; in a world of product differentiation and intra-industry trade there will be substantial cross hauling of goods between economically similar regions. Secondly, transport is represented by accessibility-related costs; in practice such costs will also be product as well as distance specific and furthermore will vary according to the market structure in particular transport markets, including the degree of balance in flows along particular corridors.

Developments of theoretical new economic geography have shown the importance of backward and forward linkages in determining the relative performance of a region and the stability of core-periphery structures (e.g., Venables, 1996). With differentiated products, where each product produced in a specific location is distinct, individuals' preferences for variety will lead to sourcing of the supplies within any one sector from a range of different regions. As we know from studies of intra-industry trade this type of cross-haul trading is positively related to income levels and the similarity of income levels between regions, but negatively related to distance and the presence of scale economies.

In such a situation, accessibility will work in a slightly different way from that envisaged in standard regional modeling since it is more about the definition of relevant market areas than just the cost of reaching markets or sourcing supplies.

Much has been done to widen the appropriate definition of transport costs to include other factors. Accessibility measures allow for the importance of time (and potentially other elements of generalized cost), the use of logistics costs approaches sets this within the context of the overall logistics strategy of the firm (McCann, 1998). However, this may not go far enough for understanding the spatial structure of markets and competition. The convenience of the iceberg transport costs concept in the new economic geography may have been a distraction here. This makes the usual assumption that transport is likely to be a disutility and thus should be minimized since in effect it reduces the value of the transported goods. However, this is based on the assumption that transport has an essentially given distance-related cost to the user. Two related elements are relevant here: the overall size and direction of flows on any one corridor and the structure of competition in the transport market itself. Both of these lead to a restructuring of perceived geography.

Given the scale economies present in transport, to which should be added the potential for scope and network economies, goods carried over sections where there are large flows would typically be expected to be able to be carried at lower unit costs, up to the point where congestion becomes a problem. Congestion, or excess demand for available capacity, will lead to increasing time and reliability costs, but also to peak-load pricing by operators exploiting their market power. Hence there is a recursive link between the demand for inter-regional trade and the cost of supporting that trade. There is a further element in this equation relating to the balance in flows on each corridor. Where there is a much larger flow in one direction than the reverse, capacity on the corridor is likely to be determined by the larger flow. Hence even small flows may benefit from the scale economies implicit in joint costs. However, this goes further, operators will face sunk costs for return hauls of vehicles and thus may attempt to price higher for moves in the direction of the larger flows and at marginal cost (or less) for the return haul. Hence prices for hauls out of a region which face a large adverse balance of trade may be much less than for prices of hauls into that region – a form of unintentional non-tariff barrier.

Competition, both within and between modes, will also affect the prices charged and hence the relevant cost of transport for inter-regional trade. Larger flows provide more scope for competition. Corridors with larger flows are more likely to have more alternative modes available (something which is imperfectly incorporated into typical accessibility indicators) and more competing operators within each mode (see Vickerman, 1999, for a more detailed discussion). Larger flows also provide more scope for predatory entry and bigger potential gains from rent seeking. This may consolidate lower prices on heavily used corridors at the expense of less heavily used corridors, particularly to the detriment of peripheral regions (Vickerman, 1998). This situation is also complicated by changing policy stances towards transport through deregulation and liberalization (for some further discussion of this point see Vickerman, 2004)

Note that in this discussion we have concentrated on tradable goods since they traditionally have defined regions' comparative advantage, trading position and hence potential for growth. Non-tradable goods have been seen to be secondary to, and dependent on the economic base of the tradable sector. Changing technology is leading to traditionally non-tradable goods also being able to be traded across regions – perhaps the starkest example of this is the outsourcing of call centers to India.

3.3 Changing Spatial Labor Markets

The previous section has dealt with structures in the traded goods market. A second and equally important set of structures are those in the labor market. Although models have moved on somewhat from the assumption of full market clearing, implying that there will always be adjustment to full employment equilibrium, they are still dominated by rather simple adjustment mechanisms which do not always reflect the reality of modern labor markets.

Four issues need to be raised here. First is the question of allowing for skill differentials. This is similar in some respects to the product differentiation question considered in the previous section. It identifies that it is not sufficient to deal just with an overall clearing of the labor market, but that different parts of the labor market can be in different degrees of imbalance at any one time. This affects the way the market adjustment mechanism is expected to work.

Secondly, traditional models of migration work on the basis of individuals making in effect once and for all decisions on the basis of perfect knowledge of the lifetime returns to responding to current period differences in wages and employment opportunities. Increasingly it has been recognized that mobile individuals engage in what could be termed both temporary or periodic and sequential migration. Temporary migration is where return to the home region is planned as part of the initial decision. Mobility here is often associated with the acquisition of either skills or financial capital at the destination. Sequential migration is a variant on this where the mobile worker engages in a series of moves. The problem which these types of move cause for modeling at the regional level is that they do not just depend on current relative values in the home and potential host region, but also on the expectation of movements in those values and on the aspiration of the individual to achieve some specific goal.

Increasingly there has also come to be much less difference between traditional permanent migration between two labor market areas and what could be termed periodic or long-distance commuting. Given that the potential labor market area for an individual will depend on the availability of jobs over a wider area than that accessible by daily commuting, but recognizing the constraints imposed by household considerations and the housing market, individuals may choose to move workplace, but not residence. This recognizes also that the job move may not be considered permanent. This increase in inter-regional (including international) commuting not only affects the definition of labor market equilibrium in each region, but also has important implications for transport. It is likely that the observed increase of about one-third in travel to work distances in the UK (from 9.8 km to 13.0 km) over the period 1985/86 to 1997/99 whilst the average journey length for all journey purposes increased by 25% (8.4 km to 10.5 km) (Vickerman, 2002) underestimates the actual change in spatial labor markets since it only includes daily journeys. This goes further than the argument on reasons for urban sprawl analyzed by Prud'homme and Lee (1999). Some careful analysis of recent Census data on workplace and transport will be needed to unpack this issue.

Finally, much has been made of the impacts of changing technology such as teleworking on transport demand and on patterns of work. This is a further dimension through which there can be change in the spatial structure of labor markets by removing the need for the daily commuting trip. The sought for usefulness of this change in terms of a reduction in travel may however not be forthcoming as the teleworking individual may use this to be able to relocate to a labor market where other members of the household may be better placed for either work or education. This may make the less frequent journey to work significantly longer and increase total travel. The relative freedom provided by teleworking may also give rise to additional travel demands. For example, one of the most significant causes of congestion in urban areas in the UK is the journey to school; although there are several reasons for the increase in car based journeys to school, increasing flexibility over working hours on the part of parents is one contributory factor.

3.4 Defining Regions

The essential message of the two preceding sections is that the concept of the region itself is much less clear than assumed in traditional modeling approaches. We observe in effect numbers of overlapping regional structures defined by differentiated products and markets for individual skills, constrained and conditioned by changing household structures, housing markets and changing technologies affecting production, modes of working, modes of living and modes of travel and transport. In a traditional regional economy most of these map neatly into definable geographical spaces. This space would have both an economic identity, defined by a given structure of industry with specific resource needs, markets, and trade relations, and a parallel institutional identity or regional government which provided the appropriate policy setting for the region's economy and its relations with other regions. This is clearly seen in the SASI model used as an example in Figure 1. Given the nature of most regions' economic structures, their external spatial relations were also fairly clearly defined in terms of the relatively small number of other regions with which trade, migration and other economic flows interacted. Regions were therefore not only defined geographical spaces, but these spaces existed within defined spaces.

In the structure we have been identifying in this paper these spatial relationships have changed significantly. Goods markets do not map simply into similar geographical spaces, labor markets are no longer predominantly confined within regions and migration spaces have changed considerably. One further way in which space has become less defined is in the lack of such a clear distinction between tradable and non-tradable goods. When individuals move easily between labor markets, for example living in one region and working in another, often maintaining residences in both, they may well purchase normally non-traded goods in different regions according to relative prices and/or convenience (the classic example of the haircut as a non-tradable good is a possible case).

Clearly, however, we still need some concept of the geographically identifiable region. Policy will continue to be implemented at the regional level and, however problematic this is likely to be, will be based on indicators defined at the regional level. Whilst overlapping regions, e.g. separate regional structures for industrial policy, labor market policy, transport policy, may seem initially attractive, the interactions between these policy areas will make it difficult to coordinate policy in such a structure. The issue is therefore how to include recognition of the new spatial structures within a workable policy and institutional environment. This requires that the policy makers have an appropriate range of models which recognize the issues.

3.5 Regional Models in Changing Regional Structures

The temptation to throw away existing models and start again rapidly diminishes when we consider the enormity of the task, the real question is whether and how can we adapt existing modeling strengths to allow for these changing spatial arrangements? Models like the SASI model discussed above already have an interregional dimension. Currently this is represented by essentially aspatial external linkages such as a migration function, a link between external GDP and the regional production function and the impact of EU transport and TENs policies on accessibility. These are aspatial in the sense that there is only interaction between each region and some average performance of the rest of the wider European economy; calibrating the model simply searches for an equilibrium for each region in the system in its interaction with the rest of the system and not in terms of a full set of individual inter-regional interaction.

One simple way of incorporating this is through the use of potential measures of all indicators. Preferably these potential measures should use different appropriate distance decay functions so that, for example, the level of import penetration in a sector is related to the potential production of it and a set of close competitor regions. The key here is the degree of variation in the distance decay functions for different activities which will define the size of the relevant region. This will depend on both the degree of competition in the market in question and the unit cost of transport. Transport costs will depend here on the size of total flows, the specific characteristics of particular goods (such as the need for special handling, safety or security issues, value to bulk ratios, time sensitivity or value of time in transit), and the degree of both between- and within-mode competition. Similar considerations apply to markets involving the transport of people such that values of time and reliability issues come to the fore.

3.6 Conclusions

This paper has not come to any specific conclusions. It rather raises issues which need to be addressed in an increasingly globalized world. It has also not produced an alternative model, rather a series of questions which need to be asked of existing models. We have used one specific model, not because it is thought to be particularly good or bad, but just representative of a way of making complex problems tractable. The key message here is that we should not aim to throw away the basis of such models, but rather to use their undoubted strengths and build on these. The paper has started to suggest areas in which this building can be done.

References

Bröcker J (2002) Spatial effects of European transport policy: a CGE approach. In: Hewings JG, Sonis M, Boyce DE (eds) Trade, networks and hierarchies: Integrated approaches to modelling regional and interregional economics, Springer, Berlin, pp 11-28

Bröcker J, Capello R, Lundquist L, Pütz T, Rouwendal J, Schneekloth N, Spairani A, Spangenberg M, Spiekermann K, Vickerman R, Wegener M (2004) Territorial impact of EU transport and TEN policies, Final report of Action 2.1.1. of the European Spatial Planning Observation Network ESPON 2006, Kiel, Institut für Regionalforschung, Christian-Albrechts-Universität

McCann P (1998) The economics of industrial location: A logistic costs approach. Springer, Berlin

Prud'homme R, Lee C-W (1999) Size, sprawl, speed and the efficiency of cities. Urban Studies 36:1849-1858

Vickerman RW (1998) Accessibility, peripherality and spatial development: the question of choice. In: Reggiani A (ed) Accessibility, Trade and Location Behaviour, Ashgate, Aldershot, pp 79-93

Vickerman RW (1999) The transport sector, new economic geography and economic development in peripheral regions. In: Beuthe M, Nijkamp P (eds) New contributions to transportation analysis in Europe. Ashgate, Aldershot, pp 45-66

Vickerman RW (2002) Sustainable mobility in an age of internationalization. In: Higano Y, Nijkamp P, Poot J, van Wijk K (eds) The region in the new economy. Ashgate, Aldershot, pp 313-327

Vickerman RW (2004) Conflicts between transport policies and spatial development policies: perspectives on regional cohesion in the European Union. Paper presented at the World Congress of the RSAI, Port Elizabeth, South Africa, April 2004.

Wegener M, Bökemann D (1998) The SASI model: Model structure, SASI deliverable D8, Berichte aus dem Institut für Raumplanung 40, Dortmund, Institut für Raumplanung, Universität Dortmund

4 ICT, the New Economy and Growth: The Potential for Emerging Markets*

Russel J. Cooper[1] and Gary Madden[2]

[1]Department of Economics, Macquarie University, Australia
[2]Communications Economics and Electronic Markets Research Centre, Department of Economics, Curtin University of Technology, Australia

4.1 Introduction

There is a consensus view that, since the mid-1990s the US has created a 'New Economy' based on the rapid deployment and widespread utilization of new information and communications technology (ICT) and by advances in computing and information management that caused a structural acceleration in labor productivity and output growth (Jorgenson and Stiroh 2000, Salvatore 2003).[1] The term New Economy is intended to capture the role of this new technology in contributing to the non-inflationary growth and high employment that characterized the period (Jorgenson and Wessner 2002).[2] The rapid rate of technological innovation in ICT and the rapid growth of the Internet are seen as underpinning such productivity gains (Jorgenson and Wessner 2002). Gains are derived from greater efficiency in the production of computers and from the expanded use of ICT (Onliner and Sichel, 2000). Structural change has also arisen from the reconfiguration of knowledge networks and business patterns made possible via ICT innovation. For example, business-to-business e-commerce and Internet retailing are altering how firms and individuals interact to enable greater efficiency in purchasing, production processes and inventory management (Jorgenson and Wessner, 2002). Panagariya (2000) claims the transfer of ICT, in particular Internet technology, to developing countries enables the leaping of technological stages and provides a basis

---

* This chapter is a revision and extension of research presented to Regional Science Association International World Congress 2004, Port Elizabeth, South Africa, 14-17 April 2004

[1] The definition of ICT adopted by the OECD ICCP panel of statistical experts is that it is the set of activities that facilitate by electronic means the processing, transmission and display of information.

[2] Although the New Economy is a macroeconomic phenomenon, its underlying dynamics combine elements of technological innovation, structural change and public policy (Jorgenson and Wessner 2002).

to enhance macroeconomic growth. Conversely, unless ICT is widely adopted, developing countries are likely to remain unable to compete in the New Economy where the sources of advantage are high knowledge-intensity and the fast adoption of new technological innovation, not low-cost factor advantage (OECD 2000).

The plan of this chapter is as follows. Section 4.2 provides a review of the benefits from the introduction of new ICT and counterbalances this with a review of the findings of the empirical ICT-growth literature. Contradictory empirical results suggest a need for an alternative modeling approach. We present a proposed model in section 4.3 which contains a detailed expenditure share allocation model embedded within a stochastic intertemporal optimization model. The rationale for this approach is twofold. First, we aim to utilize an 'atemporal/intertemporal rationality paradigm' of consistency between instantaneous and intertemporal decisions to enable information more readily available on atemporal allocation decisions to infer less readily available information on intertemporal growth decisions. Secondly, given the special features of ICT as a general purpose technology (GPT), we believe that if the use of ICT is to be considered as a contributor to growth then explicit linkages between the demand for ICT-embedded products and the supply of new ICT capital need to be examined. The proposed model is intended to provide a blueprint for an extensive program of future research. At this stage a full analytical solution to the stochastic intertemporal model is not available. However, the role of growth in leading to shifts in market demands for specific types of ICT is still able to be examined, albeit in an aggregated framework in our modeling at this point. In a model allowing for consumption and production externalities, consistency between atemporal and intertemporal decision making suggests that shifts in ICT product proportions will be associated with differential growth. The model incorporates some core features of the New Economy such as 'Application Sector' (AS) risk and important dynamic aspects of networks (consumption and production externalities). With modern developments in ICT in the New Economy it is becoming increasingly difficult to distinguish meaningfully between consumer and producer demand driven applications. Accordingly, interaction between consumer and firm decisions is considered by a representative agent model in which sustained growth may result from positive consumption and production externalities to investment in network input k. In section 4.4, the demand side of the model is estimated from observed ICT expenditure share data and the supply of the essential network input is then inferred from the atemporal/intertemporal consistency paradigm. Implications for countries at different stages of development are briefly. Section 4.5 provides a brief conclusion and agenda for future research.

4.2 General Purpose ICT and Cascading Innovation: Theoretical Notions and Empirical Evidence

Along with a rise in sectoral productivity it is argued that new ICT also improves business practice, generates spillovers to other industry and so raises productivity economy-wide. Internet use, in particular, contributes to the reduction of transaction costs related to production and distribution; enhances management efficiency by enabling firms to more effectively manage their supply chains and communicate better within the firm and with their customers and business partners; increases competition by making prices more transparent and broadening markets for both buyers and sellers; and finally increasing consumer choice, convenience and satisfaction in a variety of ways (Litan and Rivlin 2001a). Such arguments are implicitly based on the belief that the New Economy is not just another 'leading sector', i.e., an invention and innovation explosion that revolutionizes only a narrow slice of the economy. Examples of US leading sectors include railways in the 1870s, organic chemicals in the 1890s, automobiles in the 1920s, television in the 1950s and air transportation in the 1960s. Yet it could be argued that they did not change the standard economic growth dynamic; rather they defined it (Cohen et al. 2000). Cohen et al. argue that the New Economy experience is deeper and broader than that traditionally associated with leading sectors. While Cohen et al. concede that semiconductors, computers and communications do constitute a large leading sector, they consider that they not only enhance productivity within the sector but provide other economic sectors with a new set of 'tools' that open new possibilities for economic organization. Jorgenson and Wessner (2002) categorize such processes as evidence of 'cascading innovation.' That is, "productivity gains come not just from faster processing of information but also from changes in the way a firm operates and from additional technological advances made possible by the information technology. Moreover, as information technology spreads and its potential is more widely recognized, it generates demand for even faster processing—that is, another round of information technology innovation, which in turn creates the potential for more innovation both organizationally and in the products and services the re-organized and information-technology-enabled firm can create. This demand forms the basis for Moore's Law, under which the computing capacity of chips doubles every 18 months. Moreover, these enormous and regular increases in chip power provided the technological basis for the Internet, which in turn now generates more rounds of cascading innovation in how businesses operate and what they produce" (Jorgenson and Wessner 202: 47).

Brynjolfsson and Hitt (1995), Jorgenson and Stiroh (2000), Oliner and Sichel (2000), Bosworth and Triplett (2001), and Jorgenson (2001) showed that the late-1990s US surge in productivity and growth is in part due to IT investment and Internet growth. With this established, the IT-productivity debate shifted to the question of whether an IT-led economy provides permanent improvement in economic growth, or whether much of the acceleration in productivity is driven by the business cycle and concentrated in a few (leading) sectors of the economy (Gordon 2000). Bosworth and Triplett (2000), Oliner and Sichel (2000), and Jorgenson (2001) report IT investment had a relatively small impact on US economic and average labor productivity growth (because of the small IT share of total capital stock) — even when average labor productivity growth is low. While magnitude of the IT contribution to productivity remains an issue of debate, there is convincing evidence of positive long-term impacts from IT investments on national productivity. Early studies by Jorgenson and Stiroh (1995) report an IT investment contribution to productivity growth of approximately 6% of annual productivity growth (of 2.94%) for 1959-1973. During the period 1973-1995, Jorgenson and Stiroh (2000) find that IT investment contributed 13% to US average economic growth (of 3.04%) and 27% to labor productivity growth (of 1.4%). However, the acceleration of average labor productivity growth from 1995-1999 was itself due in part to rapid IT investment growth. That is, the increased IT impact on productivity resulted from a higher IT investment rate with accumulated IT investment becoming a substantially greater share of the total US capital stock. Jorgenson (2001) shows that during this period, IT capital contributed about 28% (of 4.08% p.a.) of US economic growth and 42% (of 2.11% p.a.) growth in average labor productivity. Oliner and Sichel (2000) estimate similar contributions. Debate as to the share of the improvement attributable to structural change from technical progress (changes in the trend line) and that due to the business cycle continues. Short-run growth can increase both measured productivity and investment. While it is difficult to isolate the separate impacts of trend and cycle—it is clear that both matter. Gordon (2000) attributes a substantial share of the 1995-2000 productivity growth acceleration to business cycle effects however Stiroh (2002) finds little if any cyclical effect. Their results have very different implications for the magnitude of the impact of IT investment on productivity. Regardless, it remains clear that the broad and continuing use of IT made a substantial difference to US long-term average labor productivity growth. Further, Litan and Rivlin (2001b) estimate the likely productivity impact from the Internet across eight industry sectors that account for 70% of US GDP, and find the impact of the Internet over 5 years translates to a 0.2% p.a. - 0.4% p.a. addition to baseline trend productivity growth.

Comparable research outside the US finds most of the foregoing trends apply to developed countries. Schreyer (1999) examines G-7 countries and reports that IT made a positive contribution to productivity and economic growth in these nations for 1990-1996. Daveri (2000) updated and broadened this analysis to encompass 18 countries. The essential findings remain unaltered, viz., IT capital contributes to growth, and because IT investment is growing faster than labor input, it contributes to average labor productivity via capital deepening. Dewan and Kraemer (1998, 2000) and Pohjola (2001, 2002) corroborate these finding for a global sample of nations. They conclude that wealthier industrialized countries reveal a positive relationship between IT investment, growth and productivity, however there is no evidence of this relationship for developing countries. Dewan and Kraemer (1998, 2000) argue this outcome is due to low developing country IT investment-to-GDP ratios and a lack of complementary assets such as network infrastructure and a knowledge-base to support effective IT use.[3] Kraemer and

---

[3] Brynjolfsson et al. (2000) find that market valuation effects are greatest for firms that have high levels of investment in both IT and organizational capital—again indicating the factors are complementary. So far, studies have not identified a relationship between IT investment and firm profitability. Bresnahan et al. (2002) test for the relationship between IT use, organizational change, and firm skill levels. They find IT use, complementary workplace reorganization and higher rates of new product introduction result in

Dedrick (2001) study 43 developed and developing countries and find that growth in IT investment per worker is positively correlated with labor productivity growth, but the level of IT investment is not.

4.3 A Stylized Model of ICT Network Growth Driven by Demand for New Economy Products

4.3.1 The Intertemporal Growth Model

Theoretical new growth models employ a range of specifications that imply endogenous long-run growth rate or level effects. Specific models consider externalities, learning by doing, R&D and technology diffusion. Such models explain divergent growth patterns, cluster convergence and multiple equilibriums—raising the question of how a country can shift, should it be able, to a higher equilibrium growth path. However, much of the theoretical work concerning the role of increasing returns in generating endogenous growth treats the economy as ‘flat’, i.e., no interaction between sectors are allowed for. Technical change is usually assumed pervasive—it occurs with similar intensity throughout an economy (Bresnahan and Trajtenberg 1995). This approach does not match well the differential productivity patterns experienced internationally by industry, especially those who are, more rather than less, intensive ICT users. To address this concern Bresnahan and Trajtenberg develop the notion of General Purpose Technology (GPT). GPT is characterized by the “potential for (its) pervasive use in a wide range of sectors by their technological dynamism”. As a GPT advances and spreads throughout an economy it brings about and fosters more general (or enhanced sector) productivity gains. That is, GPTs are enabling technologies that provide further opportunity, rather than offering complete solutions. To make the concept operational they propose an Applications Sector (AS) as a user (or potential user) of GPT inputs — this leads to AS own innovative activity, educing further productivity gain.

In developing a model to incorporate these characteristics, we model expenditure on ICT as the equilibrium outcome of intertemporal decisions by suppliers of a GPT and consistent atemporal decisions of AS demanders of the resultant productivity and utility-enhancing products. A representative agent paradigm is used to model both producers and consumers, with the overall utility maximizing objective being that of a representative consumer-firm. A positive network externality is associated with production activity. Internet network externalities also arise through consumption. The representative agent obtains utility from real total consumption and current network size. The agent has an option to not consume all income. Any saving defaults to network investment. Further, the consumer can relinquish ownership of part of the network in exchange for an AS investment that

---

greater use of high-skilled labor. They also find that organizational change accompanied by technology change has a greater impact on skill levels than technology change alone.

provides a risky return. Let $\tilde{F}(k)$ denote the aggregate production function of a representative firm where k is the network input recognized by the firm as under its control.[4] All other primary factor inputs (labor, the services of rented capital equipment) are optimized out. The dependence of $\tilde{F}(k)$ on the prices of these inputs is suppressed for notational clarity. We extend the concept of the production function from $\tilde{F}(k)$ to, say, $F(k,k^*)$ to allow for a network externality generated through productive activity. The argument k^* allows 'endogenous growth' to occur in the network growth equation, viz., the production function exhibits decreasing returns in k (from the perspective of the firm) and increasing returns when k is equated to k^* post-optimization. That is, during optimization k^* is treated by the firm as exogenous, and post-optimization k^* is equated to k when model equations are derived. Thus positive externalities arise from network capital and are a source of increasing returns in production.

Let $\tilde{U}(c)$ denote the instantaneous utility function of a representative consumer where c is real total consumption. We allow for an Internet network externality in consumption by extending $\tilde{U}(c)$ to, say, $U(c,k^*)$ where k^*, the current network size for an average firm, is assumed to affect the consumer's utility but is outside the control of the individual consumer.

The output produced by the representative agent is a source of income. Income is also derived from a stochastic return from AS activity (i.e. due to venture capital investment in the New Economy), obtained by selling x of the network k. The agent foregoes a 'sure' rate of return $F_k(k,k^*)\,x\,dt$ for receipt of risky return $x\,d\rho/\rho$.[5] The risky asset is assumed to pay no dividend but provides a capital gain or loss. The resulting stochastic flow of income from production and investment sources allows the supply side of the income identity to be represented as $F(k,k^*)kdt+\left(d\rho/\rho-F_k(k,k^*)\,dt\right)x$.

The price of the risky asset, ρ, is modeled as following a geometric Brownian motion with drift μ and volatility σ, viz. $d\rho=\mu\rho\,dt+\sigma\rho\,d\xi$ where $d\xi$ is Brownian motion, with the properties $E(d\xi)=0$, $E(d\xi)^2=dt$.[6]

An alternative to consumption is the retention of earnings, which are employed to extend the network. Consequently, network expansion is stochastic. The de-

---

[4] The model is considerably simplified if we treat k as a scalar.

[5] The subscript k in $F_k(k,k^*)$ denotes the derivative with respect to k, the first argument. Evaluation takes place at $k=k^*$.

[6] In a more general formulation, if the equity investment is in new economy stocks, then the drift and volatility might be modeled as functions of the network size, leading potentially to another source of network externalities.

mand side of the income identity is $p_c c + p_k \frac{1}{dt} E\, dk$. A representation of relevant aspects of the agent's intertemporal optimization problem is:[7]

$$J(k_0, k_0{}^*) = \max_{c(t),x(t)} E_0 \int_0^\infty e^{-\delta t} U(c(t), k^*(t))dt \tag{4.1}$$

subject to:[8]

$$dk = \left(\frac{F(k,k^*) - p_c c}{p_k} \right) dt + \left(\frac{d\rho / \rho - F_k(k,k^*)dt}{p_k} \right) x \tag{4.2}$$

$$d\rho = \mu\rho\, dt + \sigma\rho\, d\xi \tag{4.3}$$

$$c^*(t) = c(t),\ k^*(t) = k(t),\ t \in [0,\infty),\ k(0) = k_0 \tag{4.4}$$

Combining (4.2) and (4.3), network growth is characterized by the diffusion:

$$dk = \left(\frac{F(k,k^*) + [\mu - F_k(k,k^*)]x - p_c c}{p_k} \right) dt + \frac{\sigma x}{p_k} d\xi \tag{4.5}$$

An analytical solution to the network growth equation requires a solution for the control variables c and x to be obtained in synthesized form. Applying a dynamic programming approach inspired by Bellman's principle of optimality it can be shown that optimal c satisfies:[9]

$$U_c(c,k^*) = J_k(k,k^*) p_c / p_k \tag{4.6}$$

---

[7] This representation is incomplete. Transition equations for output prices and for (implicit) factor prices need to be added. Our representation is provided for discussion of the agent's problem rather than specification of all relationships required for a solution. We emphasise: (i) the objective function; (ii) the network growth equation that arises by equating the demand and supply sides of the income identity, (iii) one source of stochastic disturbance; (iv) equality between the network size k of the representative agent and the (average) size of the network k^* relevant for spillover effects, arising in equilibrium from the "representativeness" of the agent; and (v) network size as an initial condition. In general, an analytical solution will not be achievable without severe restrictions on the functional forms of $F(k,k^*)$ and $U(c,k^*)$.

[8] In equations (4.2), (4.3) and (4.5), explicit dependence on time t is suppressed for notational clarity.

[9] We distinguish between the private and social optimum and concentrate mainly on describing the private optimum since this determines the model that we ultimately use to confront the data. Our subscript notation in (4.6) is meant to convey differentiation with respect to variables that are viewed as amenable to adjustment by the private agent. The dynamic programming approach involves solving the Bellman equation $\delta J = \max_{c,x} \left\langle U(c,k^*) + E \frac{1}{dt} dJ \right\rangle$ for J, thence constructing J_k and solving (4.6) as $c = U_c^{-1}(J_k p_c / p_k, k^*)$.

In our current work, we side-step the need to explicitly solve the stochastic intertemporal optimization, since in any event we do not have data on either *c* or *k* but rather wish to infer an implied series for *k* in particular. Instead of solving the stochastic intertemporal problem directly, we make use of characteristics of the first order condition (4.6) and parameter estimates obtained from analyzing an embedded atemporal optimization problem which contains parameters that are also relevant for the stochastic intertemporal optimization, and employ these to infer values for the network stock *k*. To describe our approach, it is useful to first write (4.6) in parametric form as:

$$c = U_c^{-1}(\lambda p_c / p_k, k^*),\ k = J_k^{-1}(\lambda, k^*) \tag{4.7}$$

where λ is the costate variable from the intertemporal optimization problem. We show later that λ $(\equiv U_c p_k / p_c \equiv J_k)$, while it is in principle unobservable, can be written as a function of current output, a variable taken as given in a conditional atemporal optimization embedded within the intertemporal problem. Our approach allows us to interpret the network growth equation (4.5) as determining network expansion from the supply side. We now turn to an examination of optimization from the demand side. In principle this generates conditional demands for c and $\frac{1}{dt}Edk$, allowing an interpretation of the network growth equation as providing the equilibrating price of network goods.

4.3.2 An Embedded Atemporal Allocation Model

The conditional demand problem is analyzed by regarding the instantaneous utility function $U(c,k^*)$ as an indirect utility function arising from optimal allocation of expenditure to ICT and other products by an instantaneous utility maximizing representative 'consumer-firm.' In this context, demand refers to both final consumption demand and intermediate demand in production. The instantaneous demand problem is to choose a vector of ICT and non-ICT products q available at prices given by the vector p, conditional on an overall budget y. The instantaneous indirect utility function $V(y,p)$ is defined by a constrained optimization given an underlying direct utility function $\hat{U}(q)$:

$$V(y,p) = \max_q \left\langle \hat{U}(q) : p'q \le y \right\rangle \tag{4.8}$$

Our emphasis at this point is on optimal demand driven permeation of ICT through the economy. For this purpose we characterize q as an *m*-dimensional vector with $q_1,\ldots,q_{m-1}$ representing various ICT products such as telecommunications, computer hardware, software and ICT services. The rest of available expenditure (effectively, the rest of GDP for the representative agent) is denoted by the aggregate residual product q_m.[10]

One link between the intertemporal problem and the embedded atemporal problem is enforced by equating total income y for the instantaneous problem with instantaneous income for the intertemporal problem. Given (4.5), we define:

$$y = F(k,k^*) + [\mu - F_k(k,k^*)]x \tag{4.9}$$

This is consistent with:

$$p'q = p_c c + p_k \tfrac{1}{dt} Edk \tag{4.10}$$

since LHS (4.10) is y at the optimum in (4.8) and RHS (4.10) follows from (4.9) and the expectation of (4.5). Define $z \equiv \tfrac{1}{dt} Edk$. Since the intertemporal problem solves for optimal c and x and thence optimal Edk, it may be regarded as generating an optimal choice of y ($\equiv p_c c + p_k z$) at the optimum. This then provides the income constraint for an alternative and equivalent description of the atemporal demand problem (4.8), viz.

$$\tilde{V}(y, p_c, p_k) = \max_{c,z} \left\langle \tilde{U}(c,z) : p_c c + p_k z \le y \right\rangle \tag{4.11}$$

where, recognizing that ICT products are embedded in other products throughout the economy, c and z simply represent an alternative aggregation of the products in q. That is $\hat{U}(q) \equiv \tilde{U}(c,z)$ by definition and hence $V(y,p) \equiv \tilde{V}(y,p_c,p_k)$ by the equivalence of the optimizations (4.8) and (4.11). Dual to (4.11) is an optimization problem defining the minimal income necessary to achieve instantaneous utility commensurate with c and z. This dual problem:

$$Y(c,z) = \min_{p_c,p_k} \left\langle p_c c + p_k z \mid \tilde{V}(p_c c + p_k z, p_c, p_k) \ge \tilde{U}(c,z) \right\rangle \tag{4.12}$$

implies a useful envelope theorem:[11]

$$Y_c(c,z) = p_c \tag{4.13}$$

Recognizing that the atemporal problem is embedded within the intertemporal problem, our approach will be to utilize the indirect utility function from the atemporal demand problem to provide the structure (functional form) for the in-

---

[10] The static expenditure allocation problem treats y as total expenditure. In the context of our model, this includes the decision to expand the network and this is the only explicit form of savings in real terms in the model. We therefore interpret y as income.

[11] Let φ denote the Lagrange multiplier for the direct optimization on RHS (4.11) and note that $\varphi \equiv \tilde{V}_y$. The direct optimization implies $\tilde{U}_c = \varphi p_c$. The envelope theorem applied to (4.12) then yields $Y_c = p_c + \omega\left[\tilde{U}_c - \tilde{V}_y p_c\right] = p_c + \omega\left[\tilde{U}_c - \varphi p_c\right] = p_c$.

stantaneous utility function for the intertemporal supply problem. That is, we begin with specification of $V(y,p)$ and equate this to $\tilde{V}(y,p_c,p_k)$ by the equivalent optimizations presented above. Then, exploiting the additive structure of the intertemporal expected utility maximizing model, we complete the linkage to the intertemporal supply problem by equating the indirect utility function from the atemporal problem with the instantaneous utility function for the intertemporal problem. That is:

$$V(y,p) \equiv \tilde{V}(y,p_c,p_k) \equiv U(c,k^*) \tag{4.14}$$

We argue that a suitable specification of an indirect utility function to track demand drivers of ICT permeation may be based upon a functional form such as:

$$V(y,p) = \left(\frac{y}{B(p)}\right)^{1-\phi}\left(\frac{y}{A(p)}\right)^{\phi-\eta} \ln\left(y/A(p)\right), \quad \eta \le \phi \le 1 \tag{4.15}$$

where

$$\ln A(p) = \sum_j \alpha_j \ln p_k + \tfrac{1}{2}\sum_j\sum_i \theta_{ji} \ln p_j \ln p_i, \quad \sum_j \alpha_j = 1, \quad \sum_j \theta_{ji} = 0, \quad \theta_{ji} = \theta_{ij} \tag{4.16}$$

$$\ln B(p) = \sum_j \beta_j \ln p_j, \quad \sum_j \beta_j = 1 \tag{4.17}$$

Here $A(p)$ is interpreted as a price index relevant to an economy in which there is little penetration of new advances in ICT (an ‘Old Economy’).[12] The indirect utility function (IUF) (4.15) is based on an underlying translog specification, which applies in the special case $\phi = \eta = 1$. In that case, only the old economy index $A(p)$ is of relevance. We free up ϕ and η, allowing a Modified (regularized) Almost Ideal Demand System (MAIDS) specification, following Cooper and McLaren (1992, 1996). MAIDS permits identification of one constant parameter, ϕ or η. To obtain identification for both ϕ and η, we allow one (country specific) constant value for η and work with a variable parameter specification for ϕ which provides for convergence of ϕ to η as an economy becomes more sophisticated.[13] The price index $B(p)$ contains weights appropriate for a sophisticated economy making extensive use of modern ICT (in fact, the weights are based on

---

[12] That is, given that $p = (p_1,\ldots,p_{m-1},p_m)'$ where prices $p_1,\ldots p_{m-1}$ apply to specialized ICT products such as telecommunications, hardware, software and IT services, and where p_m represents an aggregate price of all other products, we expect the weights in $A(p)$ to be skewed in favour of p_m.

[13] This approach is conceptually similar to that of Cooper and Madden (2005) although that work considered a variable parameter (but not country specific) specification for η.

extrapolations of the future structure of such an economy).[14] All economies are represented as somewhere between these extremes, with the parameters ϕ and η controlling the extent of ICT permeation for a given economy.

Due to lack of suitable price data we drop the second order translog term in (4.16). Obviously this reduces flexibility. To enhance flexibility by other means we introduce a lagged share specification within the main price deflator $A(p)$, allowing price index weights to be share path dependent, following the approach of Cooper and Madden (2005). In the current context, the estimated path dependency of shares plays a crucial role in the construction of the implied network stock series.

The indirect utility function which we ultimately employ in this chapter is therefore:[15]

$$V(y,p)=\left(\frac{y}{B(p)}\right)^{1-\phi}\left(\frac{y}{A(p)}\right)^{\phi-\eta}\ln\left(y/A^*(p,y,s_{-1})\right),\quad \eta\leq\phi\leq 1 \tag{4.18}$$

$$\ln A^*(p,y,s_{-1})=\ln A(p)+\sum_j \gamma_j\left[s_{j,-1}-\underline{s}_j\right]\ln(p_j/y) \tag{4.19}$$

$$\ln A(p)=\sum_j \alpha_j \ln p_j,\quad \sum_j \alpha_j=1 \tag{4.20}$$

$$\ln B(p)=\sum_j \beta_j \ln p_j,\quad \sum_j \beta_j=1 \tag{4.21}$$

The corresponding optimal share equations now become:[16]

$$s_i=\frac{\alpha_i+\gamma_i\left[s_{i,-1}-\underline{s}_i\right]+\left[(1-\phi)\beta_i+(\phi-\eta)\alpha_i\right]\ln\left(y/A^*\right)}{1+\sum_j \gamma_j\left[s_{j,-1}-\underline{s}_j\right]+(1-\eta)\ln\left(y/A^*\right)} \tag{4.22}$$

---

[14] Thus we would expect the weights in $B(p)$ to favour $p_1,\ldots,p_{m-1}$ to a greater extent than is the case for the weights in $A(p)$.

[15] In the lagged-share-dependent price index A^*, the $s_{j,-1}$ are lagged shares of product j, $\underline{s}_j$ denotes the share relevant to a reference country and time period and s_{-1} denotes the vector of one period lagged shares.

[16] These are optimal from the perspective of a private representative agent. While lagged shares and total expenditure influence the price index A^*, the private agent is assumed not to recognise this as dependent on individual decisions. A divergence between the private and social optimum can occur on this account. We are interested in estimating parameters associated with actual behaviour so we track the private optimum share allocation. We impose model consistent equilibrium by aligning y and s_{-1} in A^* with their equivalent values in other components of the specification (4.22).

4.3.3 Network Stock and Intertemporal Features Inferred From the Embedded Atemporal Model

In order to explain our approach to inferring information of an intertemporal nature (with specific reference to the size of ICT network stocks), before turning to empirical work on the atemporal allocation system (4.22) we return to the relationship between the intertemporal problem and its embedded atemporal problem. Now (4.18) implies:

$$\tilde{V}(y,p_c,p_k)=\frac{y^{1-\eta}}{\tilde{B}(p_c,p_k)^{1-\phi}\tilde{A}(p_c,p_k)^{\phi-\eta}}\ln\left(y/\tilde{A}^*(p_c,p_k,y,s_{-1})\right),\quad \eta\le\phi\le 1 \tag{4.23}$$

where $\tilde{B}$, $\tilde{A}$ and $\tilde{A}^*$ are alternative aggregations of prices such that:

$\tilde{B}(p_c,p_k)\equiv B(p)$, $\tilde{A}(p_c,p_k)\equiv A(p)$ and $\tilde{A}^*(p_c,p_k)\equiv A^*(p)$.[17] (4.24)

Now (4.23) implies:

$$\tilde{V}_y=\frac{y^{-\eta}}{\tilde{B}^{1-\phi}\tilde{A}^{\phi-\eta}}\left[(1-\eta)\ln\left(y/\tilde{A}^*\right)+1+\gamma^*\right] \tag{4.25}$$

where $\gamma^*=\sum_j\gamma_j\left[s_{j,-1}-\underline{s}_j\right]$. Given (4.13) and (4.14), we can identify $p_c\tilde{V}_y$ with U_c. Then given (4.6) we can identify $p_k\tilde{V}_y$ with J_k. Since J_k defines the co-state variable λ, substitution of $p_k\tilde{V}_y$ in (4.7) for λ gives a parametric representation of the optimal feedback decision rule for c in terms of y as the equation pair:

$$c=U_c^{-1}\left(p_c\frac{y^{-\eta}}{\tilde{B}^{1-\phi}\tilde{A}^{\phi-\eta}}\left[(1-\eta)\ln\left(y/\tilde{A}^*\right)+1+\gamma^*\right],k^*\right)=\Phi(y,k^*) \tag{4.26}$$

$$k=J_k^{-1}\left(p_k\frac{y^{-\eta}}{\tilde{B}^{1-\phi}\tilde{A}^{\phi-\eta}}\left[(1-\eta)\ln\left(y/\tilde{A}^*\right)+1+\gamma^*\right],k^*\right)=\Psi(y,k^*) \tag{4.27}$$

Our approach is to conjecture functional forms for both $\Phi(y,k^*)$ and $\Psi(y,k^*)$ and to impose structure on them in order to enforce compatibility of the

---

[17] The vector $p\equiv(p_1,\ldots,p_m)$ is the set of prices required in principle for the empirical work. In fact due to data limitations we proceed to the empirical work without price information by making use of Cobb-Douglas-like specifications which treat price elasticities as independent of the price level. Thus the relationship between p_c and p_k on the one hand and the vector p of disaggregated (direct ICT product and other - not direct ICT, though ICT-embedded) product prices on the other, is of conceptual rather than of empirical relevance at this stage of our research.

implied $U(c,k^*)$ with $\tilde{V}(y,p_c,p_k)$ – the latter pre-specified by (4.23) – as required by (4.14). Consider the functional forms:[18]

$$\Phi(y,k^*) = \Pi\left(y/A\right)^{\chi_1}\left(y/B\right)^{\chi_2}\left(k^*\right)^{\chi_3} \tag{4.28}$$

$$\Psi(y,k^*) = \left(y/A\right)^{\kappa_1}\left(y/B\right)^{\kappa_2}\left(k^*\right)^{\kappa_3} \tag{4.29}$$

for some (yet to be specified) price-dependent parameters Π, χ_1, χ_2 and χ_3 in the case of the $c = \Phi(y,k^*)$ function and κ_1, κ_2 and κ_3 in the case of the $k = \Psi(y,k^*)$ function.

Since $k = k^*$ in equilibrium, (4.29) implies:

$$k^* = \left(y/A\right)^{\theta_1}\left(y/B\right)^{\theta_2} = \tilde{\Psi}(y) \tag{4.30}$$

where $\theta_1 = \kappa_1/(1-\kappa_3)$ and $\theta_2 = \kappa_2/(1-\kappa_3)$. In principle, (4.30) provides a means of constructing a network stock series k from observations on output y provided estimates of the parameters θ_1 and θ_2 are available together with the price indexes A and B. The price index parameters are in principle available from estimation of the atemporal demand system (4.22). However, estimates of θ_1 and θ_2 are more difficult to obtain. Our approach is to impose further structure on the solution forms specified in (4.28) and (4.29). First, in (4.28) we enforce zero degree homogeneity on Π, as discussed below. Second, in (4.29) we set $\kappa_1 = \kappa_3 = \frac{1}{2}$.[19] Third, we impose an assumption of 'conditional balanced growth' on (4.28) and (4.29) together, also discussed below. Together these conjectured structural specifications allow the sum $\theta_1 + \theta_2 = 1 + \theta$ to be identified as a variable parameter from the atemporal estimation. We define conditional balanced growth as constancy of c/k in the absence of changes in A/B. Now given (4.30), (4.28) implies:

$$c = \Pi\left(y/A\right)^{\chi_1+\chi_3}\left(y/B\right)^{\chi_2+\chi_3\theta} = \tilde{\Phi}(y) \tag{4.31}$$

and hence from (4.31) and (4.30):

$$c/k^* = \Pi\left(y/A\right)^{\chi_1+\chi_3-1}\left(y/B\right)^{\chi_2+(\chi_3-1)\theta} \tag{4.32}$$

and conditional balanced growth requires:

---

[18] To avoid notational clutter, since (4.24) holds, where convenient we write $A(p)$ and $\tilde{A}(p_c,p_k)$ both simply as A, and similarly for B and A^*.

[19] The conjectured solution (4.29) is still quite general because we impose no restrictions on κ_2. Effectively, the restriction imposes an implied equivalent normalization in (4.30), viz. $\theta_1 = 1$. We then write the remaining unknown θ_2 simply as θ.

$$\chi_1 + \chi_2 + (\chi_3 - 1)(1+\theta) = 0 \tag{4.33}$$

However, (4.31) also implies:

$$\ln c = \ln \Pi + \left[\chi_1 + \chi_2 + \chi_3(1+\theta)\right]\ln y - (\chi_1 + \chi_3)\ln A - (\chi_2 + \chi_3\theta)\ln B \tag{4.34}$$

Now define:

$$\ln \Pi = (\chi_1 + \chi_3)\ln A + (\chi_2 + \chi_3\theta)\ln B - \ln \bar{A} \tag{4.35}$$

where

$$\ln \bar{A} = \left[\ln A + \sum_j \gamma_j (s_{j,-1} - \underline{s}_j)\ln p_j\right], \; \gamma^* = \sum_j \gamma_j \left[s_{j,-1} - \underline{s}_j\right] \tag{4.36}$$

For arbitrary χ_1, χ_2 and θ define χ_3 such that $\chi_3 = \left[1 + \gamma^* - (\chi_1 + \chi_2)\right]/(1+\theta)$. This is equivalent to the restriction:

$$\chi_1 + \chi_2 + \chi_3(1+\theta) = 1 + \gamma^* \tag{4.37}$$

Given (4.35) and (4.36), (4.37) ensures Π is homogeneous of degree zero in prices.[20] Then (4.37) and (4.33) imply:

$$\theta = \gamma^* \tag{4.38}$$

Using (4.38) and the normalization $\theta_1 = 1$, we rewrite (4.30) as:

$$k^* = (B/A)(y/B)^{1+\gamma^*} \tag{4.39}$$

The importance of (4.39) is that it allows construction of the network stock k from information on the income series y and from the price indexes A and B and the variable parameter γ^*, all of which can in principle be constructed from parameters estimable from the solution to the atemporal optimization problem.[21]

---

[20] Zero degree homogeneity of Π in prices is necessary, in view of (4.34), to allow y to be defined as nominal while c is real.

[21] Other variables associated with the intertemporal optimization may be obtained by an extension of our approach. For example, combining (4.34)-(4.37) and noting (4.19), we obtain a means of constructing c, viz. $\ln c = \ln\left(y/A^*\right)$. Additionally, from (4.19),

$$\ln(y/A) = \left\{\ln\left(y/A^*\right) + \sum \gamma_j \left[s_{j,-1} - \underline{s}_j\right]\ln\left(p_j/A\right)\right\}/(1+\gamma^*), \text{ so that}$$

$y/A = (\Gamma c)^{1/(1+\gamma^*)}$ where $\ln \Gamma = \sum_j \gamma_j \left[s_{j,-1} - \underline{s}_j\right]\ln(p/A)$. Constructing $V(y,p)$ according to (4.18) using (4.39) and these results we obtain, in view of (4.14) an equivalent expression for the instantaneous utility function:

4.4 Empirical Specifications and Results

We estimate a variant of the share equation system (4.22) from pooled cross country and time series data. In our empirical work, we treat η in (4.22) as a country specific parameter.[22] We do not have a consistent time series on relevant prices across all the countries in our data set and so we cannot include the price function A^* explicitly in our estimation. We proxy (4.22) by:[23]

$$s_{i,c,t}=\frac{\alpha_i+\gamma_i\left[s_{i,c,t-1}-s_{i,55,0}\right]+\left[(1-\phi_{c,t})\beta_i+(\phi_{c,t}-\eta_c)\alpha_i\right]\ln\tilde{y}_{c,t}}{1+\sum_j\gamma_j\left[s_{j,c,t-1}-s_{j,55,0}\right]+(1-\eta_{c,t})\ln\tilde{y}_{c,t}} \tag{4.40}$$

where $\tilde{y}_{c,t}$ is GDP per capita for country c in time t, measured in US dollars and normalized as described below.[24]

The 'ICT-permeation' parameter ϕ is treated as variable over time, approaching the relevant η for each country asymptotically, so that the share equations (4.40) converge over time to:

Viz. $U(c,k^*)=u_0k^{*(1-\phi)/(1+\gamma^*)}\,c^{(\phi-\eta)/(1+\gamma^*)}\ln c$ where $u_0=\Gamma^{(\phi-\eta)/(1+\gamma^*)}\left(A/B\right)^{(1-\phi)/(1+\gamma^*)}$. Further, this implies $U_c=u_0k^{*(1-\phi)/(1+\gamma^*)}\,c^{(\phi-\eta)/(1+\gamma^*)-1}\left[\left(\frac{\phi-\eta}{1+\gamma^*}\right)\ln c+1\right]$. Now noting that from (4.25), $p_cV_y=\frac{p_cy^{-\eta}}{B^{1-\phi}A^{\phi-\eta}}\left[(1-\eta)\ln\left(y/A^*\right)+1+\gamma^*\right]$ we can equate U_c with p_cV_y, and utilize (4.39) and the expression for u_0 to obtain an expression for p_c as a function of elements of the price vector p and parameters from the atemporal optimization, viz.:

$$p_c=\Gamma^{(\phi-\eta)/(1+\gamma^*)}A^{\phi-\eta}y^{(1-\phi)+\eta}c^{(\phi-\eta)/(1+\gamma^*)-1}\frac{(\phi-\eta)\left[\ln\left(y/A\right)-\ln\Gamma/(1+\gamma^*)\right]+1}{(1-\eta)\left[(1+\gamma^*)\ln\left(y/A\right)-\ln\Gamma\right]+1+\gamma^*}.$$

[22] In principle η could also vary over time. With a small number of time series observations, in this study we have opted to treat η as constant over time.

[23] Equation (4.40) depicts the i^{th} expenditure share equation for country c at time t.

[24] Because the function A^* contains many of the parameters in (4.40), use of y/A^* in place of $\tilde{y}$ would clearly be preferable as it would allow additional cross-equation parameter restrictions to be utilized to aid precision in estimation and allow identification of parameters associated with more complex adjustment processes. This task awaits further data development.

$$s_{i,c,t} = \frac{\alpha_i + \gamma_i \left[s_{i,c,t-1} - s_{i,55,0} \right] + (1-\eta_c)\beta_i \ln \tilde{y}_{c,t}}{1 + \sum_j \gamma_j \left[s_{j,c,t-1} - s_{j,55,0} \right] + (1-\eta_c) \ln \tilde{y}_{c,t}} \tag{4.41}$$

As $\tilde{y}$ grows indefinitely, these shares further converge to β_i. In the base case, i.e., for the reference year and reference country – defined by construction as the lowest per capita income observation in our data set – viz. country 55 (Vietnam) in 1994, the shares equate to α_i.[25] In order to economize on parameters for estimation we impose a parsimonious specification for the time variation in ϕ, viz.:

$$\phi_{c,t} = \phi_{55,1} + (\eta_c - \phi_{55,1}) REL_{c,t} \tag{4.42}$$

where[26]

$$REL_{c,t} = \left\{ \left[{(ict/gdp)_{c,t-1}} \Big/ {(ict/gdp)_{55,0}} \right] - 1 \right\} \Big/ \left\{ \left[{(ict/gdp)_{c,t-1}} \Big/ {(ict/gdp)_{55,0}} \right] + 1 \right\} \tag{4.43}$$

In table 4.1, nine years (1993-2001) of ICT expenditure data for five selected countries – Australia, China, South Africa, the United States and Vietnam – are highlighted.[27] These are country numbers 2, 9, 45, 53 and 55 of a 55 country data set. In Table I, ICT is split into four categories: Telecommunications, IT Hardware, IT Software and IT Services. Annual expenditures on these categories of ICT are given as shares of GDP in columns headed S1 to S4 respectively. All other components of GDP (including some minor ICT expenditures) are aggregated in a fifth share, denoted as S5.

---

[25] The data is normalized so that $\tilde{y} = 1$ and all implied prices and price indexes are unity for Vietnam in 1994. The $s_{j,55,0}$ are the Vietnamese expenditure shares for 1993. These were denoted by $\underline{s}_j$ in the derivations leading to (4.22).

[26] Specification (4.43) is a parsimonious variant of an initially considered specification

$$REL_{c,t} = \left\{ \left[{(ict/gdp)_{c,t-1}} \Big/ {(ict/gdp)_{55,0}} \right] - 1 \right\} \Big/ \left\{ \left[{(ict/gdp)_{c,t-1}} \Big/ {(ict/gdp)_{55,0}} \right] - 1 + \varepsilon_c \right\}$$

which would have allowed country specific closure of the gap between ϕ and η but at the expense of too many parameters for efficient estimation. After experimentation, we set $\varepsilon_c = 2$ for all countries $c = 1,\ldots,55$.

[27] The full 55 country data set is available on request. The first year for each country is utilized to construct lagged effects and estimation takes place over the 8 years 1994-2001.

Table 4.1. Initial Data for Five Selected Countries

obs	c-num	time	year	gdpcap	s1	s2	s3	s4	s5
10	2	0	1993	16935.5	0.03413	0.01268	0.00375	0.00798	0.94147
11	2	1	1994	18977.9	0.03552	0.01403	0.00394	0.00787	0.93863
12	2	2	1995	20055.4	0.03662	0.01385	0.00402	0.00724	0.93828
13	2	3	1996	22141.8	0.03805	0.01353	0.00435	0.00757	0.93650
14	2	4	1997	22016.8	0.04131	0.01415	0.00495	0.00936	0.93024
15	2	5	1998	19530.0	0.04627	0.01543	0.00546	0.01180	0.92104
16	2	6	1999	20837.6	0.04446	0.01650	0.00578	0.01270	0.92055
17	2	7	2000	19813.6	0.04665	0.01623	0.00627	0.01384	0.91700
18	2	8	2001	18119.7	0.05221	0.01595	0.00774	0.01558	0.90851
73	9	0	1993	512.5	0.01125	0.00350	0.00010	0.00018	0.98498
74	9	1	1994	445.5	0.01366	0.00623	0.00031	0.00038	0.97943
75	9	2	1995	572.4	0.02140	0.00577	0.00029	0.00036	0.97218
76	9	3	1996	651.6	0.02283	0.00624	0.00049	0.00027	0.97017
77	9	4	1997	719.4	0.02096	0.00779	0.00059	0.00033	0.97032
78	9	5	1998	747.6	0.03103	0.00860	0.00057	0.00043	0.95937
79	9	6	1999	795.8	0.03528	0.00966	0.00070	0.00058	0.95378
80	9	7	2000	851.9	0.03675	0.01318	0.00099	0.00087	0.94821
81	9	8	2001	924.6	0.03767	0.01432	0.00128	0.00130	0.94543
397	45	0	1993	3269.9	0.02197	0.01019	0.00255	0.00724	0.95806
398	45	1	1994	3341.1	0.02238	0.01153	0.00286	0.00814	0.95509
399	45	2	1995	3679.0	0.02681	0.01005	0.00300	0.00657	0.95357
400	45	3	1996	3421.9	0.02885	0.01124	0.00436	0.00829	0.94726
401	45	4	1997	3440.0	0.02783	0.01152	0.00449	0.00987	0.94629
402	45	5	1998	2989.6	0.02770	0.01669	0.00571	0.01186	0.93804
403	45	6	1999	3080.9	0.03005	0.01795	0.00628	0.01448	0.93124
404	45	7	2000	2970.9	0.03346	0.01635	0.00683	0.01300	0.93036
405	45	8	2001	2920.6	0.03586	0.01598	0.00802	0.01491	0.92523
469	53	0	1993	25742.4	0.02729	0.01219	0.00497	0.01241	0.94314
470	53	1	1994	26886.5	0.02785	0.01282	0.00539	0.01268	0.94126
471	53	2	1995	28246.6	0.02767	0.01422	0.00547	0.01320	0.93943
472	53	3	1996	29333.8	0.02692	0.01655	0.00601	0.01377	0.93675
473	53	4	1997	31161.1	0.02636	0.01660	0.00647	0.01486	0.93571
474	53	5	1998	32463.4	0.02633	0.01817	0.00743	0.01653	0.93154
475	53	6	1999	34048.7	0.02611	0.01821	0.00807	0.01725	0.93035
476	53	7	2000	36125.9	0.02536	0.01663	0.00914	0.01856	0.93031
477	53	8	2001	37010.1	0.02585	0.01323	0.00939	0.01937	0.93217
487	55	0	1993	187.5	0.01193	0.00879	0.00045	0.00075	0.97809
488	55	1	1994	221.4	0.01639	0.00844	0.00043	0.00081	0.97392
489	55	2	1995	277.8	0.02494	0.00809	0.00039	0.00083	0.96575
490	55	3	1996	324.4	0.02874	0.00923	0.00049	0.00082	0.96072
491	55	4	1997	354.0	0.03406	0.01172	0.00081	0.00133	0.95207
492	55	5	1998	352.1	0.03879	0.00611	0.00058	0.00084	0.95368
493	55	6	1999	362.3	0.05017	0.00672	0.00060	0.00088	0.94164
494	55	7	2000	376.9	0.05286	0.00782	0.00063	0.00080	0.93790
495	55	8	2001	391.0	0.05391	0.00836	0.00073	0.00095	0.93606

In all countries, S5 is falling slowly but inexorably. Extrapolating this, our model predicts a long run fall in S5 and gains in all four ICT share categories as GDP rises asymptotically. This reflects a shift in input shares which is suggestive either of a non-linear expansion path, i.e., non-homotheticity of the underlying preference function of the representative agent, or of substitution towards lower priced ICT products. Because we are working without price data, and because some of the shifting pattern undoubtedly represents a preference twist in favor of innovative ICT products, in order to pool the data across countries with widely varying GDP, we employ a specification that takes this non-homotheticity into account.

Table 4.2 presents the core parameter estimates and overall equation fit statistics.

Table 4.2 . Core Parameter Estimates

		Weights for Old Economy price index		
	Parameter	estimate	standard error	*t* statistic
	α_1	0.0134		constrained by addition
	α_2	0.0073	0.0003	24.8
	α_3	0.0004	0.0001	5.3
	α_4	0.0006	0.0001	4.3
	α_5	0.9784	0.0008	1151.7
		Weights for New Economy price index		
	Parameter	estimate	standard error	*t* statistic
	β_1	0.0195		constrained by addition
	β_2	0.0335	0.0039	8.5
	β_3	0.0038	0.0011	3.4
	β_4	0.0053	0.0020	2.7
	β_5	0.9378	0.0114	82.2
		Influence of own lagged shares		
	Parameter	estimate	standard error	*t* statistic
	γ_1	1.0673	0.0414	25.8
	γ_2	0.8437	0.0412	20.5
	γ_3	1.1358	0.0382	29.8
	γ_4	1.1565	0.0378	30.6
	γ_5	1.9037	0.2983	6.4
		Summary Statistics		
	eqn 1	eqn 2	eqn 3	eqn 4
DW	1.6072	1.8302	2.1329	1.5775
R-squared	0.9083	0.9012	0.9597	0.9713

The estimated values for α_1 and β_1 are determined residually by adding up, *viz.* the expenditure shares must sum to unity. The α_i are predicted shares for the reference country and time, *viz.* Vietnam in 1994. The β_i represent the estimated limiting ICT expenditure shares associated with the technology or preferences when GDP becomes very large and when this is associated with new ICT fully 'permeating' the economy—at this point ϕ falls to η. National e-preparedness may be thought of as measured by $1-\phi$, the weight given to real income in the indirect utility function when deflated by the 'New Economy' GDP deflator $B(p)$. In this model, $B(p)$ is a Cobb-Douglas index using the β_i weights. The fit of the share equations and apparent randomness of residuals is improved markedly by inclusion of the γ_i (lagged share) parameters. Comparison of α_5 with β_5 indicates that ICT expenditure as a share of GDP rises from around 2.2% of GDP for a low-income (Old Economy) country to about 6.2% in the long run for a high-income (New Economy) country. This is really only an estimate of the direct influence of ICT, since much of ICT is embedded in other products which utilize ICT as an intermediate input.

The country-specific η parameters are reported in table 4.3. In estimation over such a large range of countries with apparent varying data quality, we found it necessary to group certain countries to focus on a common η parameter in some cases. Table 4.3 first presents individual estimates of η for those countries where the data seemed strong enough to support individual estimates. At the base of the table some grouped estimates, typically for some very low income countries, are presented.

Table 4.3. Country Specific Parameter Estimates

	Country specific parameter (η)				
Identifier	parameter	estimate	std. error	*t* statistic	country
1	C1	0.9910	0.0168	59.0	Argentina
2	C2	0.9730	0.0086	113.6	Australia
3	C3	0.9837	0.0092	107.4	Austria
4	C4	0.9708	0.0089	108.8	Belgium
5	C5	0.9539	0.0174	54.8	Brazil
7	C7	0.9688	0.0086	112.8	Canada
8	C8	0.9802	0.0141	69.7	Chile
9	C9	0.7811	0.0738	10.6	China
11	C11	0.9573	0.0144	66.7	Czech Republic
12	C12	0.9786	0.0079	123.9	Denmark
14	C14	0.9771	0.0090	108.4	Finland
15	C15	0.9880	0.0072	137.0	France
16	C16	0.9823	0.0083	118.0	Germany
17	C17	0.9949	0.0119	83.5	Greece
18	C18	0.9824	0.0083	118.3	Hong Kong
19	C19	0.9720	0.0128	76.0	Hungary
21	C21	0.8666	0.1972	4.4	Indonesia
22	C22	0.9668	0.0129	75.2	Ireland
23	C23	0.9674	0.0098	98.5	Israel
24	C24	0.9884	0.0100	98.6	Italy
25	C25	0.9824	0.0085	115.4	Japan
26	C26	0.8874	0.0374	23.7	Korea
27	C27	0.9119	0.0251	36.3	Malaysia
28	C28	0.9751	0.0255	38.3	Mexico
29	C29	0.9775	0.0078	125.6	Netherlands
30	C30	0.9733	0.0076	127.4	New Zealand
31	C31	0.9790	0.0080	122.4	Norway
37	C37	0.9711	0.0237	41.0	Poland
38	C38	0.9699	0.0137	70.9	Portugal
41	C41	0.9991	0.0201	49.7	Saudi Arabia

42	C42	0.9587	0.0105	91.1	Singapore
45	C45	0.9463	0.0162	58.5	South Africa
46	C46	0.9878	0.0122	80.9	Spain
47	C47	0.9786	0.0073	134.6	Sweden
48	C48	0.9737	0.0069	141.3	Switzerland
49	C49	0.9814	0.0153	64.3	Taiwan
51	C51	0.9850	0.0253	38.9	Turkey
52	C52	0.9778	0.0076	128.4	UK
53	C53	0.9717	0.0073	132.7	USA
		Grouped Country Results Other Eastern Europe Group			
6	CEE	0.9634	0.0138	70.0	Bulgaria
39	CEE				Romania
40	CEE				Russia
43	CEE				Slovakia
44	CEE				Slovenia
		Other Middle East Group			
13	CME	0.9830	0.0623	15.8	Egypt
35	CME				Other Middle East
		Other Latin America Group			
10	CLA	0.9752	0.0117	83.2	Colombia
34	CLA				Other Latin America
54					Venezuela
		Other Asia and Pacific Group			
20	CAP	0.9174	0.0370	24.8	India
32	CAP				Other Asia Pacific
33	CAP				Other East Asia
36	CAP				Philippines
50	CAP				Thailand
55	CAP				Vietnam

Finally, table 4.4 provides estimates of some constructed variables and of key estimated variable parameters. The most important results are the estimates of the network capital stock, reported in the second last column of table 4.4, i.e., column 9. The last column, 10, shows the annual growth in the network stock estimate. The capital stock is calculated using a simplified variant of (4.39). Because we do not have consistent price series, we cannot at this point construct the price indexes A and B, so our estimates of k are based on the proxy formula $k = \tilde{y}^{1+\gamma^*}$. Clearly these estimates depend critically on $\tilde{y}$ which is itself given in column 4 of table 4.4, and on the estimates of γ^* given in column 5. The γ^* are calculated from the γ_i (lagged share) parameters – themselves presented in table 4.2, according to the formula $\gamma_{c,t}^* = \sum_j \gamma_j \left[s_{j,c,t-1} - s_{j,55,0} \right]$. Table 4.4 also presents the calculated series for $REL_{c,t}$ (column 6), phi (column 7), and the implied capital/output ratio using the network stock k and our normalized output measure $\tilde{y}$. Note that many series are indexes, normalized to unity for Vietnam in 1994.

Table 4.4 Derived Data, Variable Parameter Estimates and Constructed Series

obs	c-num	year	$\tilde{y}$	γ^*	rel	ϕ	$k/\tilde{y}$	k	growth
11	2	1994	85.71	-0.031	0.455	0.976	0.873	74.79	0.00
12	2	1995	90.57	-0.033	0.474	0.976	0.861	77.96	4.24
13	2	1996	100.00	-0.034	0.476	0.976	0.857	85.66	9.88
14	2	1997	99.43	-0.035	0.487	0.976	0.851	84.66	-1.17
15	2	1998	88.20	-0.040	0.522	0.975	0.835	73.69	-12.97
16	2	1999	94.11	-0.048	0.566	0.975	0.804	75.71	2.75
17	2	2000	89.48	-0.048	0.568	0.975	0.804	71.99	-4.92
18	2	2001	81.83	-0.051	0.582	0.975	0.798	65.31	-9.27
74	9	1994	2.01	0.007	-0.187	1.015	1.005	2.02	0.00
75	9	1995	2.59	0.002	-0.032	0.984	1.002	2.59	28.09
76	9	1996	2.94	-0.004	0.119	0.955	0.995	2.93	13.13
77	9	1997	3.25	-0.006	0.153	0.948	0.993	3.23	10.12
78	9	1998	3.38	-0.006	0.151	0.948	0.992	3.35	3.88
79	9	1999	3.59	-0.016	0.299	0.919	0.980	3.52	5.14
80	9	2000	3.85	-0.021	0.357	0.908	0.973	3.74	6.22
81	9	2001	4.18	-0.026	0.405	0.898	0.964	4.02	7.52
398	45	1994	15.09	-0.016	0.314	0.968	0.957	14.43	0.00
399	45	1995	16.62	-0.019	0.344	0.967	0.948	15.75	9.12
400	45	1996	15.45	-0.020	0.359	0.967	0.946	14.63	-7.13
401	45	1997	15.54	-0.025	0.413	0.965	0.933	14.49	-0.93
402	45	1998	13.50	-0.026	0.420	0.965	0.934	12.62	-12.95
403	45	1999	13.91	-0.034	0.477	0.963	0.915	12.73	0.88
404	45	2000	13.42	-0.040	0.517	0.962	0.902	12.11	-4.87
405	45	2001	13.19	-0.040	0.521	0.962	0.902	11.90	-1.75
470	53	1994	121.42	-0.029	0.444	0.975	0.872	105.83	0.00
471	53	1995	127.57	-0.030	0.457	0.975	0.863	110.14	4.07
472	53	1996	132.48	-0.032	0.469	0.975	0.855	113.25	2.83
473	53	1997	140.73	-0.035	0.485	0.975	0.842	118.49	4.63
474	53	1998	146.61	-0.036	0.492	0.975	0.838	122.81	3.64
475	53	1999	153.77	-0.039	0.515	0.975	0.821	126.26	2.81
476	53	2000	163.15	-0.040	0.521	0.975	0.815	133.05	5.38
477	53	2001	167.14	-0.040	0.522	0.975	0.817	136.52	2.61
488	55	1994	1.00	0.000	0.000	0.978	1.000	1.00	0.00
489	55	1995	1.26	-0.003	0.087	0.973	0.999	1.25	25.35
490	55	1996	1.47	-0.010	0.220	0.965	0.996	1.46	16.42
491	55	1997	1.60	-0.015	0.284	0.961	0.993	1.59	8.81
492	55	1998	1.59	-0.022	0.373	0.955	0.990	1.57	-0.89
493	55	1999	1.64	-0.020	0.358	0.956	0.990	1.62	2.96
494	55	2000	1.70	-0.030	0.454	0.951	0.984	1.68	3.40
495	55	2001	1.77	-0.033	0.478	0.949	0.981	1.73	3.43

We draw special attention to the calculated network capital stock series. Because we lack price data, our constructed series needs to be interpreted cautiously. Nevertheless the results are indicative and suggest that, in the interests of balanced world development, there may be a good deal of cause for concern with the take-up of ICT innovation. Concentrating on the five countries highlighted in the tables of results, it can be seen that while the implied ICT network stocks have grown rapidly for China and Vietnam (from very low bases) and have also continued to grow steadily in the USA, they have fallen slightly overall in Australia and more extensively and worryingly in South Africa. Thus the picture is mixed. Some developing countries are apparently catching up and are likely to enjoy the benefits of the network externalities evident with ICT. Other developing countries (and some advanced countries) appear to be in a more problematic position with respect to network growth that does not auger well for the future.

4.5 Conclusions

Much of the literature on growth-investment relationships and issues associated with the gap between rich and poor countries has resulted in contradictory findings. The question arises, in an era of significant technological revolution, whether the benefits are being spread, or are capable of being spread, in such a fashion as to ameliorate the substantial economic North-South imbalance. The model of this chapter seeks to address these issues by providing a sound theoretical basis for specification of estimating forms intended to help untangle the myriad influences.

An appropriate concentration on network effects that operate through increasing returns in production and via consumption requires modeling in an integrated fashion that incorporates another factor capable of creating a critical mass, viz., the pervasiveness of ICT. When an economy is ‘sophisticated’, and the size of the ICT network stock is a measure of this, then any positive economic impact from new application success is likely to be enhanced. For a more sophisticated economy (more typical of the North), ‘good news’ effects may be magnified through many sectors due to the pervasiveness of ICT, while ‘bad news’ effects may be more readily absorbed. The reverse is true for less sophisticated economies. Thus indicators of network sophistication are an important means of determining the ‘value’ of ICT to an economy. A question that arises is whether low income countries are getting the same ‘value’ as high income countries are able to obtain from innovations in ICT.

In the absence of direct information on the true degree of permeation of ICT, our approach uses indirect evidence, associated with information on components of ICT as GDP expenditure shares, to infer the existence and possible growth (or decline) of an underlying ICT network stock, posited to be a source of positive consumption and production externalities. The approach uses the postulate of rationality and hence of assumed consistency between atemporal and intertemporal decisions to draw inferences on the natures of the intertemporal choices and trade-offs facing both developing and advanced countries from an examination of more readily available atemporal allocation data. The approach seems promising and points clearly to the need for a greater volume and quality of data. Like any modeling effort, the results will only be as good as the data.

References

Bosworth BP, Triplett, JE (2001) What's new about the new economy? IT, economic growth and productivity. International Productivity Monitor 2:19-30

Bresnahan TF, Brynjolfsson E, Hitt LM (2002) Information technology, workplace organization and the demand for skilled labor: Firm level evidence. Quarterly Journal of Economics 117: 339-76

Bresnahan TF, Trajtenberg M (1995) General purpose technologies: Engines of growth? Journal of Econometrics 65:83-108

Bridges.org (2001) Spanning the digital divide: understanding and tackling the issues. http://www.bridges.org

Brynjolfsson E, Hitt LM (1995) Information technology as a factor of production: The role of differences among firms. Economics of Innovation and New Technology 3:183-99

Brynjolfsson E, Hitt LM, Yang S (2000) Beyond computation: Information technology, organization transformation and business performance. Journal of Economic Perspectives 14:23-48

Cohen SC, DeLong JB, Zysman J (2000) Tools for thought: What is new and important about the "E-conomy." Berkeley Roundtable on the International Economy, University of California at Berkeley

Cooper R, Madan D, McLaren K (1995) Approaches to the solution of stochastic intertemporal consumption models. Australian Economic Papers 34:86-103

Cooper R, Madden G (2005) A "model consistent" approach to productivity measurement. In: Productivity Commission, Quantitative tools for microeconomic policy analysis. Australian Government Publishing Service, Canberra

Cooper R, McLaren K (1996) A system of demand equations satisfying effectively global regularity conditions. Review of Economics and Statistics 78:359-64

Cooper R, McLaren K (1992) An empirically oriented demand system with improved regularity properties. Canadian Journal of Economics 25: 652-67

Daveri F (2001) Information technology and growth in Europe. Working Paper, University of Parma, Parma.

Dewan S, Kraemer KL (1998) International dimensions of the productivity paradox. Communications of the ACM 41:56-62

Dewan S, Kraemer KL (2000) Information technology and productivity: Evidence from country level data. Management Science 46:548-62

Gordon R (2000). Does the 'new economy' measure up to the great inventions of the past? Journal of Economic Perspectives 14:49-76

ITU (2002) World telecommunication indicators database. ITU, Geneva

Jorgenson DW, Stiroh KJ (1995) Computers and growth. Economics of Innovation and New Technology 3:295-316

Jorgenson DW, Stiroh KJ (1999) Information technology and growth. American Economic Review 89:109-15

Jorgenson DW, Stiroh KJ (2000) Raising the speed limit: US economic growth in the information age. Brookings Papers on Economic Activity 1: 2000, Brookings Institution, Washington DC

Jorgenson DW, Wessner CW (eds) (2002) Measuring and sustaining the new economy—report of a workshop, National Academy Press, Washington DC

Kenny C (2003) World bank telecommunications sector activity: The 1990s and beyond. In: Madden G (ed) World telecommunications markets. The international handbook of telecommunications economics, volume III. Edward Elgar, Cheltenham, pp 250-65

Kraemer KL, Dedrick, J (2001) Information technology and productivity: results and implications of cross-country studies. In: Pohjola M (ed) Information technology and economic development. Oxford University Press, Cambridge

Lanvin B (2003) International efforts to bridge the digital divide: A global imperative" In: Madden G (ed.) The international handbook of telecommunications economics, volume III. Edward Elgar, Cheltenham pp. 241-9.

Litan RE, Rivlin, AM (2001a) Projecting the economic impact of the internet. American Economic Review 91: 313-22.

Litan RE, Rivlin, AM (2001b) The Economy and the internet: What lies ahead?" In: Litan RE, Rivlin AM (eds*)* The economic payoff from the internet revolution. Brookings Institution Press, Washington, DC, pp 1-28

OECD (2000) A new economy? The changing role of innovation and information technology in growth. OECD, Paris

Onliner S, Sichel D (2000) The resurgence of growth in the late 1990s: Is information technology the story? Journal of Economic Perspectives 14:3-22

Panagariya A (2000) E-commerce, WTO and developing countries. The World Economy 23:959-78

Pohjola M (2001) Information technology and economic growth: A cross-country analysis. In: Pohjola M (ed) Information technology and economic development. Oxford University Press, Cambridge, pp 242-56

Pohjola M (2002) The new economy and growth and development. Oxford Review of Economic Policy 18: 380-96

Pritchett L (1997) Divergence, big time. Journal of Economic Perspectives 11:3-17

Salvatore D (2003) The new economy and growth: Editor's introduction. Journal of Policy Modeling, 25:431-4

Schreyer P (1999) The contribution of information and communication technology to output growth. Statistical working paper 99:4. OECD, Paris

Stiroh KJ (2002) Are ICT spillovers driving the new economy? Review of Income and Wealth 48:33-57

5 The Aging of the Labor Force and Globalization

Ronald W. McQuaid

Employment Research Institute, Napier University, Craiglockhart Campus, Edinburgh, Scotland, UK

5.1 Introduction

The rapid aging of the population structure across most developed countries is one of the main challenges facing their economies and governments (CEC 2004a).[1] These changes will have significant impacts upon pension and healthcare costs, the demand for different products and services, the quantity and composition of labor forces, productivity levels etc. and hence national and regional economic growth and living standards (Mackeller 2003; Samorodov 1999). Indeed, Bloom and Canning (2003) argue that having a younger population with an increase in working age per non-working age population may be one of a number of factors in the huge growth of 'Tiger' economies, such as Ireland in the 1990s, and the reverse of having an aging structure may have a negative impact on growth. For example, Lisiankova and Wright (2005) argue that EU growth rates will fall significantly due to population aging. Unemployment may also be influenced by changes in the age structure as Katz and Kreuger (1999) estimate that 0.4 percentage points of the total fall in USA NAIRU from the mid-1980s to mid-1990s was due to the baby boom maturing. For those countries with at younger age structure these changes may provide opportunities for increased development using their own growing labor supplies to help meet global consumption, but they may also be a source of migrants to countries experiencing a reduction in their labor force, as part of labor market globalization.

While a rapid and long-term decline in labor forces in the EU is projected to start around 2010 and to continue until around 2050, as discussed below, many other major industrial countries will also face similar labor force pressures in the next few decades (GROS 2002), while Japan has one of the aged population structures (Endo and Katayama 1999). By 2050 all the EU Member States will be in

---

[1] Aging in this context means the move to a high support ratio in terms of those aged 60+ divided by those aged 15-59 years.

population decline and all will have been aging for some time (CEC 2004b; UN 2005). These changes are not homogenous over space, and at the regional and sub-regional levels within countries the aging population and workforce structural changes can be greater than those between countries (Hollywood et al. 2003). Although there is a long history of regional models of demographic change, the full implications of the aging of labor forces have rarely been fully analyzed (Batey and Madden 1998). Hence, this chapter seeks to discuss some of the issues concerning the aging of labor forces that increasingly need to be considered in regional policies and modeling.

An aging population structure will directly and indirectly influence the demand for products and services (and hence the demand for labor). For example: demand for medical care should increase overall; saving rates may decline as older people use their capital for consumption, especially if returns from private pensions are lower than expected, although alternatively savings may rise for those planning their financial needs over longer retirement periods; and there may be 'wealth' effects as inheritances are passed down among smaller sized families etc. In addition there are profound implications of an aging population for pensions and public finances through the increasing dependency rate of people needing support.

This chapter focuses upon how aging will directly affect labor supply, particularly in terms of the size, participation rate and productivity of the labor force, which are all related to the changing age composition of the population. Changes to the general size of the labor force will depend upon demographic characteristics and economic participation rates for each age cohort. These in turn will be influenced by factors such as migration and policies or incentives encouraging withdrawal from, continuation in, or entry into the labor force. Population aging may lead to an increase in international migration, but also a decrease in internal migration between regions (Greenwood 1997). Hence the responses of national or regional economies to an aging workforce or a potential fall in labor force size will vary. One effect of a declining working age population may be labor shortages (Robson 2001), but these may be countered by rising wages, to reflect the scarcity of labor, so encouraging higher participation rates, in-migration, greater investment in human capital to increase productivity rates, capital substitution for labor and/or a shift in trade, including outsourcing, etc. (see for instance, Cappelli 2003). Each of these factors will also be influenced by national, regional or local circumstances and institutions (including pension and wage systems, migration policies etc.).

To determine the effective size of the labor force in a region or country, at a specific time, it is necessary to consider the numbers of people in each age group (cohort), their participation rates, and their average hours worked.[2] In addition the productivity of each age cohort per hour worked needs to be taken into account. The size of each cohort is largely determined by historic demographic factors and migration. Participation rates are influenced by factors such as: gender, previous work experience, age etc. Generally, in recent decades, participation rates have tended to fall dramatically after the age of 55 years and reduce to very low rates

---

[2] Annual hours, taking account of holidays and other leave.

after State Pension Ages (usually around 65 years old, although in some places participation rates among older workers has started to increase, see below). As is discussed later, the links between productivity and age structures are not clear.

The next section considers the broad demographic trends in the size of the working population in selected countries. Although focusing upon the major aging OECD countries, figures are also considered for Brazil, Russia, India and China (the 'so-called' BRIC countries who may become major global economies in coming decades). Section 5.3 considers participation rates and section 5.4 considers productivity amongst older workers. Finally, some policy issues are briefly considered with the conclusions.

5.2 Population Change

The sizes of each population age cohort, have dramatically changed in the last half-century in most developed countries. The 'baby boom' following 1945 led to a large increase in the population and then the labor force, although this is now beginning to come to an end in many places as the 'baby boomers' start to retire. Also, over recent decades the birth rate has decline dramatically due to birth control technologies, socio-cultural changes, and economic factors etc. In many countries the Total Fertility Rate is now below replacement level, such as Japan, Italy, Spain, UK and Germany, and is declining in most other countries.[3] The result has been a large relative increase in the size of the working age population over the last half-century, but a fall is projected to start within the first few decades of this century, depending upon the specific country.

Table 1 presents a broad overview of population trends for a range of major OECD and BRIC countries. United Nations (2005) projections to 2050 show rapidly aging population structures will continue in each of these countries and a large fall in projected labor forces (assuming no changes to retirement ages and participation rates).[4]

The share of 15-49 year olds within the total population has generally been decreasing slightly from 1960 and especially from around 1980 to 2010, in each OECD country and Russia. The trends have a slight inverted 'U' shape of the population cohort increasing to around 1980/1990 and then declines, which are often projected to get relatively larger in future decades. In Japan, the peak in this age cohort was notably earlier than elsewhere (around 1970) while peaks were around 1990 for the USA, Canada, UK, Italy and France. Overall, in contrast in China, Brazil and India the rates have generally risen fairly steadily, but with a peak in 2000 for Brazil. Although the rate for China will follow a similar decline

---

[3] The total fertility rate (TFR) is the number of children that a woman would have over her lifetime if she experienced the age specific fertility rate for that year for each age between 15 and 44.

[4] Although developed countries are generally facing an aging population structure, the global population is still increasing greatly.

to the OECD countries (due partly to the 'one child' policy and to smaller families linked to economic growth and other changes).

Table 5.1. Percentage of Population by Age as a proportion of the Total Population (and Population in Thousands)

		1960	1970	1980	1990	2000	2010
USA	0- 14	31%	28%	23%	22%	22%	19%
	15-49	46%	47%	52%	53%	51%	48%
	50-64	14%	15%	15%	13%	15%	19%
	65+	9%	10%	11%	12%	12%	13%
	Total	186158	210111	230406	254776	283230	308557
Canada	0- 14	34%	30%	23%	21%	19%	16%
	15- 49	47%	49%	54%	55%	52%	49%
	50- 64	12%	13%	14%	13%	16%	21%
	65+	7%	8%	9%	11%	13%	14%
	Total	17909	21717	24516	27701	30757	33216
Japan	0-14	30%	24%	24%	18%	15%	14%
	15- 49	53%	56%	53%	51%	47%	43%
	50- 64	12%	12%	15%	18%	21%	21%
	65+	6%	7%	9%	12%	17%	22%
	Total	94096	104331	116807	123537	127096	128220
UK	0-14	23%	24%	21%	19%	19%	16%
	15- 49	47%	45%	47%	49%	48%	48%
	50- 64	18%	18%	17%	16%	17%	19%
	65+	12%	13%	15%	16%	16%	17%
	Total	52372	55632	56330	57561	59415	60262
Germany	0-14	21%	23%	18%	16%	16%	13%
	15- 49	47%	47%	51%	50%	49%	47%
	50- 64	20%	16%	15%	19%	19%	20%
	65+	12%	14%	16%	15%	16%	20%
	Total	72815	78169	78289	79433	82017	81353
France	0-14	26%	25%	22%	20%	19%	18%
	15- 49	44%	48%	49%	50%	49%	46%
	50- 64	18%	14%	15%	16%	16%	20%
	65+	12%	13%	14%	14%	16%	17%
	Total	45684	50772	53880	56735	59238	61203
Italy	0-14	25%	25%	22%	16%	14%	13%
	15- 49	50%	49%	48%	51%	49%	46%
	50- 64	16%	16%	16%	18%	19%	20%
	65+	9%	11%	13%	15%	18%	21%
	Total	50200	53822	56434	56719	57530	56390

Table 5.1 Percentage of Population by Age as a proportion of the Total Population (and Population in Thousands) (continued)

China	0-14	39%	40%	36%	28%	25%	20%
	15- 49	46%	47%	50%	56%	56%	56%
	50- 64	10%	9%	10%	11%	12%	16%
	65+	5%	4%	5%	6%	7%	8%
	Total	657491	830675	998878	1155305	1275133	1366213
Brazil	0-14	43%	42%	38%	35%	29%	25%
	15- 49	46%	46%	50%	52%	56%	55%
	50- 64	8%	8%	8%	9%	10%	13%
	65+	3%	4%	4%	4%	5%	6%
	Total	72742	95988	121616	147957	170406	191444
India	0-14	40%	40%	39%	36%	33%	29%
	15- 49	48%	47%	48%	50%	52%	54%
	50- 64	9%	9%	9%	9%	10%	11%
	65+	3%	4%	4%	4%	5%	6%
	Total	442344	554911	688856	844886	1008937	1164020
Russia	0-14	30%	27%	22%	23%	18%	14%
	15- 49	51%	52%	53%	49%	54%	52%
	50- 64	13%	14%	15%	18%	16%	21%
	65+	6%	8%	10%	10%	12%	13%
	Total	119906	130392	138660	148292	145491	136976

Source: ILO, Laborstat, 2005

The share of the total population made up by 50-64 year olds is rising across major parts of the globe. In Japan this share has risen steadily from 12% in 1960 to 21% in 2000 (and projected to remain at this level in 2010), while the share of 65+ year olds will have risen from 6% in 1960 to 22% in 2010. Many European countries are following a similar profile although the number of 65+ year olds has not yet reached Japanese levels. Of the four largest European Union countries, Italy saw a rise of 50-64 year olds from 16% of the population in 1960 to 20% for 2010, with the share of 65+ year olds rising from 9% to 21%. In Germany the share of 59-64 year olds fell from 20% in 1960 to 15% in 1980, but back to 20% by 2010, with the 65+ population rising from 12% to 16% to 20% in the same periods. France and the UK saw similar 'U' shapes in 50-64 year olds' share of population, rising to 20% and 19% respectively by 2010, with continued projected rises after then. So while the trends within Europe are similar, there are some differences in the speed and level of how far each country has moved towards more aged population structures. Canada shows a steady rise from 12% to 21% in 50-64s from 1960 to 2010, with the 65+ share also doubling from 7% to 14%. The USA experienced high levels of in-migration over this period, although the share of 50-64s is rising (from 14% to 19% in 2010) and so is the share of 65+, but it is currently at more modest levels than many of the other countries discussed (rising from 9% to 13% from 1960 to 2010). Russia has a similarly aging population, roughly similar to Canada's structure in 2010.

However, in Brazil, India and China the shares of older population cohorts are at much lower levels than the OECD countries, but rising. In China the 50-64 year olds share rise from 10% to 16% between 1960 and 2010, but rise to only 13% and 11% in Brazil and India. Similarly the 65+ year olds shares rise during this period from 5% to only 8% in China, and 3% to 6% in Brazil and India.

At the youngest end of the population age, the share of the total population made up by under 15 year olds is falling across most OECD countries and Russia. For instance, in the USA from 1960 to 2010, the share of under 15 year olds is expected to fall from 31% to 19% (and was 22% in 2000). In Japan and Russia the figures are 30% and 14%, while in Canada 34% and 16%. The UK, Germany, France and Italy show similar patterns to each other with the share of under 15s falling from 23%, 21%, 26% and 25% respectively in 1960 to 16%, 13%, 18% and 13% by 2010. In China the decline is almost a halving of the share from 39% to 20%, while there is relatively less decline in Brazil (43% to 25%) and India (40% to 29%).

5.2.1 Absolute Changes in Population

In some of the countries, not only are the shares of under 15s falling, but so are their absolute numbers. In Europe the number of under 15s falls in absolute terms by 20%, 31%, 7% and 42% in the UK, Germany, France and Italy respectively from 1960 to 2010, compared to a total population increases of 15%, 12%, 34% and 12%, and a rise of 18%, 12%, 40% and 3% respectively of 15-49 year olds.[5] Similarly in Japan, the total population rises by 36% from 1960 to 2010, 15-49 year olds stay about constant, while under 15s fall by 36% and over 65s rise by 400%. In Canada the total population rises by 11%, 15-49 year olds by 93% but under 15s fall by 13%. So in the longer-term the aging of society and the work force structure will continue and lead to a decline in the labor force, *ceteris paribus*. In Russia over 1960-2010 the total population grows by 14%, 15-49 year olds rise at about the same rate (17%), but 50-64 year olds rise by 89% and 65+ by 135%. Meanwhile the numbers of under 15s falls dramatically by half (48%).

There are small increases in the absolute numbers of under 15 year olds (a key part of the future labor supply) in the USA and China, but large increases in Brazil and India. In the USA there is a small increase in under 15s (by 4%) from 1960 to 2010, but this compares to an overall population rise of 66%, a rise in 15-49 year olds (the key labor market cohort) of 75%, and a rise of 133% in over 65 year olds (note that the size of the cohort age ranges is different for each age group, but the overall trend is clear). In China the total population rises by 108% from 1960 to 2010, 15-49 year olds rise by 149%, while under 15s rise only 8% and over 65s

---

[5] If equal age cohorts are taken, e.g. 5-year bands, then the aging of the population structure appears more pronounced as there is a relative growth of older people in each age group, such as the 45-49 year olds compared to the 20-24 year olds.

rise by 249%. Finally in Brazil the number of under 15 rises by 54%, which is large compared to the other countries, but small compared to the total population growth of 163%, and 218% for 15-49 year olds, 343% for 50-65 year olds and 414% for the over 65 year olds. However, the later period of 1990 to 2010 the number of under 15s declines by 6% (compared to a total population growth of 29%) In India the respective figures for 1960 to 2010 are 93% for under 15s, 163% for total population growth, and 193%, 246% and 342% for the other age groups. For 1990 to 2010 the number of under 15s grows by 11%, but this compares to a total population growth of 38%.

In summary, the trends suggest that there will be major reductions in the relative numbers, and increasingly absolute numbers, of young people going into higher education and entry-level jobs in the main OECD countries and Russia, with small increases only in the USA and China. Even if the absolute labor forces remained constant, the change in the age structure has important implications for different types of labor supply. This implies a tightening of the labor market for entry-level positions including those who would traditionally have proceeded onto craft or technical training. In Brazil and India the numbers of younger people are still growing fairly rapidly (from 1960 to 2010), but this has changed to an absolute decline in Brazil after 1990 and a small increase in India. So the aging of labor forces is becoming a fully global phenomenon.

5.3 Labor Force Participation Rates

In addition to the absolute size of each age cohort, their participation rate in the labor market (the proportion of the working-age population that is working or looking for work) is important and this has also been changing in recent decades. Table 2 shows the activity rate (all persons classified either as employed or as unemployed) as a percentage of the cohort population.[6] As expected, in all countries the percentage of under 15s participating in the labor market is now virtually zero. Among those over 65 years the rates have declined markedly since 1960 and are extremely low in most countries, except Japan (expected to be 19% in 2010, but down from 36% in 1960), with the USA and Russia (9% and 7% respectively) higher than the European countries. China with 14% (28% in 1960) and Brazil (14% down form 32% in 1960) were relatively low, although India was very high

---

[6] It is also useful to use the measure of employment per population cohort (Funk 2004) as it takes account of some institutional variables particularly for the over 50 year olds (such as people moving from employment to unemployment (i.e. still counted as ‘active’) as a stepping stone towards early retirement in Germany, but similar types of people moving from employment directly to incapacity or sickness benefits (i.e. inactivity), in the UK. Also the proportion employed generally better reflects the contribution towards GDP and taxation of the population. However, it will be influenced by the business cycle (by excluding higher numbers of unemployed during contractions) and will not fully reflect the capacity of the labor market·

at 30% (compared to 44% in 1960). So the rate appears to be declining considerably across the globe, although it remains higher in most poor countries, and Japan. Nation and personal income and social, pension and welfare factors all seem important in activity rates for older people. The activity rate for 15-49 year olds has been increasing over the decades in each country, partly as a result of greater levels of female working in the formal labor market.[7] In the USA it rose from 64% in 1960 to 79% in 2000 and expected to remain at that level in 2010, with similar rises in Canada (61% in 1960 to 82% by 2010), the UK (72% to 80%), Germany (74% to 80%), France (68% to 76%), Italy (63% to 75%), Japan (71% to 78%) and Russia (78% to 84%). China (87% to 88%, although these figures are influenced by the link between state owned enterprises and wider social services for employees) Interestingly India has declining activity rates for the 15-49 year olds (72% in 1960 and 68% in 2010).

Table 5.2. Percentage Economically Active Population by Age Group

	Age Group	1960	1970	1980	1990	2000	2010
USA	10-14	2%	2%	0%	0%	0%	0%
	15- 49	64%	66%	74%	77%	79%	79%
	50- 64	65%	65%	62%	64%	66%	65%
	65+	21%	16%	13%	11%	10%	9%
Canada	10-14	2%	2%	0%	0%	0%	0%
	15- 49	61%	65%	75%	81%	82%	82%
	50- 64	59%	60%	60%	59%	63%	61%
	65+	18%	13%	9%	7%	6%	5%
Japan	10-14	3%	2%	0%	0%	0%	0%
	15- 49	71%	70%	69%	70%	74%	78%
	50- 64	70%	71%	70%	71%	73%	70%
	65+	36%	35%	29%	25%	22%	19%

---

[7] For instance, in Germany labor force activity rates of women in the age 55-59 age group increased from 49.8% in 1995 to 59.8% in 2003, while there was a similarly high growth rate in Canada - from 48.2% to 60.0%. For the 60-64 age group, rates rose from 10.9% to 17.5% and from 23.3% to 32.2% respectively (ILO Laborstat, 2005). In the UK the labor force participation rate for women with dependent children has risen from 61% in 1988 to 67% (compared to 71% for males) in 1998, and there has been an increase in the number of women remaining childless. Elsewhere female participation rates have risen considerably, for example, especially between 1985 and 1997 in East and South-East Asian countries. However, female participation appears to have peaked in the early 1990s after which the growth reduced markedly or may also have actually begun to fade even before the effects of the 1997 financial crisis began to be felt (UNESCAP 2005, p. 95).

Table 5.2. Percentage Economically Active Population by Age Group (Continued)

UK	10-14	0%	0%	0%	0%	0%	0%
	15- 49	72%	74%	75%	79%	80%	80%
	50- 64	61%	64%	68%	62%	63%	59%
	65+	14%	10%	7%	5%	4%	4%
Germany	10-14	3%	1%	0%	0%	0%	0%
	15- 49	74%	75%	75%	79%	80%	80%
	50- 64	59%	56%	57%	55%	51%	57%
	65+	14%	11%	6%	3%	2%	2%
France	10-14	4%	3%	0%	0%	0%	0%
	15- 49	68%	69%	71%	70%	74%	76%
	50- 64	60%	56%	62%	51%	54%	50%
	65+	16%	10%	4%	2%	2%	1%
Italy	10-14	11%	4%	2%	0%	0%	0%
	15- 49	63%	64%	65%	68%	73%	75%
	50- 64	48%	44%	47%	43%	42%	41%
	65+	15%	8%	7%	5%	4%	3%
China	10-14	34%	35%	34%	15%	8%	0%
	15- 49	87%	87%	89%	89%	89%	88%
	50- 64	62%	61%	60%	65%	65%	64%
	65+	28%	24%	18%	19%	17%	14%
Brazil	10-14	22%	20%	19%	18%	14%	11%
	15- 49	58%	60%	64%	70%	70%	69%
	50- 64	51%	51%	52%	54%	52%	51%
	65+	32%	27%	21%	18%	16%	14%
India	10-14	30%	25%	21%	17%	12%	7%
	15- 49	72%	71%	70%	67%	67%	68%
	50- 64	69%	67%	66%	63%	62%	62%
	65+	44%	41%	38%	35%	32%	30%
Russia	10-14	0%	0%	0%	0%	0%	0%
	15- 49	78%	83%	84%	84%	82%	84%
	50- 64	49%	50%	61%	58%	55%	58%
	65+	7%	7%	7%	8%	7%	7%

Source: ILO, Laborstat, 2005

For the 50-64 year olds the picture is more varied, partly reflecting varying normal retirement ages etc. This is partly due to differing institutions such as taking early retirement (where it was common in European countries to respond to the restructuring in the economy by allowing people to take early retirement, although some of these people will still look for work). In the USA, Japan and Brazil the rates remained fairly constant at around 65%, 70% and 51% respectively in 2010 (the same as in 1960). In Canada, China and Japan the 50-64 participation rates rise slightly from 59% to 61% and 62% to 64% respectively from 1960 to 2010. In Russia the rise is more dramatic, from 49% in 1960 to 61% in 1980 and then fluctuating to 58% by 2010. In Europe, however, the rates fell in the UK (where rates have had a strong inverted ‘U’ shape over time), Germany and particularly in France and Italy from 61% to 59%, 59% to 57%, 60% to 50% and 48% to 41% respectively. This may be linked to high levels of welfare and early retirement provision, and in some cases the withdrawal of former workers in declining sectors from the labor market (McQuaid and Lindsay 2004). This has important implications for dealing with the future impacts of aging population structures in Europe, as a major policy is to increase overall participation rates (CEC 2003, 2004a). While Brazil has had a fairly constant rate since 1960, it is low, around the level of some European countries. India meanwhile has a higher, but declining level (from 69% in 1960 to 62% in 2010).

A further group of the labor market is the post State Pension Age (or ‘normal’ retirement age) cohort. As a much higher percentage of people enter the labor force later after going to higher education, and longer life expectancy means the time between retirement and death increases, there is a reduced period of working-life, both absolutely and relative to non-working life. It is likely that policies to delay normal retirement age (by reducing early retirement or raising the retirement age) or encouraging part- or full-time work after retirement age are likely to be developed and intensified. Working past State Pension Age has become more prevalent in countries such as USA and Japan and is likely to be a future trend in other countries (see Table 2).

5.3.1 Regional Issues in Participation

As shown, there are major differences in participation rates over time and between and within regions in a country. Taking the UK as an example, Campbell (1999) found that around two-fifths of men aged between 55 and 65 were without paid work in 1997, compared to one-fifth in 1979. Until the mid-1970s, the participation rate in the UK for men aged 55 and over was one of the highest among OECD members (OECD 1995) with an activity rate for all 50-64 year olds of 68% in 1980. One half of men and one third of women retire before State Pension Age (Disney 1996). In addition, each cohort of men appears less likely to remain in employment at older ages so these trends are not fully explained as a consequence of the downturn in the economy during the 1980s and 1990s but, rather, are part of an ongoing process.

At regional and sub-regional levels differences in participation rates may be linked to former industrial structures (e.g. areas with mining or heavy industry often have high sickness rates among older former workers), previous out-migration etc. (Hollywood et al. 2003). As Table 3 illustrates, there are wide regional difference inactivity rates even within specific age bands. More prosperous UK regions, such as the Eastern and South East have high activity rates, for those aged 50 to State Pension Age (currently 60 years for females and 65 for males, although the age for females will rise to 65 years by 2020), at 76.5% and 76.4% respectively, while former heavy industrial areas, such as Wales (65.7%) or the North East (60.9%) have relatively low activity rates.

Table 5.3. Economic participation rates for age 50 state pensionable age (male 65, female 60) in Great Britain

Eastern	76.5
South East	76.4
South West	74.3
East Midlands	73.4
West Midlands	72.0
Scotland	70.5
London	69.8
Yorkshire and Humber	68.9
North West	68.8
Wales	65.7
North East	60.9
Great Britain	71.5

Source: Labour Force Survey (LFS), ONS, February 2003

Taking a UK statistical region near the average, Scotland, large intra-regional differences in participation or economic activity rates are clear. The rural areas generally have high activity rates (e.g. Aberdeenshire, 83.3%, Shetland Islands 87.0%) while the older industrial areas have lower rates. For example rates of activity for those aged 50-State Pension Age in industrial Glasgow and North Lanarkshire were 55.5% and 55.4% (LFS 2004). These different rates may be due to factors such as greater levels of multi-job, part-time work and out-migration in rural areas, while in former heavy industrial areas there may be a cohort effect of ill health due to industrial disease and a discouraged worker effect, where people leave the labor market due to lack of perceived opportunities.

Campbell found that increasingly fewer older people return to work after leaving the labor market and older people without a job are likely to become less attached to the labor market moving from unemployment to long-term sickness or retirement. So the decrease in activity of older workers has not been matched by an increase in unemployment (Beatty and Fothergill 2002), to an extent hiding the levels of withdrawal from the labor market.

5.4 Productivity and Aging

In terms of aggregate productivity it is difficult to determine the effect of change age compositions of the labor force. As wages are correlated with seniority and age and are partly used to encourage retention and good performance (OECD 1998), if these wages do not reflect actual higher productivity then as the population ages there may be a potential reduction in the productivity per wage. This may affect competitiveness and prove an incentive for firms to use redundancies to reduce the numbers of older workers. An aging workforce may potentially, but not necessarily, reduce national productivity as: traditionally older workers have often been considered less productive or entrepreneurial than younger workers; there will be a shift in expenditure from more to less productive sectors; and the relative size of the employed workforce is expected to significantly reduce under current trends (*ceteris paribus*). On the other hand, older workers may have greater life and work experience, longer tenures and lower turnover, less absenteeism and require less supervision than younger workers (Johnson 2003; Jackson 1998, pp. 95-100). Also some loss of productivity among older workers may be due to skills obsolescence, rather than age (Skirbekk 2004).

However, there is less empirical evidence on the links between productivity and age in the skills needed in modern businesses using, for example, high levels of new technologies. If labor markets grow more slowly, or cease to grow, there may be incentives to innovate and use existing resources more efficiently, while higher wages due to aging should lead to greater investment in human and physical capital, and hence productivity growth. The labor force trends discussed above suggest that, in the most affected countries, both increases in the working population and/or increases in the relative productivity of older workers are important.

The role of employers in assisting older workers to improve productivity is of course important as changes they make, such as modifying the duties of a job, changing the capital-labor ratio, imposing mandatory shift coverage, and hiring contract employees, immigrants, or labor in foreign countries or outsourcing, can all influence productivity levels (Horrigan 2004).

Any decline in productivity due to an aging labor force, *ceteris paribus*, could be in terms of GDP per worker (if older workers are less productive, i.e. the composition of the workforce changes), and/or in terms of GDP per capita (if the working population falls in relation to the total population, i.e. the dependency ratio changes). A further issue is whether we consider productivity per hour or per person, especially if there is a higher propensity for part-time working among older workers or average working hours change. Hall and Jones (1999) argue that output per worker is more important than output per capita, although the latter may be more relevant for changes in total output of the economy. Of course, increasing productivity is generally desirable anyway, irrespective of the relative productivity of older workers. Overall, these factors will influence the parameters of regional models over time. Indeed, if production does not rise there is a danger that a large increase in the dependency ratio could lead to higher taxes (and other costs such as for pensions etc.) for those still working, possibly leading to a spiral of disincentive for some and potential out-migration, or to a convergence in wages or workforce structures if real wages rise due to labor scarcity with increases in participation rates and in-migration.

This discussion suggests some key issues for modeling. Each age cohort is likely to have a different elasticity of labor demand and of supply. If the skills and qualifications of older workers are not renewed and improved, in comparison with those of younger workers, then the labor demand for older workers shifts to the left and leads to a less elastic demand schedule (Funk 2004). Funk suggests that such more limited skills development for older workers and wages based on seniority etc. are largely institutional. It is crucial to fully incorporate age, gender, regional and sub-regional factors if we are to more fully understand the implications of aging upon labor markets across the globe.

5.5 Policy Issues and Conclusions

In the next half century the effects of demographic changes are like to be great in terms of its impacts upon economic growth of different countries and their regions, (affected by labor supply and dependency ratios), demand for goods and services, income distributions, globalization of labor markets etc. There are a number of interconnect policies related to an aging workforce. Various non-mutually exclusive approaches are possible, including: increasing the size of the domestic population (for instance through the globalization of the labor market); increasing the participation rates of the population in the labor force; raising productivity; and globalizing trade and production. This paper has briefly considered the first three. Another approach could be to accept a lower standard of living and/or population. If total fertility rates remain low in developed countries, then policies will be required to manage the decline of populations and their increasingly aged structures even if there are relatively high levels of net in-migration.

Specific policies concerning the size of the domestic labor force include: increased birth rates (but these are difficult to promote effectively); increased net temporary and permanent in-migration across all skills bands (although this would need to be at a very large scale in the case of many countries, especially among younger workers (see: Wright 2002), and within the EU almost all countries will be seeking to attract immigrants from each other as well as from beyond the EU). There are dangers that this could lead to a continuous outflow of young and skilled workers from developing countries, and the aging of population structures in developed countries would be expected to increase global migration. Also the population trends in many populous countries also have a rapidly aging population structure, although they may be some years behind most OECD countries, and as they develop then more domestic opportunities for young workers may limit the level of their out-migration.

Increased participation rates for the population, and specifically for each age cohort, may be achieved through policies focused upon: bringing in more people not currently in the labor force (often through increasing the female participation rate); people working to an older age, for example through changing social security and pension incentives, increasing the retirement age, or tapering the age when people stop work so rather than choosing to retire from work totally people may gradually decrease their working hours; or an increase in working after normal retirement age State Pension Age; improving labor market flexibility through better work-life-balance policies, including improved support or employer policies for those looking after older spouses or other relatives; increasing hours worked (although recent experience in the EU has been the reverse). However, it needs to be recognized that many older workers would not wish to continue working, particularly those with relatively high real incomes and with current incentives to retire (Landis and MacKellar 2000), and the wealth effects of declining populations (such as spreading inheritances amongst fewer children) may also continue to allow some to retire early, although the distribution of such wealth may affect outcomes. Many of these possible policies would attempt to reverse trends of the last several decades, although this could be seen as a response to the greater life expectancy and other changes during these years. Increases in the productivity of older workers requires both supply and demand side measures, with incentives for both sides to invest more in the human capital of older workers and the redesigning of workplaces, systems and tasks by employers, so as to help improve worker productivity.

In some places the development of these policies may have been inhibited in the short term by current high unemployment rates, for example in much of the European Union, and the actual short- to medium-term increase in labor forces (due to the remnants of post 1945 the baby boomers), so limiting immediate pressure on labor markets. As discussed earlier, potential labor shortages are likely to lead to significant changes in the behavior of labor demand and supply at both regional and national levels, so as the effects of aging population structures increase, greater changes in policy responses are to be expected at regional and national levels.

Although most countries include, or are likely to include, elements of the full range of polices, some may have greater emphasis upon certain aspects. One approach may be to accept a long-term decline in the absolute population of a country (and this has occurred in many regions over time), but to try to maintain living standards through improving productivity and investment abroad (i.e. exporting capital to employ workers abroad and create a long-term income stream back to the original country) – which has similarities to the Japanese approach over recent decades.[8] Another approach could be to seek to maintain, or increase, the domestic workforce (particularly through the in-migration of both skilled and unskilled workers), with an associated increase in GDP, as appears to have happened in the

---

[8] For instance, in Japan 0.2% of the labor force was foreign in 2001, while 13.9% of the USA labor force was foreign born (OECD 2005). In Australia 24% of the labor force was foreign born. In terms of inflows of foreign workers, these numbered (excluding seasonal workers) 142,000 moving to Japan in 2001 and 179,000 permanent settlers and 679,000 temporary workers to the USA.

USA. In the EU at the start of the twenty first century eleven of the 15 Member States needed in-migration to maintain their populations during recent years (CEC 2004b). There is a need to form improved dynamic models of regional growth which incorporate demographic change and dependency rate changes (including the proportion of working population and working age population), as well as information and technological change, entrepreneurship and other factors.

In summary, the aging of societies across the globe, but especially in developed economies such as Japan and the European Union (EU), is one of the main challenges facing the employers and governments. One set of factors is the reduced population growth, or even decline, due to factors such as low fertility rates. Another important factor is the reduction in labor force participation rates, due both to demographic factors resulting in relatively more people reaching retirement age, but also due, in recent decades, to many ceasing work before they reach State Pension Age. This trend is beginning to change with fewer employers offering early retirement, greater flexibility in pension provision, and moves towards increasing State Pension Ages. There will be pressure to reverse some of the trends of recent decades towards lower participation rates, especially for older workers. It would be useful for regional modeling to take fuller account of the aging and changing activity rates such as working versus non-working population ratios.

These population changes will greatly affect the dependency ratios and the need for seeking to increase the labor force above what it would otherwise be, through a range of policies. In addition to seeking to ameliorate any decline in the labor force, the affected regions and countries need to consider issues such as how to increase the productivity of older workers and how the changing income distributions will affect the development of the economy and provision of services for older people (especially where pension and health costs are funded out of current taxes and income rather than savings). Of course older people may make considerable paid and unpaid contributions (such as caring for relatives, voluntary work etc.). In many countries it is likely that a variety of policies will be developed to try to reverse the population and activity rates trends. There is a need to refine our understanding and modeling of the potential effects of these changes and the policies to deal with them. This suggests that regional models should seek to fully incorporate age structures, participation rates and related differential productivity rates among regional and sub-regional factors if we are to more fully understand the implications of aging upon labor markets across the globe.

References

Alonso W, (1978) A theory of movements. In: Hansen NM, (ed) Human settlement systems, Ballinger, Cambridge MA, pp 197-211

Batey P, Madden, M (1998) The employment impact of demographic change: A regional analysis. Papers in Regional Science, 78(1):69-87

Beatty C, Fothergill S (2002) Hidden unemployment among men: a case study. Regional Studies, 36(8): 811-823

Bloom D, Canning D (2003) Contraception and the Celtic Tiger. Economic and Social Review, 34(3): 229-247

Campbell N (1999) The decline of employment among older people in Britain. Centre for the Analysis of Social Exclusion Paper 19, CASE, London

Cappelli P (2003) Will there really be a labor shortage? Organizational Dynamics, 32:221-233

CEC (Commission of the European Communities) (2004a) Increasing the employment of older workers and delaying the exit from the labour market, Communication COM (2004) 239 Final, CEC, Brussels

CEC (2004b) A New Partnership for Cohesion - Convergence Competitiveness Cooperation, Third Report on Economic and Social Cohesion, Office for Official Publications of the European Communities, Luxemburg

CEC (2003) Proposal for a Council Decision on Guidelines for the Employment Policies of Member States. Office for Official Publications of the European Communities, Luxembourg

Disney R (1996) Productivity, wages and educational attainment: The impact of workforce aging. In: Disney R. (ed) Can We Afford to Grow Old? MIT Press, Cambridge MA, pp 153-192

Endo Y, Katayama, E (1998) Population Aging and Japanese Economic Performance. In: Bosworth B, Burtless G (eds) Aging Societies: The Global Dimension, The Brookings Institute Press, Washington, DC, pp 240-265

Funk L (2004) Employment Opportunities for Older workers: A comparison of selected OECD-Countries. In: CESifo DICE Report Journal for Institutional Comparisons, Vol. 2, pp 22-33

Greenwood M (1997) Internal migration in developed countries. In Rosenzweig, M.R. and Stark, O. (eds), Handbook of Population and Family Economics, Elsevier Science, Amsterdam, 647-720

GROSS (General Register Office for Scotland) (2002) Scotland's Population 2002 - The Registrar General's Annual Review of Demographic Trends. GROS, Edinburgh

Hollywood E, Brown R, Danson M, McQuaid RW (2003) Older Workers in the Scottish Labor Market: A New Agenda. Report to Scotecon, Glasgow www.scotecon.net/publications/McQuaid%20older%20workers%20Full%20Report.pdf

Horrigan MW (2004) Employment projections to 2012 - concepts and context. Monthly Labor Review, 127(2):3-22 www.bls.gov/opub/mlr/2004/02/art1full.pdf

Katz L, Kreuger A, (1999) The high-pressure U.S. labour market of the 1990s. Brookings Papers on Economic Activity, 1:1-63

Landis F, MacKellar M (2000) The predicament of population aging: a review essay. Population and Development Review, 26(2): 365-97

Lisiankova K, Wright RE (2005) Demographic change and the European labor market. National Institute Economic Review, 194:74-81

LFS (Labour Force Survey) (2003), (2005) Labour Force Survey, Office for National Statistics, London

Lindsay C, McQuaid RW (2004) Avoiding the 'McJobs': Unemployed job seekers and attitudes to service work..Work, Employment and Society, 18(2):296-319

MacKeller LF (2003) The Predicament of population aging: A review essay, Vienna Yearbook of Population Research, Vienna Institute of Demography, 73-99

OECD (Organisation for Economic Cooperation and Development) (1995) The Labor Market and Older Workers, OECD Social Policy Studies No. 17, OECD, Paris

OECD (1998) Work-force aging. OECD Employment Outlook 1998, Paris, OECD, 123-51

OECD (2005) Trends in International Migration, OECD, Paris

Robson WBP (2001) Aging Populations and the Workplace: Challenges for Employers, The British-North American Committee, Toronto

Samorodov A (1999) Aging and labor markets for older workers. Employment and Training Papers. ILO, Geneva

Skirbekk V (2004) Age and individual productivity: A literature survey. Vienna Yearbook of Population Research, 133-153

United Nations (2005) World Population Prospects, The 2004 Revision, United Nations, New York

UNESCAP (United Nations Economic and Social Commission for Asia and the Pacific) (2005) Sustainable social development in a period of rapid globalization: challenges, opportunities and policy options, UNESCAP, New York

Wright RE (2002) The Impact of Population Aging on the Scottish Labour Market, Quarterly Economic Commentary, 27(2):38-43

6 The Role of Intraindustry Trade in Interregional Trade in the Midwest of the US

Darla K. Munroe[1], Geoffrey J.D. Hewings[2] and Dong Guo[3]

[1] Department of Geography, The Ohio State University, Columbus, Ohio USA
[2] Regional Economics Applications Laboratory, University of Illinois, Urbana, Illinois USA
[3] Laboratoire d'Economie et de Gestion, Université de Bourgogne, Pôle d'Economie et de Gestion, Dijon, France

6.1 Introduction

The subject of international trade among countries has long been of concern to policy makers and academics alike. As economic activity has become more and more international in scope, the potential impact of international trade on regional economic growth and income distribution has become central to many studies. Within economics, the study of industrial organization, particularly with respect to imperfect competition and economies of scale and agglomeration, has influenced developments in international trade theory in the past few decades. In identifying the determinants of trade among countries, issues such as market size, relative level of Gross National Product (GNP) per capita, market structure, etc., have all become important, as well as the more traditional determinants of trade, e.g., relative capital and labor endowments. Furthermore, there has been an increasing realization of the role and influence of location in explaining trade and trade patterns (see Krugman, 1990; Fujita *et al.*, 1999; Hanson, 1996; Martin, 1999, Venables *et al.*, 2003)

What about trade among regions, within countries? If international trade has significant impacts on economic growth and welfare concerns (employment, income, etc.), it should follow that trade within countries may also merit much further consideration. In the U.S. Midwest, the volume of trade among states exceeds the volume of foreign trade originating from those states by several orders of magnitude. For example, table 6.1 summarizes the volume of export flows between Illinois and some of the U.S. largest trading partners, and between Illinois and other states in the Midwest.

One can infer from this table that domestic trade flows among these states are a significant economic force in that region. At the same time, international trade to and from the Midwest is also substantial, but the trade among states within the

Midwest certainly merits further study. In some cases, this interregional trade may even be the driving force behind increased international trade from the Midwest (Hewings, *et al.*, 1997).

Table 6.1. Largest Volumes of International and Interregional Exports from Illinois, 1995

Country	Volume (US$ Billion)	State	Volume (US$ Billion)
Canada	6.29	Indiana	17.7
Japan	2.19	Michigan	17.5
Mexico	2.08	Ohio	19.9
UK	1.34	Wisconsin	18.0
Germany	1.28		

Source: Estimated and updated from Hewings, *et al.*, (1997) and Commodity Flow Survey (1966).

Not only has there been little discussion of the role of interregional trade, for over two decades, little or no information was available to document the magnitude of state-to-state flows. As a result, it is difficult to do more than infer from other information the growth of this trade. In this paper, attention will be focused on the *nature* of this trade and its association with economic structure. In particular, the analysis will explore the degree to which trade in the region is dominated by interindustry rather than intraindustry trade. Concomitantly, the emerging trade interdependence will be explored in connection with the degree to which the economies of the Midwest share a similar economic structure. The next section will review recent developments in international trade theory with a particular focus on intraindustry trade. Then, some descriptive measures of trade among these five Midwestern states and some indices of trade overlap within industries will be reported. Given these trading patterns, the next section will examine the similarities and differences in the economic structures of these state economies. Finally, policy implications and directions for future study will be discussed.

6.2 Conceptual Framework

In this section, two parallel literatures will be explored in preparation for the analysis that follows. The first concentrates on the role of trade in regional development, promoting an evolutionary view in which transportation and communication costs assume the most significant roles. The second draws on the extensive literature on the nature of trade, particularly, the role of intraindustry trade.

6.2.1 Trade and Regional Development

One of the most imaginative contributions to the regional development literature was provided by Thompson (1965). His ideas about the evolutionary path of a regional economy provided an important vehicle in which to marry the ideas of export base analysis, linkage development and the ideas associated with growth in the level of intermediation in an economy. Perhaps, his work might be considered as stemming from Marshallian principles but with a strong trade orientation (see Martin, 1999). Furthermore, they provide insights into what is, essentially, a network evolution of trade with attention being paid to the internal and external division of this trade. Thompson's ideas will be used as the basis for the development of this conceptual theory. However, the evolution will be considered for a two-region context and the focus will be on an understanding of the nature and extent of interregional trade. In some senses, these ideas find an echo in Hanson's (1996) recent work, but with a different regional geography implied (interregional within a nation rather than between countries).

Early Development Stages. Assume two isolated regions, separated by a wilderness with very poor interregional transportation access. Accordingly, there will be very little specialization as the possibilities for exchange are limited by high transportation costs. Further assume that the regions are themselves located in a nation that is poorly connected to the rest of the world. Now assume that a highway or railroad is constructed between the two regions, significantly lowering transportation costs; the mechanics of the process described by Thompson can begin to unfold. Specialization will be possible now that exchange can be effected. Each region will begin to specialize in a set of goods and services in which they enjoy some comparative advantage vis a vis the other region. Trade will be dominated by interindustry exchange; within each region, new suppliers will locate to provide inputs into the firms making goods and services for export to the other region thereby creating an increase in the intraregional multipliers. In all probability, intraregional exchange may increase more rapidly than interregional trade as localization economies assume considerable importance.

Early Maturity. As these economies mature, agglomeration economies will serve to strengthen each region's competitive position in the production of goods and services that they export. Increases in export activity will generate innovations in the transportation sector, offering the possibility of lower costs of transportation between the regions. Internally, the level of intermediation will increase, increasing the intraregional multipliers but, at the same time, external trade will also increase.

Late Maturity. The next stages offer a more complex pattern of evolution spurred by two important developments, a significant reduction in the costs of transportation and communication in general and the integration of the regions into a global economy. The reduction in transportation costs is effected by significant investments in transportation infrastructure, reducing the role of these costs in the production function of the average firm and increasing the tradability of goods produced. As a result, market orientation becomes a more dominant force in location decision-making. Further, the spatial scale of agglomeration economies shift from the urban or metropolitan scale to the regional or even the multiregional scale as the regions become integrated in the global economy. Most importantly, returns to scale are now complemented by returns to scope with the following expectations:

- Internal returns to scope accruing to an individual establishment will be lower, as a smaller number of secondary products will be produced and there will be a lowering of the dependence on local (intraregional) suppliers and markets for exchange, resulting in a decrease in the intraregional multiplier *without* a concomitant decrease in levels of production;
- External to the establishment (but internal to the firm) returns to scope will be higher with secondary products produced by establishments allocated across many multiregional operations, thereby increasing interregional trade, interregional dependence and the interregional multiplier.

Changes in the spatial structure of returns to scope can be explained by the increasing role played by returns to trade, namely the increasing impact of each additional dollar invested in transportation on lowering production costs. Equally importantly, there will be a change in the composition of interregional trade with interindustry interregional trade replaced by intraindustry interregional trade. Within the regions, even though the volume of output may continue to increase, the hollowing out process will result in a decrease in the intraregional multiplier. Spatial clustering of activities will focus on different attributes of the regional economy (e.g., the role of a region's occupational capital) as firms search more widely for the highest quality and cheapest inputs knowing that transportation investments have significantly broadened the effective geography within which they can search. By exploiting returns to scope over a larger range of establishments, firms and regions can enjoy a more favorable international competitive position.

Link with the new trade theory. The apparent tensions between the ideas of the comparative advantage and the new trade theory have been shown by Krugman (1990) and others to be more apparent than real. Initial factor endowments will still allow for specialization to occur with cost advantages leading to a cumulative process of product differentiation. This process, in turn, will lead to the exploitation of increasing returns and thus to increasing trade between regions. However, this trade will be more heavily concentrated in intraindustry trade for many of the reasons provided by Krugman. It so happens that such processes find their fullest expression in the state economies of the Midwest of the US. For example, if the economies in question are large, of similar size and with few differences in factor endowments than something other than comparative advantage has to be proposed to explain trade. Per capita income is high - thus providing opportunities for product differentiation to serve the demand for higher quality products and a greater variety of products. Scale economies exist and capital/labor ratios are high and similar (i.e., no clear comparative advantage). Instead, it is lower transportation costs, region-wide agglomeration effects, and the effectiveness and ease of information flows that all combine to facilitate intraindustry specialization. At the level of trade between regions that is focused on intermediate goods and services, it is unlikely that much of the apparent cross-hauling can be explained by product differentiation. The development of niche products and markets probably accounts for the vast majority of this intraindustry trade.

Further, this process indicates a potential difference in the role of trade between nations and between regions within a nation. For example, in Hanson's two-region model, the production network involves consideration of one region dominated by higher skill requirements and another in which the skill requirements are lower. In the interregional trade within a country case, there may be few differences in skill endowments. Thus an explanation of trade must focus on other factors and on the nature of this trade itself. In this regard, Hummels et al. (1998, 1999) call attention to the role of vertical specialization in trade to offer a plausible alternative explanation. One of the characteristics of vertical specialization (the use of imported goods in production that is ultimately exported) will be the importance of intraindustry trade. Now attention will be directed to the received theory on the role of intraindustry trade

6.2.2 Intraindustry Trade

In a traditional, Heckscher-Ohlin model of international trade, trade is driven by differing factor endowments between regions. Countries specialize in the production of goods that use the most abundant factor most intensively, allowing them to capture comparative advantage through trade. The Heckscher-Ohlin model cannot adequately explain the large degree of trade taking place among similar economies, and the increasing domination of intraindustry trade in particular.[1] In this section, theoretical developments in the study of intraindustry trade will be discussed, as well as their application for regional trade models. First, broad-based determinants of intraindustry trade will be outlined. From these general categories, the type of production specialization, market structure, regional economic issues, and welfare concerns will be examined in greater detail.

Determinants of IIT. If intraindustry trade[2] is at odds with the more traditional Heckscher-Ohlin framework of comparative advantage, one must first grapple with the determinants of such trade. Stone (1997) separates the determinants of IIT into two categories: industry-based determinants, and regional characteristics. The industry-based determinants include: product differentiation, scale economies, industry specific cost structures, and transportation costs. On the other hand, regional determinants are based on macroeconomic characteristics: income level, and the relative capital/labor ratios, for example. By separating out these two components of IIT, one can learn more about both in characterizing trade flows (Balassa and Bauwens, 1987).

Stone summarizes and elaborates on the hypotheses surrounding the emergence of IIT. Not all of these determinants are incompatible with a trade regime

---

[1] A dissenting voice has been raised by Davis (1995) who proposed an innovative approach that combined Heckscher-Ohlin and Ricardo; he suggested that technical differences in production matter, thus diminishing the importance of increasing returns to scale and the need to assume imperfect competition.

[2] Henceforth intraindustry trade will be referred to as IIT.

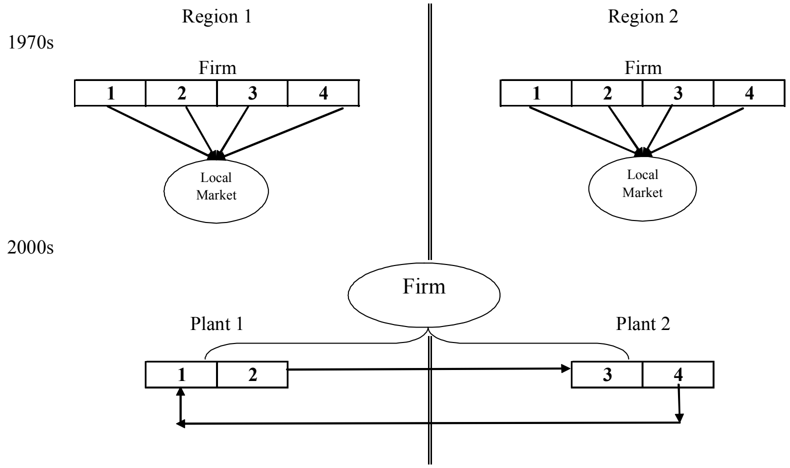
that is characterized by a Heckscher-Ohlin (H-O) framework. However, the large volume of *bilateral intraindustry trade flows* that emerge with high levels of IIT do not conform to H-O assumptions.

Hypotheses of IIT:

1. IIT will increase as income differences decrease because demand structures become more similar.
2. The share of IIT in total trade will increase as the difference in factor composition (e.g., capital/labor ratios) falls.
3. The share of bilateral IIT will increase as the average income level increases.
4. The share of bilateral IIT will increase as the difference in relative incomes, as a measure of the economy's size, of trading partners fall.
5. The share of bilateral IIT will increase as total size of trading partners increases (Stone, 1997).

Product Differentiation. Within the IIT theoretical literature, there are differing assumptions regarding the type of product differentiation within an industry that leads to IIT. The three general types of product differentiation include horizontal differentiation (differences of *variety*), vertical differentiation (differences of *quality*), and the vertical integration of production process itself (trade in intermediate goods). Krugman (1991b) has championed the case for horizontal differentiation leading to increased IIT. In his model, as economies become more similar and per capita income rises, consumer preferences become more diverse. Thus, consumer goods become differentiated by type or variety. As each region specializes in a certain variety of a good, incentives for trade arise. In figure 6.1, a stylized view of this process is suggested for an economy with two regions and production of four commodities. With high interregional transportation costs (in the 1970s), firms tended to produce goods for sale in geographically restricted markets. With decreases in interregional transportation costs by 2000s, and with greater concentration of activity (through mergers and acquisitions), firms now organize production to exploit scale economies at the plant level with production in one plant now serving both regions. Central to this argument is the assumption that the demand structure of the trading regions is very similar in nature, as are relative capital and labor endowments. This model is most applicable to the study of trade among highly developed economies, with a predominance of trade in capital-intensive goods and a high level of technology.

A second, more problematic explanation of IIT trade is that of vertical product differentiation. In this case, IIT can take place among less similar economies than required for the case of horizontal specialization. Flam and Helpman (1987) employed such a framework to study IIT between economies with differing levels of per capita income. In such a model, products within an industry are differentiated by quality. This difference in quality may be due to differences in technical efficiency or intensity of production, as well as labor productivity or differences in human capital. Lambertini (1997) theorizes that under certain conditions, such trade can benefit both trading regions, although some more welfare concerns can arise. In any case, IIT among regions with differing income distributions can potentially benefit one region more than another.



Note: Numbers refer to products

Fig. 6.1. Changing organization and structure of establishments and firms, 1970s to 2000s.

A third type of IIT that arises is due to trade in intermediate goods, or the vertical integration of production. Hummels et al. (1998) postulated that the internationalization of production led to vertically linked economies. In this model, regions specialize in a particular stage of the production process, thus leading to increased IIT as production increases. In their definition of vertical specialization, a good must be produced in multiple sequential stages, and must cross at least one international border more than once (Hummels et al., 1998). For example, in the simplest form, one country can export an intermediate good to another country that completes production of the good, and then exports the final product back to the first country. Vertical integration occurs more readily in economies with a relatively higher percentage of GNP derived from trade.

A complementary perspective, referred to as the fragmentation of production, has been provided by the work of Jones and Kierzkowski (1990), in which they state that production block can be connected by service links (as in the case of production systems in 2000s shown in figure 6.1). Still later, Arndt and Kierzkowski (2000) noted that fragmentation of production implies that constituent activities may be organized into blocks that need not to be performed in spatial proximity to one another. The process itself is elaborated further in Jones and Kierzkowski (2001) in which the concept is advanced to provide an explanation the nature of fragmentation as a consequence of the new patterns of process in the world economy. This concept is very close in spirit to Hummels et al. (1998) idea of vertical integration of production.

In fact, fragmentation is not a new phenomenon. Görg (2000) provides a more detailed definition of fragmentation in terms of multi-plants (or multi-national) firms; he mentioned that fragmentation can be take place across borders or domestically, between firms or even intra-firm (as shown in figure 6.1). In the case of intra-firm fragmentation, domestic fragmentation leads to the emergence of domestic multi-plant firms, while international fragmentation leads to intra-firm trade between plants within the same multinational company (MNC). In the case of inter-firm fragmentation, international fragmentation often leads to inter-firm trade.

Intraindustry trade, that is a main feature of international fragmentation, is defined by Tharakan (1984) as "the simultaneous import and export of products which are close substitutes from each other in terms of factor inputs and consumption." However, Thom and McDowell (1999) argued that intraindustry trade takes two forms — horizontal and vertical. Horizontal intraindustry trade is associated with economies of scale and occurs when products are differentiated and consumers express preferences for product variety. Vertical intraindustry trade, on the other hand, is similar to interindustry trade in that it exploits comparative advantage and specialization. Similarly, Aturupane (1999) states that vertical intraindustry trade is more likely to be driven by differences in endowments, as in practice much of this will reflect subcontracting activities and the slicing up of the value chain — which is the case of the production block defined in international fragmentation. However, when trade between regions with similar factor endowments is considered, it will take on more of the characteristics of horizontal intra-industry trade that is more likely to be driven by scale economies and imperfect competition. Abd-el-Rahman (1991) and Greenaway et al. (1995) argue that making such a distinction is important as the determinants of each type of intraindustry trade differ.

Price (2001) noted two trends in the fragmentation process: trends in the spatial dimension associated with economies becoming more global (in part reflected by the hollowing out phenomenon noted earlier), and in the specialization dimension where firms (and particularly plants) are becoming more specialized because of the enlarged market created by global demands.

Most the research has focused on Europe or on trade between Europe and US. In fact, as the factor endowments theory has indicated, most developing countries are a larger part of production fragmentation in the whole chain of production in the developed countries, but data problems have limited extensive analysis.

Since each of the three intraindustry trade scenarios presented at the beginning of this section requires differing production conditions leading to IIT, it is likely that each scenario may yield differing welfare concerns for the trading regions. One of the greatest problems of studies of IIT is that the causes of IIT are usually quite complex. It is possible for all three types of specialization to be present in the trade flows between regions. In most cases, it is probably best to study each industry in each region separately to determine what the driving force of trade is in order to determine optimal public policy.

Market Structure and Scale Economies. Central to any study of IIT is the issue of market structure. For a traditional H-O framework, one must assume a perfectly competitive market structure with constant returns to scale. This assumption is too restrictive for more complex economies where scale economies are important and market imperfections rampant. Innovations in the theory of industrial organization have allowed for examining alternate market structures and IIT.

With respect to economies of scale, there are different types of increasing returns. Marvel and Ray (1987) and Ethier (1979) state that increasing returns due to *internal economies* (increasing returns at the firm level) do not lead to increases in IIT. Instead, some authors focus on *external economies* (at the industry level) as a more important factor in IIT. In this case, increasing returns arise due to market concentration, larger markets, or decreased transportation and information costs. Trade in intermediate inputs, or vertically integrated trade, also becomes possible with external economies of scale (Helpman and Krugman 1995).

As mentioned previously, the effect of scale economies on IIT depends on industry characteristics. Certain industries more than others would have scale economies leading to IIT. Lancaster (1980) stresses that monopolistic competition is the most competitive market structure in industries characterized by diverse consumer preferences and production specifications, but not in all cases does the presence of scale economies imply IIT. Hummels and Levinsohn (1995) provide a useful method for classifying the industries most likely to contribute to IIT, depending on the nature of scale economies. They state that industries with a small number of firms are most likely oligopolistic in nature. On the other hand, industries with a large number of firms likely exhibit increasing returns to scale. In the second case, product differentiation is more likely to occur, leading to increases in IIT.

IIT and the Study of Regional Economies. One of the most important contributions of the study of IIT has been a renewed focus in regional economies. Traditional trade theory largely ignores spatial issues: such as industry location, shared borders, and agglomeration effects. However, along with new trade theory, many authors have "rediscovered" the geography of trade (Krugman 1991a). Industry linkages and agglomeration economies in highly complex economies cause regions to become more interdependent, and can further intraindustry specialization. Krugman (1991b, 1994) discusses the development of pecuniary externalities that develop from demand and supply linkages between firms in a given region. Backward linkages arise as manufacturing locates to an area with large nearby demand: demand is generated from the concentration of increased manufacturing production. Forward linkages arise because it then becomes more desirable to live and produce near a concentration of manufacturing production. Due to transportation costs, manufactured goods are cheaper near to where they are produced. As intraindustry specialization increases, so should IIT.

What sort of regional developments are taking place in the Midwest and how do they potentially affect levels of IIT? According to Hewings et al. (1998), for the Chicago region, internal interaction (within the Chicago MSA) is being replaced by external interaction—both with the Midwest and beyond. In particular, increases in trade of intermediate inputs between the Chicago region and the rest of the Midwest are likely to increase. This process was referred to by Okazaki (1987) as *hollowing-out*. Until now, there has not been an explicit link between this phenomenon and new trade theory, but is likely that any growing interdependence of the Midwestern states occurs concomitantly with an increase in IIT.

Some Welfare Concerns. How does the advent of IIT affect consumers? One of the biggest predictions of IIT is the proliferation of product varieties. Most economists agree that the increases in product variety, ceteris paribus, can provide a benefit to the consumer in that more variety leads to increased utility. However, Greenaway and Tharakan (1986) theorize that there may be a "socially optimal" level of product variety, beyond which no further gains are realized. Thus, the effect of IIT on consumers is also complex.

Conventional trade theory has some specific predictions for changes in income distribution; namely, that the real income of the relatively scarce factors of production will decline as trade increases between regions. In contrast, new trade theory emphasizes ways in which intraindustry trade can potentially offset the costs of income distribution. Two types of gains can occur from IIT. First, an increase in IIT can increase the overall volume of trade. Secondly, the ability for firms to specialize within an industry can result in increased production through the realization of scale economies, which in turn can have beneficial impacts on the employment within that industry (Greenaway and Tharakan, 1986). Though increased trade can certainly have some disruptive effects on regional employment, they could be offset by the gains from intraindustry specialization. Helpman and Krugman (1985) state that this case is most likely when countries are sufficiently similar in factor endowments and scale economies are important to production. In this case, changes in relative factor prices are moderate, and the gains from specialization directly offset income redistribution effects. Indeed, increased IIT, it has been claimed, has enabled the OECD countries to undertake the trade liberalization programs of GATT precisely *because* those economies are so similar and distributional effects minimal (Krugman 1991a,b).

Thus, in terms of welfare concerns, again an industry-specific study may be worthwhile. In determining present and future employment changes for a specific industry within a specific region, it would be useful to view trade flows to and from that industry to determine whether intraindustry specialization plays an important role in the development of that industry and its trade patterns. Then, the level of IIT can aid with economic forecasts and long-term welfare implications for that industry.

6.3 IIT and Midwestern Trade

What drives trade between states in the Midwest? How are key industries changing over time, and how can new trade theory aid in making sense of complex trade flows? Appendices 1 through 3 summarize some quantitative exploration of these trade flows.

6.3.1 The Grubel-Lloyd Index of Trade Overlap

Perhaps the most important descriptive measure of IIT is the Grubel-Lloyd (GL) index of trade overlap. This index is measured as (Grubel and Lloyd 1975, Stone 1997):

$$SIIT_{jk} = 1 - \left\{ \left[\sum_i \left| X^e_{jki} - M^e_{jki} \right| \right] / \left[\sum_i \left(X^e_{jki} + M^e_{jki} \right) \right] \right\} \tag{6.1}$$

where:

$$X^e_{jki} = X_{jki} \left\{ \left(X_{jk} + M_{jk} \right) / 2X_{jk} \right\}$$

$$M^e_{jki} = M_{jki} \left\{ \left(X_{jk} + M_{jk} \right) / 2M_{jk} \right\}$$

and where: j is the country, k is the time period, and i is the industry. This index displays the level of trade *within* an industry relative to trade *between* industries. A value of 1 would imply perfect trade overlap, or that the value of that region's exports from a given industry was equal to the value of imports to that same industry. A value of 0 would imply perfect specialization within that industry (that the value of either exports or imports was equal to zero). Comparing the GL indices for the five Midwestern states is a good point of departure for understanding trade flows within this region. Appendix 1 summarizes these findings. The data used in calculating these indices come from the Commodity Flow Survey (U.S. Bureau of the Census and U.S. Department of Transportation Statistics, 1996) for the years 1969-1993. These data were aggregated at the two-digit level of the Standard Industrial Classification scheme (SIC) and integrated with consistently developed input-output tables for the same states.

Thus, one can roughly assume that for a given industry, a value of the GL index approaching 1 would imply a predominance of IIT. Conversely, a value approaching 0 may imply trade driven by other causes, such as relative factor endowments (as in a Heckscher-Ohlin framework). For each of the five states, five industries with the highest (trade overlap) and lowest (trade driven by industry specialization) indices are reported. In addition, the state of destination is reported[3]. As predicted by new trade theory, some of the more "high-tech" industries appear in the first column – that of high trade overlap; e.g., fabricated metal, transportation equipment, machinery, transportation equipment and food or kindred products (agricultural processing). Conversely, in the column reporting more specialized trade, some industries appear that are more natural-resource based, or have lower levels of high-tech production methods; e.g., coal, textile mill products, pulp or paper products, metallic ores and furniture and fixtures. However, these results are somewhat equivocal. In a few cases, an industry that exhibits a

---

[3] The abbreviation RUS stands for "Rest of United States," i.e., any domestic state other than the five Midwestern states.

high level of trade overlap for one state is specialized in another state; e.g., photographic and optical instruments, leather or leather products, and clay, concrete glass or stone. This finding perhaps points to the complexity of these trade flows. It is likely that trade driven by both intraindustry specialization and comparative advantage occurs. Another interesting finding is that for all states, most of the IIT is directed to other states in the Midwest. For Illinois, Ohio and Wisconsin, more of their trade to the Midwest is driven by IIT, while their trade to states outside the Midwest is predominantly specialized. This observation underscores the importance and interdependence of trade flows among states within this region and further suggests that agglomeration effects are being manifested at the multistate level rather than for individual metropolitan or state economies.

It should be noted that several authors have addressed problems with the GL index. Nilsson (1997) presented two major problems with the measurement of IIT. The first is the inappropriate grouping of industry activities. He proposed an alternative measure, indicating that the volume of intraindustry trade between two countries r and s may be divided with the total number of products they trade with each other to yield a measure of the average level of intraindustry trade per product group. Further, a dynamic GL index was suggested by Brülhart (1994) based on the concept of marginal intraindustry trade (MIIT) to address the problem of changes in the trade flows. A distinction between horizontal and vertical intraindustry trade was suggested by Aturupane et al. (1999).

6.3.2 Income Trends in the Midwestern States

Appendix 2 summarizes changes in per capita income from 1969-1993. These data were derived from REIS data (Regional Economic Information System: 1969-1994, U.S. Department of Commerce, Bureau of Economic Analysis, Washington, D.C.). Income was reported in constant 1987 U.S. dollars. An index of percentage change in per capita income during this time period is reported. A value equal to 1.00 would imply no change. For all the five states, income increased over this time period, but at a lower rate than the national average.

6.3.3 Vertically Integrated Trade in the Midwest

Hummels et al. (1998, 1999) recently conducted a study to estimate the degree of vertically integrated trade among OECD countries. They defined vertically integrated trade as trade for goods that are produced in multiple sequential stages, and that cross a border more than once (Hummels et al. 1998). In such cases, firms exploit both economies of scale and locational advantages. Economies of scale are achieved if the scale of production can increase as certain regions specialize in the production of a certain stage (or stages) of a good's production. Locational advantages are realized by locating production according to access to particular markets or by taking advantage of regional wage differentials. The authors also found that the degree of vertically integrated trade varies considerably among industries. For their sample[4], they determined that the following industries exhibited the greatest level of vertically specialized trade: motor vehicles, shipbuilding, aircraft, industrial chemicals, nonferrous metals, petroleum and coal products. Conversely, those industries with the lowest levels of vertical trade were agriculture, mining, wood products and paper products (Hummels et al. 1998). Following their results, some indices of growth in certain industries in the Midwest were calculated. The industries that may exhibit high levels of vertically integrated trade in the Midwest include: motor vehicles; manufacturing; fabricated metals; chemicals; petroleum products; and transportation and utilities. Conversely, industries for which a lower level of vertically integrated trade include: farm products, mining, lumber and wood, and paper products. Appendix 3 summarizes indices of growth for these selected industries. The data were obtained from REIS Gross State Product tables (Regional Economic Information System: 1969-1994, U.S. Department of Commerce, Bureau of Economic Analysis, Washington, D.C.) at the two-digit SIC level using constant 1987 U.S. dollars.

For all the Midwestern states, more industries in the group hypothesized to have higher levels of vertically integrated trade experienced growth relative to national changes in these industries. However, for all the states except Illinois, growth was experienced in at least one industry that would not likely have high levels of vertically integrated trade. Therefore, growth in vertically integrated trade is certainly not the only driving force of production increases in the Midwest for the period 1977 to 1991. However, for certain industries, such as chemicals and petroleum products, growth higher than the national average was seen by all but one state (Ohio and Illinois, respectively). Hummels et al. (1998) also noted strong trends in these two industries. In order to further study vertically integrated trade in the Midwest, much more complete data would be needed, but from these growth indices, one can assume that such trade is likely to be important to the region.

6.4 Directions for Further Study

Trade flows within the Midwestern states need to be examined in great detail. Countless studies have been conducted regarding the welfare effects of international trade, but there is a paucity of such studies at the regional level. Some of the evidence presented in this paper indicates that significant regional differences do exist, though in general the economies of the Midwest are becoming more similar over time. What will be the outcome of increased trade flows within the Midwest? Will they be mutually beneficial and lead to per capita income across the region? Or will some states fare better than others because of initial advantages in "growth" industries, while others specializing in more traditional industries decline in relative terms? In order to begin to answer such questions, much

---

[4] Based on OECD trade data for the years 1968 to 1990 using the 4-digit SITC classification scheme.

more must be known about these trade flows. Based on the findings of Hummels et al. (1998, 1999), it appears that some industries that have proven to be contributors to increased production and economic growth are *also* faring well in the Midwest, but this finding is not found across the board. One policy implication may be to identify those industries that are most likely to expand production, scale and growth through their trade links within the Midwest.

Another related question involves the role of economies of scope and agglomeration. Many authors have argued that in high-tech, decreasing cost industries, economies of scale are significant. Trade in intermediate inputs, intraindustry trade, and vertically integrated trade all facilitate such growth, and such trade is only likely to increase over time. New trade theory is useful to a study of regional economies because it focuses on the role of industrial organization and market structure in fostering trade flows that are otherwise inexplicable. If initial factor endowments led to industrial specialization, the exploitation of scale economies leads to product differentiation and thus bilateral intraindustry trade. In addition, the renewed focus of many on the geography of trade flows leads to a better understanding of the role of industry location and path-dependent regional development. However, attention to issues of clustering as a development strategy may be misplaced in a context in which the costs of spatial interaction across considerable distances are minimal. Recent empirical analysis of the Brazilian Northeast region and the Midwest of the US (Magalhães et al. 2001) revealed a strong contrast in the degrees of interaction and thus pointed out the continued important role that connectivity plays in understanding trade. The methodology of feedback loop analysis proposed in Sonis et al. (1995) offers a complementary methodology to explore more extensively the vertically integrated inter-regional or inter-country trading patterns.

References

Abdel-Rahman K (1991) Firms' competitive and national comparative advantages as joint determinants of trade composition. Weltwirtschaftliches Archiv 127:83-97

Aturupane C, Djankov S, Hoekman B (1999) Horizontal and vertical intraindustry trade between Eastern Europe and the European Union. Weltwirtschaftliches Archiv 135:62-81

Balassa B, Bauwens L (1987) Intraindustry Specialisation in a Multi-Country and Multi-Industry Framework. The Economic Journal 97:923-939

Brülhart M (1994) Marginal intraindustry trade: measurement and the relevance for the pattern of industrial adjustment. Weltwirtschaftliches Archiv, 130:600-613

Commodity Flow Survey (1996) U.S. Bureau of the Census and U.S. Bureau of Transportation Statistics, Washington, D.C.

Davis DR (1995) Intraindustry trade: a Heckscher-Ohlin-Ricardo approach. Journal of International Economics, 39:201-226

Ethier W (1979) Internationally Decreasing Costs and World Trade. Journal of International Economics 9:1-24

Flam H, Helpman E (1987) Vertical product differentiation and North-South Trade. American Economic Review 77:810-822.

Fujita M, Krugman P, Venables A (1999) The spatial economy. Cambridge University Press, New York

Görg H (2000) Fragmentation and trade: US inward processing trade in the EU. Weltwirtschaftliches Archiv, 136:403-422

Greenaway D, Tharakan PKM (1986. Imperfect competition and international trade: The policy aspects of intraindustry trade. Wheatsheaf Books, Sussex

Greenaway D, Hine R, Milner C (1995) Vertical and horizontal intraindustry trade: a cross industry analysis for the United Kingdom. Economic Journal 105:1505-1518

Grubel, HG, Lloyd PJ (1975) Intraindustry trade. Macmillan, London

Hanson GH (1996) Localization economies, vertical organization and trade. American Economic Review, 86:1266-1278

Helpman E and Krugman PR (1985). Market Structure and Foreign Trade. MIT Press, Cambridge, MA

Hewings GJD, Israilevich PR, Okuyama Y, Anderson DK, Schindler GR, Foulkes M, Sonis M (1997) Returns to scope, returns to trade and the structure of spatial interaction in the US Midwest, Discussion paper 97-P-3, Regional Economics Applications Laboratory, University of Illinois, Urbana.

Hewings GJD, Sonis M, Guo J, Israilevich PR, Schindler GR (1998) The Hollowing-Out Process in the Chicago Economy, 1975-2011. Geographical Analysis 30:217-233

Hummels D, Levinsohn J (1995) Monopolistic competition and international trade: Reconsidering the evidence. The Quarterly Journal of Economics 110:799-836

Hummels D, Rapoport D, Yi K-M (1998) Vertical specialization and the changing nature of world trade. Economic Policy Review (Federal Reserve Bank of New York), June, 79-97

Hummels D, Ishii J, Yi K-M (1999) The nature and growth of vertical specialization in world trade. Working paper, Federal Reserve Bank of New York.

Jones RW, Kierzkowski H (1990) The role of services in production and international trade: a theoretical framework. In: Jones RW, Krueger AO (eds) The political economy of international trade. Blackwell, Oxford

Jones RW, Kierzkowski H (2001) A framework for fragmentation. In: Arndt SW, Kierzkovski H (eds), Fragmentation: new production patterns in the world economy. Oxford University Press, New York

Krugman PR (1990) Rethinking international trade. MIT Press, Cambridge, Massachusetts

Krugman PR (1991a) Geography and trade. Leuven University Press, Leuven, Belgium

Krugman PR (1991b) Increasing returns and economic geography. Journal of Political Economy, 99:483-499

Krugman PR (1994) Complex landscapes in economic geography. American Economic Review, AEA Papers and Proceedings, 84:412-416

Lambertini L (1997) Intraindustry trade under vertical product differentiation. Keio Economic Studies, 34, 51-69.

Lancaster K (1980) Intraindustry trade under perfect monopolistic competition. Journal of International Economics, 10:151-175

Magalhães A, Sonis M, Hewings GJD (2001) Regional competition and complementarity reflected in relative regional dynamics and growth of GSP: A comparative analysis of the northeast of Brazil and the midwest states of the U.S. In: Guilhoto JJM, Hewings GJD (eds) Structure and structural change in the Brazilian economy, Ashgate, London

Martin R (1999) The new ‘geographical turn’ in economics: some critical reflections. Cambridge Journal of Economics, 23:65-91

Marvel HP and Ray EJ (1987) Intraindustry trade: sources and effects on protection. Journal of Political Economy, 95:1278-1291

Nilsson L (1997) The measurement of intraindustry trade between unequal partners. Weltwirtschaftliches Archiv 133:554-565

Okazaki F (1987) General verification of the logit-type stochastic trade pattern using intertemporal, interregional input-output data. Papers of the Regional Science Association 63:1-11

Price VC (2001) Some causes and consequences of fragmentation. In Arndt SW, Kierzkovski H (eds), Fragmentation: new production patterns in the world economy. Oxford University Press, New York

REIS. Regional Economic Information System, (1969-1994) U.S. Department of Commerce, Bureau of Economic Analysis, Washington, D.C.

Sonis M, Guilhoto JJM, Hewings GJD (1995) The Asian economy: trade structure interpreted by feedback loop analysis. Journal of Applied Input-Output Analysis 2: 24-40

Stone LL (1997) The growth of intraindustry trade. Garland Publishing, New York

Tharakan PKM (1984) Intraindustry trade between the industrial countries and the developing world. European Economic Review, 26:213-227

Thom R, McDowell M (1999) Measuring marginal intraindustry trade. Weltwirtschaftliches Archiv 135:48-61

Thompson W (1965) A Preface to Urban Economics. RFF, Johns Hopkins University Press, Baltimore, MD

Venables AJ, Rice PG, Stewart M (2003) The Geography of intraindustry trade: Empirics. Topics in Economic Analysis and Policy 3 (article 11) [Berkeley Electronic Press]

Appendix 1 Grubel-Lloyd Index Results

	Most Overlap	SIC	State of Destination	Most Specialization	SIC	State of Destination
Illinois	Farm Products	01	Indiana	Fresh Fish	09	Indiana
	Lumber & Wood Prods	24	Indiana	Coal	11	RUS
	Clay, Concrete, Glass or Stone	32	RUS*	Ordinance or Accessories	19	RUS
	Fabricated Metal Products	34	Indiana	Petroleum or Coal	29	RUS
	Machinery	35	Indiana	Clay, Concrete, Glass or Stone	32	RUS
Indiana	Farm Products	01	Illinois	Fresh Fish	09	Illinois
	Non-metallic Minerals	14	Ohio	Leather or Leather Products	31	Illinois
	Food or Kindred Products	20	RUS	Textile Mill Products	22	Ohio
	Clay, Concrete, Glass or Stone	32	Illinois	Furniture or Fixtures	25	Ohio
	Photographic, Optical Instruments	38	Ohio	Coal	11	Illinois
Michigan	Machinery Excluding Electrical	35	Ohio	Textile Mill Products	22	Ohio
	Food or Kindred Products	20	RUS	Apparel or Finished Textiles	33	Illinois
	Leather or Leather Products	31	Ohio	Non-metallic Minerals	14	Indiana
	Primary Metal Products	33	RUS	Electrical Machinery	36	Illinois
	Fabricated Metal Products	34	Ohio	Photographic, Optical Instruments	38	Indiana
Ohio	Non-metallic Minerals	14	Indiana	Metallic Ores	10	RUS
	Rubber or Miscellaneous Plastic	30	Wisconsin	Ordinances or Accessories	19	RUS
	Transportation Equipment	37	Illinois	Apparel or Other Finished Textiles	23	Wisconsin
	Fabricated Metal Products	34	Indiana	Waste or Scrap Materials	40	RUS
	Machinery Excluding Electrical	35	Michigan	Misc. Freight Equipment	41	RUS
Wisconsin	Rubber or Misc. Plastic Products	30	Ohio	Farm Products	01	Ohio
	Primary Metal Products	33	RUS	Ordinance or Accessories	19	RUS
	Fabricated Metal Products	34	Indiana	Pulp, Paper or Allied Products	26	Michigan
	Electrical Machinery Equipment	36	Indiana	Leather or Leather Products	31	RUS
	Photographic, Optical Instr.	38	Illinois	Misc. Freight Equipment	41	Illinois

* RUS = “Rest of U.S.”

Appendix 2 Income Trends in the Midwestern Region, 1969 – 1993

Average Percent Change in Income Per Capita, 1969 - 1993 Relative to Total Change in U.S. Income	
Illinois	0.88
Indiana	0.90
Michigan	0.88
Ohio	0.86
Wisconsin	0.94

Average and Relative (to US) Percent Change in Income Per Capita, 1969 - 1993
(1.00 = no change)

	Absolute	*Relative*
Illinois	1.31	0.88
Indiana	1.34	0.90
Michigan	1.31	0.88
Ohio	1.28	0.86
Wisconsin	1.41	0.94
U.S. Total	1.51	

Appendix 3 Vertical Integration

Industries With a High Degree of Vertically Integrated Trade Internationally
Average Growth Indices for 1977 - 1991 Weighted by US Total Growth (1.00 = no change)

	Motor Vehicles	Manufacturing	Fabricated Metals	Chemicals	Petroleum Products	Transportation and Utilities
Illinois	1.42	0.86	0.88	1.01	0.65	0.93
Indiana	0.95	0.90	0.88	1.06	1.10	0.91
Michigan	0.89	0.77	0.87	1.02	1.15	0.84
Ohio	1.10	1.10	0.91	0.98	1.10	0.86
Wisconsin	0.85	0.99	1.10	1.11	1.13	0.91
US Total	0.82	1.11	1.07	1.12	1.08	1.24

Industries With a Low Degree of Vertically Integrated Trade Internationally
Average Growth Indices for 1977 - 1991 Weighted by US Total Growth (1.00 = no change)

	Farm Products	Mining	Wood	Paper Products
Illinois	0.65	0.86	0.99	0.89
Indiana	0.72	0.90	1.12	0.85
Michigan	0.93	0.77	1.08	0.86
Ohio	0.84	0.88	1.26	0.92
Wisconsin	0.95	0.99	1.13	1.04
US Total	1.14	0.99	1.08	1.11

7 Globalization, Regional Economic Policy and Research

Edward Feser

Department of Urban and Regional Planning, University of Illinois at Urbana-Champaign, Champaign, Illinois USA

7.1 Introduction

While the case for increasing the degree of integration in the world economy is strong (e.g., Sen 1999; Irwin 2002; Bhagwati 2004), it is based on the notion of net benefits.[1] Workers, industries and places absorb trade-related gains and losses differently and must adjust to greater or lesser degrees to the liberalizing international economic environment. Because of the high degree of industrial specialization and concentration in the space economy, some states and regions' potential losses are greater than their gains, at least in the short- to medium-run. While the losses seem eminently visible—evidenced in bankruptcies, closed factories, and unemployment lines—the benefits are typically diffuse. It still the case that net gains are not a fait accompli, despite a high degree of confidence in the net positives of trade within mainstream economics (Samuelson 2004). Moreover, core-periphery and cumulative causation theories and findings suggest that there are circumstances in which opening regions to import competition can exacerbate regional income inequalities (Venables 1998; Fujita and Hu 2001; Meardon 2001; Hu 2002; Mansori 2003).[2] The differential regional impacts of globalization are therefore important and need to be fully understood, along with the aggregate economic benefits. Also requiring careful thought is what local strategies, if any, are necessary to maximize the benefits of globalization for regional businesses and workers. In a globalizing world of scarce public sector resources, what is the appropriate subnational economic policy response?

---

[1] Bradford, Grieco and Hufbauer (2005) estimates the income gains to the U.S. economy of 50 years of post-WWII economic liberalization at \$2,800 to \$5,000 per head in 2003.

[2] Studies showing that a reduction of trade barriers leads to a deconcentration of economic activity—and therefore potential reduction in regional inequalities—include Hanson (1998), Sánchez-Reaza and Rodríguez-Pose (2002), and Pernia and Quising (2003).

The volume of literature on modern trends in globalization is already enormous. Yet, the question of what globalization means for economic policy making at the subnational scale has received comparatively little systematic attention. Perhaps that is because decisions about desirable levels of economic integration are made at the national level. Yet, regions within nations face most of the practical challenges of adjusting to the new economic order. Note that one needs to distinguish between scholarship that focuses on the regional impacts of globalization and that which concerns appropriate regional-level policy responses. The former literature is extensive, incorporating work on topics such as the influence of globalization on regional inequality (noted above), the impact of free trade agreements on local communities (Glasmeier and Conroy 1994), the implications of foreign direct investment for regional growth (Glickman and Woodward 1988; Head et al. 1995), and the rise of multinationals and their roles in competitive regional clusters (Ivarsson 1999). Systematic work on regional policy implications, by contrast, is limited and frequently indirect, emphasizing how regional growth is increasingly a function of international linkages without consideration of whether that implies any significant and necessary *change* in local development policy options or priorities. An open economy is nothing new to regions, after all, and perhaps the current set of policy frameworks and evaluation methods are adequate for regional scale strategic planning and policy making as global economic integration proceeds. If that were the case, planners and policy makers' tasks would be easier. They could concentrate on doing better what they have been doing already rather than changing approaches altogether.

To wit, this paper considers two questions. First, are there any *unique* implications of growing global economic integration for development planning and policy making at the regional level? The focus is on the United States context although some of the more general findings are relevant to other highly industrialized countries with active local and regional economic development institutions. Related issues include whether globalization is appreciably different today than it used to be and whether it means anything more, from the perspective of a given region, than heightened competition for its resident industries and related challenges of more rapid macro-regional structural change and adjustment. Second, what kinds of spatial empirical research and model building would be most valuable to regional policy makers faced with designing programs and making specific allocative investment decisions in the face of growing economic integration? Putting on the shoes of a regional policy maker, what kinds of problems should regional scientists pursue?

7.2 Today's Globalization

The global economy becomes more integrated with relative increases in the international exchange of goods, services, money, investment, labor, and technology. Such exchanges are realized in foreign trade, multinational direct foreign investment, movements of short-term portfolio funds, technology embodied in human and physical capital, and international migration. The real drivers of economic integration, however, are the emergence of new markets, the erosion of old markets, shifts in public policies that reduce barriers to exchange, trends in corporate strategy and industrial organization, the ongoing improvement of transportation and communications infrastructure, and technological change. For city and state officials, the most visible and politically compelling aspect of globalization is the associated labor market impact: job and wage growth and decline in the face of enhanced international competition. Generally, negative labor market outcomes are perceived more keenly than positive ones. Not only are wealth effects harder to trace, but confusion in the economic development community over job creation versus wealth creation is probably as acute as it has ever been (Malizia 1994).

Determining the appropriate regional policy response to globalization is no easy task. Rarely is there a single cause of any given change in regional labor markets and, even if there were, even rarer are the data necessary to isolate that cause. Feenstra (1998) notes, for example, that the impact of globalization on employment and wages is often observationally equivalent to the effects of other influences, such as technological change or market shifts. A company observed releasing workers locally and establishing a production base in a foreign country in order to rationalize domestic demand contraction and tap foreign demand growth looks very similar to the company relocating elements of its business services functions offshore in response to high local costs. In both cases employment is reduced domestically and expanded overseas. However, the forces behind each move, demand shifts in the first case and a combination of cost cutting and technological change in the second, have very different implications for regional growth and related development policy.

The complexity of cause and effect are exacerbated by the inconsistent use popular media and academic literature of the many terms used to describe globalization trends—offshoring, outsourcing, vertical specialization, and the like—as well as the common practice of extrapolating from isolated but highly visible cases of trade-related job loss. There is probably a natural tendency among public officials to over-estimate their jurisdiction's vulnerability to the negative effects of globalization in the wake of large business closures, much as those who have personally experienced a natural disaster or major accident tend to overestimate the risk of subsequent hazardous events. A single major trade-related plant closure in a locality can do much to derail efforts to think clearly about strategies for maximizing the benefits of global integration. In a typically charged local political environment, what may do most to limit the propagation of bad policy following such events may be the simple fact that there is ultimately little subnational governments can do influence market liberalization trends. Unfortunately, the political incentive to be viewed as "protecting" regional industries and workers combined with general ignorance of globalization dynamics also means that good regional policies may not get adopted either.[3]

---

[3] Carol Conway, Associate Director of the Southern Growth Policies Board, an organization of southern governors, in a recent speech in Tokyo: "The current poster child for trade-related job loss is PillowTex, which in one fell swoop laid off some 6,000 workers in ru-

A commonly debated question in the mid-1990s was whether the current *level* of economic integration is more extensive than it was immediately prior to World War I. For example, relying on data showing that the ratio of U.S. trade to GDP in the 1980s was roughly what it was at the beginning of the 20th century, Krugman (1996, p. 120) delivered a now well-known caution for analysts of the American situation:[4]

> . . .one should have some historical perspective with which to counter the silly claims that our current situation is completely unprecedented: the United States is not now and may never be as open to trade as the United Kingdom has been since the reign of Queen Victoria.

While the relative volume of goods and services flows addresses only the trade dimension of economic integration, it is often regarded as a barometer of globalization trends more generally. Roughly ten years after Krugman's piece, it is now clear that U.S. goods and services trade today does, in fact, significantly exceed relative levels at any time in the last century (Gresser and West 2001; Perera-Tallo 2003). Moreover, U.S. barriers to trade continue to come down. But more significant for the purposes of this paper is what seems to be an emerging consensus that the *nature* of economic integration is substantially different today than in the past, at least from the perspective of the U.S.[5] Table 1 summarizes this consensus by highlighting twelve trends identified in the literature as distinguishing today's globalizing economy. The trends fall into five basic categories. Table 2 clarifies some commonly used terminology.

The first of the five categories captures sectoral shifts in the pattern of trade. They include significant growth of intraindustry trade ("intratrade") and a dramatic increase in the ratio of merchandise trade to merchandise output (Krugman 1995; Bordo et al. 1999). With respect to the latter, while goods producing sectors such as agriculture, mining and manufacturing constitute a declining share of overall U.S. gross domestic product (GDP), the amount of international trade within those sectors has increased substantially. The ratio of goods exports to agricultural, mining and manufacturing GDP was 38.1 percent in 2002, up from 37.8 percent in 1999 and about 15 percent in 1970 (Irwin 2002; Council of Economic Advisers 2004). Goods sectors are exporting, importing, offshoring and outsourcing more—in relative terms—than ever before, with significant impacts for U.S. regions that formerly captured much of their associated value chains.

---

ral North Carolina and Virginia. Never mind that the underlying cause of the corporation's demise was a foolish, debt-riddled merger and acquisition strategy and the banks' equally foolish willingness to believe in rosy growth scenarios. . .globalization is taking all the blame. Add to that the newest scare—thousands of software jobs moving to India—and you have a powerful political backlash in the making" (Conway 2003, pp. 3-4).

[4] The quotation also appears in partial form in (Krugman 1995).

[5] According to Bhagwati (2004, p. 13): ". . .the complacent view that there is nothing new about globalization is simply wrong."

Table 7.1. Summary of factors distinguishing today's globalization trends

Sectoral shift	• Substantial growth of intraindustry trade • Merchandise sector trade dependence • Increase in services trade
Organizational shift	• International component outsourcing • Rise in offshoring • Increase in intrafirm international transactions
Spatial shift	• Vertical specialization on a large scale • Emergence of supertrading countries • Growth of exports from low- to high-wage countries
Adjustment pressure	• Increased speed of response to emerging international market opportunities
Government role	• State activism as globalization driver • Stronger perception of government responsibility for mitigating adverse effects of economic integration and trade

Also in the category of sectoral shifts are the significant increase in services trade and the reduction in the types of services that are "non-tradable," both closely-related trends that can be attributed to advances in information technology and the continued digitization of services. The rise of services trade means that activities that were once viewed as domestic growth engines that will replace declining commodity production are now subject to growing international competition. Corporate offshoring decisions, such as that recently made by Indiana-based Cummins Incorporated to set up a high tech center in India to speed computerized design and testing of truck engines (Oneal 2004), are wreaking havoc with the conventional view that the U.S. will necessarily capture high value, technology- and knowledge-intensive segments of key value chains.[6]

The second category of globalization trends concern shifts in business organization strategies and practices, including increases in component outsourcing to international suppliers and subsidiaries and increased offshoring of services like information processing and customer support (McKinsey Global Institute 2003; Farrell 2004). Related to those two trends is the rise of the multinational firm as the globalization lynchpin (McCann and Mudambi 2004). It is hard to overstate how dominant multinational firms are as influences on globalization trends. Hanson and Slaughter (2003) report that 11,151 businesses in the U.S. were part of a multinational firm in 1999. That was roughly 1/20th of 1 percent of the total of

---

[6] There is a heavy "stages" bias (e.g., Rostow 1960) in local and regional economic development thinking in the U.S. The view that U.S. regional economies will evolve naturally toward high technology and knowledge-based activities as less developed countries like China and India assume production of labor-intensive commodities is deep-seated. High-tech international outsourcing to low-wage countries is challenging this stages view, heightening fears of pending dramatic reductions in the U.S. standard of living.

over 24 million U.S. businesses. However, the authors estimate that multinationals accounted for 80 percent of U.S. goods exports, 66 percent of U.S. goods imports, 42 percent of U.S. capital investment, and 82 percent of U.S. industrial research and development in that year (Hanson and Slaughter 2003, p. 5). In addition, multinationals accounted for about 25 and 32 percent, respectively, of U.S. non-bank employment and GDP. Multinationals' influence therefore well exceeds their modest numbers. They drive globalization whether gauged in terms of production activity or as conduits for technology diffusion, global knowledge spillovers, innovation, and skilled worker international migration.

Accompanying sectoral and business organizational trends are distinct spatial shifts in the pattern of trade and location of industries (Venables 1998). Most obvious is the emergence of supertraders or countries with extremely high ratios of trade to GDP. Examples are Singapore, Hong Kong, Malaysia, Belgium, and Ireland. Also significant has been growth in commodity exports from low-wage to high-wage regions, especially as the former have become satellite production locations for multinationals serving industrialized markets. But perhaps the trend with the most significance for local development strategy is increasing vertical specialization in trade, or the growing tendency of trading countries to specialize in specific pieces or segments of product value chains (Balassa 1967; Hummels et al. 1998; Hummels et al. 2001; Yi 2003). Sometimes described as "slicing up the value chain," vertical specialization is one result of company outsourcing, itself made possible by new information technologies, better information technology infrastructure, management innovations, and lower transportation costs. Vertical specialization implies increasing business locational flexibility, but in a world in which there are still advantages to agglomeration. Trade specialization may occur because of spatial externalities conferred by the clustering of specialized industries, though it could also be driven by other factors such as traditional comparative advantage. A given product is no longer produced by a single company and its suppliers and subsidiaries in a single location (a single agglomeration) but rather by specialized clusters of companies (multiple agglomerations) in different parts of the world. In essence, what we are seeing occurring on a global scale is the same value chain specialization that has characterized the regional distribution of economic activity within countries for a long time.

The fourth category of globalization trends refers to the nature of adjustment to the new economic order. Relaxed capital controls, advanced information technology, and improvements in transportation infrastructure mean that goods, services, and capital flows are increasing significantly (Stiglitz 2002). Loosened capital controls in less developed countries have brought with them the potential for huge financial management problems that are not so much unprecedented in their scale as in the dramatic and uncontrollable speed with which they emerge (Bhagwati 2004). Moreover, because there is competition from a much larger pool of potential rivals, new business opportunities can materialize and then evaporate very quickly. That creates a challenging business environment, especially for less developed countries and regions. It implies shorter cycles of regional competitive advantage and therefore more rapid structural change and greater churn in labor markets. In tandem, there is a lower margin for error for development strategies aimed at encouraging local businesses to tap specific new markets.

The fifth and final category of trends refers to citizen and business perceptions of the government's role in addressing the negatives associated with economic integration. While popular understanding of the role of the welfare state is less commonly discussed in the globalization literature, it has particular significance from a policy perspective. Bhagwati (2004) argues that earlier world economic integration was driven more by technological developments in transportation and communication than policy changes (see also Krugman 1995).[7] If today's integration is indeed propelled to a much greater degree by state activism, it must be understood as a conscious policy choice that could, in principle, be reversed. How the impacts of integration are perceived by the public and leveraged (when positive) and mitigated (when negative) by governments is therefore especially important for determining globalization's future course. Moreover, because the interventionist role of the welfare state is much more widely accepted today, perceived compromises in its ability to deliver on its promises are viewed with greater concern (Bhagwati 2004).

In this environment, citizens' and businesses' heightened sense of economic insecurity is not just felt by national governments (Clark and Montjoy 2001). Many state governors in the U.S. face considerable pressure to lobby federal trade negotiators to slow down the pace of integration. The states confronting the most significant industrial restructuring are naturally those where that pressure is most intense. Although such states might benefit from an open and balanced discussion of how to adjust policies for a global era, they are frequently not the most propitious environments for such debates. Some state and local leaders have also discovered that federal trade policy can serve as a convenient scapegoat for weak economic trends at home. Because any scapegoat is less compelling if it is admitted that more fundamental forces could also explain lackluster economic trends, incentives for unbiased consideration of policies to maximize the benefits of globalization are not always very strong.

Taking these trends together, we have a set of unique features of the current growth of international economic integration: growing intraindustry trade, vertical specialization, the emergence of supertrading countries, stronger trading relationships between high- and low-wage nations, increased offshoring, heavy dependence on trade by merchandise sectors (particularly in the U.S., but elsewhere as well), growing trade in services, rapid growth in multinationals, and faster goods, services, and financial flows. Add to these the political context in the United States, including the strong role of the federal government in pushing integration, a pervasive sense of economic insecurity among existing and potential trading industries and their employees, the now well-established legitimacy and implied obligations of the welfare state, and political incentives to scapegoat trade policy at the state and local level. This is the context in which globally-savvy regional eco-

---

[7] Though note that Cooper (1995), in comments on Krugman (1995), makes a case that technological improvements in transportation are a much more important factor behind recent globalization trends than appreciated by many trade economists.

nomic development strategies must be discussed, designed, and implemented. What changes should be made in the mix of policies and programs promulgated in a less globalized economy? Below I argue that the answer hinges on having the right information and tools to properly evaluate options make necessary changes.

Table 7.2. Globalization concepts and terminology

Concepts	Description
Outsourcing	Decision by one company to contract with another company or subsidiary to provide services or products (inputs).
Offshore outsourcing (or offshoring)	An outsourcing arrangement established by a domestic company with a foreign contractor or subsidiary.
Vertical foreign direct investment (FDI)	Company invests in the relocation of a given stage of its production to a foreign country; may result in offshore outsourcing via a wholly or partially-owned subsidiary.
Vertically integrated multinational firm	Firm that uses its own foreign and domestic units to produce a good rather than arm's-length relationships (contracts) with other companies.
Horizontal specialization in trade	Pattern of trade that results when companies use entirely domestic inputs in exported products. Implies that trading countries specialize in the full value chain of selected products.
Vertical specialization in trade	Pattern of trade that results when companies use imported inputs in products that are subsequently exported (Balassa 1967; Findlay 1978; Sanyal 1983; Hummels et al. 2001). Involves a domestic company outsourcing part of the production of its exported good. Sometimes called "intra-product specialization," "global production sharing," or "slicing up the value chain." Implies that trading countries increasingly specialize in certain segments of value chains.
Hollowing out	The transfer of the production base of a given industry overseas or out of a region, leaving the service and/or R&D components. Driven by offshore outsourcing or vertical FDI and implies increase in vertical specialization in trade.

7.3 Globalization, Regional Economic Policy…

We should distinguish between development policy at the regional scale and regional policy as applied by national governments. With regard to the latter, two issues are of chief concern. The first is that decisions to reduce trade barriers imply an obligation by national governments to consider and mitigate net costs for specific workers and regions (Kletzer 2002; Rosen 2002). In the U.S., that is the rationale for the somewhat underutilized and arguably inadequate federal Trade Adjustment Assistance (TAA) program. Administered by the Employment and Training Administration of the U.S. Department of Labor, the TAA program provides compensation and re-employment assistance to workers dislocated from trade-impacted businesses. While such compensation is a people- rather than place-based strategy (Bolton 1992), it has a distinct regional distribution given the spatial concentration of sectors under heavy pressure from foreign competition. In addition to TAA, a case might be made for fiscal compensation to subnational governments since they often bear a substantial part of the cost of adjustment, principally in workforce training programs subsidized through community college systems. Such compensation already occurs on a modest scale in the U.S. through occasional backdoor development planning grants or pork barrel allocations to legislative districts impacted by trade. Such are the pragmatic means of obtaining necessary Congressional support for multilateral trade agreements like NAFTA.

The second major *national* regional policy issue is whether increased integration will lead to regional income convergence or divergence. The question has received relatively modest attention to date and is ripe for additional empirical research (Stroomer and Giles 2003). However, as in the case of compensation, the policy implications are principally national in scale, although there may be some things specific regions ought to do to avoid a path of cumulative decline.[8]

For guidance about what regions themselves should do differently, we could turn to various disparate literatures on the policy implications of selected processes influenced by globalization, such as the formation of industry clusters, the emergence of the knowledge or learning economy, and the importance of a high skilled technology-savvy workforce in advanced industrialized countries (e.g., Simmie and Sennett 1999; Enright 2000; Helmsing 2001). Unfortunately, globalization is more often the assumed economic context than the direct subject of investigation. Attempts to come to a systematic and comprehensive understanding of whether globalization *per se* warrants specific regional policy responses are surprisingly difficult to find. However, three relatively recent treatments are articles by Felbinger and Robey (2001) and Rondinelli, Johnson and Kasarda (1998) and a comprehensive report by the Southern Growth Policies Board (2003), an organization of Southern governors. They are worth considering in some detail.

Felbinger and Robey (2001) claim that three new assumptions must drive state and local economic development policy in a rapidly globalizing economy: the erosion of city and state jurisdictions as meaningful units of analysis; the increasing importance of focusing on building a flexible workforce; and the value of partnerships between governments, business, and academic institutions. The authors would reduce or eliminate training programs designed to provide specialized skills based on the needs of predicted growth industries. General skill building is

---

[8] An issue somewhere between a local policy issue and a national regional policy issue is the question of whether multilateral trade and investment agreements that address spending, regulatory and procurement policies will eventually make many state and local economic development programs illegal. See Stumberg and Schweke (1999).

more appropriate given the growing difficulty of predicting what new specialized skill demands will emerge. Building a "flexible and adaptable platform that could accommodate innovation and change" means creating "a flexible and trained workforce, an entrepreneurial business community, a 'built' and technological infrastructure in place, and governmental regulations that not only protect the citizenry but facilitate development and innovation" (2001, pp. 68-9). In general, Felbinger and Robey would redirect investments in particular firms and industries and to government basics like education and infrastructure. However, they do cite favorably states' increasing efforts to build "centers of excellence," specialized applied research centers aimed at boosting selected technology industries.

Rondinelli, Kasarda and Johnson (1998) begin by arguing that the globalized economy is characterized by the following factors: increasing mobility in the factors of production, thereby making regions more vulnerable to structural economic shifts; the importance of technology-related and knowledge-based growth for industrialized countries; the greater need for companies to expand their markets to be competitive; growing corporate strategic and organizational agility; and the emergence of global corporate alliances and manufacturing networks. They conclude that successful cities will be those that "provide the labor force, services, and infrastructure that allow locally based domestic and foreign-owned firms to participate more successfully in the international marketplace" (p. 71-2). Cities must invest in telecommunications and transportation infrastructure; develop education and training programs that produce both high skilled workers and workers capable of transitioning to new sectors quickly; improve the quality of life of the community through the provision of cultural and environmental amenities; promote entrepreneurship and technology development; strengthen cooperation among public and private organizations and stakeholders; adopt new forms of regional governance that reduce interjurisdictional competition; and address urban poverty through aggressive inequality reduction and the development of programs focused on education and literacy, family and child development, and drug and crime prevention.

Rondinelli et al.'s prescriptions call for a highly activist development strategy with a little bit of everything, though with emphasis on education and technology/innovation promotion.[9] Among the many specific interventions they advocate are small business incubators, technology transfer programs that assist businesses in the acquisition of information about new technologies and business practices, technology commercialization assistance programs for small and medium sized businesses, business retention and expansion programs, trade promotion and brokering programs, grants to small businesses to commercialize new products, high technology zones with quality infrastructure and favorable tax and regulatory

---

[9] "To meet the challenges of an open world economy, new initiatives will have to be taken in American cities to form internationally oriented communities centered on agile public and private organizations promoting innovation and creativity, not only in manufacturing, trade, and services, but also in the physical sciences, technology, education, and the arts. In the future, economically vital cities will be those that can adapt their economies, physical structures, and cultures to become part of a new international urban network of trade and investment" (Rondinelli et al. 1998, p. 85).

treatment, and global trade logistics support facilities.[10] Although they assert that the hyper-mobility of capital means that attempts to pick industrial winners will fail, they nevertheless recommend that states and localities "consider creating or expanding financial and tax incentives to attract domestic and foreign firms in core industries" (p. 91).

Clearly, like Felbinger and Robey, Rondinelli et al. are ambivalent about the value of targeted development investments: picking winners is bad but promoting the growth of core industries or clusters is good. The two sets of authors are not alone. This ambivalence is reflected heavily in the applied economic development literature, especially that concerned with competitive industry clusters, a body of writing that is strongly influencing local development policy making at present (Raines 2002; Martin and Sunley 2003). It is an ambivalence borne of substantial uncertainty about how to aid the emergence of new regional growth engines in an environment in which firms are increasingly footloose.

A somewhat different set of ideas are advocated in the Southern Growth Policies Board (SGPB) report. The SGPB report is particularly instructive because it represents one consensus view of globalization from the perspective of public officials rather than researchers. The report identifies nineteen recommendations (see Table 3), each endorsed by Southern governors. Ten of the strategies may be characterized as emphasizing marketing, information or advocacy (e.g., branding the region as globally-linked, explaining the benefits of trade to stakeholders, exposing regional populations to cultural diversity). Three strategies stress basic education; one, training and workforce development; and four, inter-governmental cooperation and re-engineering government. The most operational recommendation endorses trade promotion programs. In general, the report implies that the mostly trade-friendly Southern gubernatorial administrations cannot ignore growing anti-trade sentiment in a region hard-hit by the collapse of the textile, apparel, furniture and tobacco sectors (with Mexico and China as the perceived culprits). Thus its heavy emphasis on what a cynic might dismiss as public relations gimmicks rather than legitimate development interventions.

But the SGPB report should not be viewed as a purely public relations piece. It offers two important lessons. First, it is clear evidence of the emotionally charged nature of the globalization debate at the subnational level. Policy makers have obviously perceived a very pressing need to educate citizens and businesses about the opportunities international integration presents. Second, the report suggests that globalization may not imply much of a shift in the "what" aspect of regional economic development, as in "what kinds of strategies should we adopt?" One reason to advocate international education programs, global branding and trade promotion as the principal policy responses to recent globalization trends is that states are already doing many of the other things proposed by writers like Felbinger and Robey (2001) and Rondinelli, Johnson and Kasarda (1998). Indeed, strategies like technology development, entrepreneurship, incubators, and capital assistance to small businesses are actually "second wave" development initiatives

---

[10] An example of a new-era multi-modal transport facility is North Carolina's air cargo manufacturing project, the Global Transpark.

Table 7.3. SGPB recommended regional globalization policies.

Strategy	Type
Craft a vision of the globalized South (branding)	Marketing, Image
Cultivate an innovative, globally savvy business image for the South	Marketing, Image
Create a Southern initiative targeted to a developing world region	Marketing, Image
Improve the quality of foreigners' experiences in the South	Marketing, Tourism
Leverage existing connections to bolster recruitment efforts	Marketing, Attraction
Create an informed leadership and constituency for trade	Advocacy, Information
Celebrate local international connection success stories	Advocacy, Information
Internationalize local media reporting	Advocacy, Information
Develop outreach strategy to encourage civic support, volunteerism	Advocacy, Information
Encourage significantly more businesses to actively support trade	Advocacy, Information
Build comprehensive trade assistance program	Trade Assistance
Elevate state globalization policy through reorganization, accountability	Organizational
Forge multi-state partnerships to promote exports	Organizational
Fully leverage international expertise and links in higher education	Organizational
Build relations where there is a high return on investment	Organizational
Prepare every student to transition immediately to postsecondary school	Basic Education
Internationalize basic education	Basic Education
Expose youth to people from other countries	Basic Education
Embed international goals in state workforce development policies	Training

From Southern Global Strategies Council (2003), pp. 22-49).

that were adopted in many states in the 1980s (Bradshaw and Blakely 1999). While one could make the case that such strategies are all the more important in the face of globalization, and perhaps that they have not yet been adopted to a sufficient degree, they are not in themselves new ideas. In addition, the connection of some of the proposals to globalization often seems to be a very distant one. While fighting urban crime and poverty and reducing any unnecessary barriers to new business formation are surely important aims, it is hard to see how globalization *per se* has made them that much more imperative.

7.4 …and Research

What is missing in the debate about what globalization means for regional development policy is the *how* problem, as in "how do we get some of these things done?" The how problem relates to whether or not we have the right factual information and planning tools to design programs; to make multi-level allocation decisions that consider efficiency, distribution, and opportunity cost issues; and to conduct evaluations. Being proactive in economic development policy requires sound decision making regarding the relative distribution of scarce resources between general economic framework building on the one hand, and direct development interventions on the other. The former refers to changes in the rules, regulations, system of public finance and public services, and physical, technology, and education infrastructure; the latter to, first, programs or strategies to aid *adjustment* to restructuring and, second, initiatives to boost *growth* by fostering new regional economic engines. Additional difficult allocation and strategy design decisions must be made within those major categories of activity.

Taking just one example of the challenge facing policy makers, few are the regions that will not attempt to aid the emergence of new knowledge-based industries. Such is the holy grail of 21st century economic development. But how should it be done? Via a "shoot everything that flies" recruitment approach (Rubin 1988), a scrupulously hands-off scheme to foster entrepreneurial growth wherever emerges, a precision "pick winners" strategy that aims a host of initiatives at predicted growth sectors (Rondinelli et al. 1998), or a centers of excellence plan that emphasizes innovation as proposed by Felbinger and Robey (2001)? Would a single type of development instrument (e.g., tax inducements) be most effective, or many instruments in combination? Clearly it is a practical challenge of enormous dimensions with much room for wasted resources and counterproductive intervention. It is also a problem for which the globalization trends have significant bearing. The following sections discuss ways in which globalization trends are contributing to a widening gap between the *what* and the *how* in regional economic development planning and policy. Rather than attempting to be exhaustive, I suggest several particularly important areas of research that would help to bridge the gap.

Location incentives and subnational tax policy. One of the most active areas of economic development intervention at the subnational scale in the U.S. is the provision of tax and non-tax industrial location incentives (Bartik 2003). Paradoxically, despite the view predominant in the academic literature that incentives have little influence on company location behavior or hiring decisions, their use continues unabated. Indeed, a loosening of the geographic bounds on capital concomitant with globalization actually may be encouraging cities and states to increase their use of various inducements in an attempt to capture more activity from what is almost certainly a growing pool of footloose firms (Clark and Montjoy 2001). Both the Organization for Economic Co-operation and Development (OECD) and the European Union (OECD 1998) have implemented programs to examine the question of whether interstate tax competition is increasingly harmful in an integrated world economy.[11] While not strictly focused on incentives, their concern is that globalization is limiting governments' ability to set their tax regimes independently. The notion is that it is getting harder to tax mobile capital and labor. There is potential that tax competition will intensify significantly, leading to wors-

---

[11] According to the OECD (1998, p. 13): "…the removal of non-tax barriers to international commerce and investment and the resulting integration of national economies have greatly increased the potential impact that domestic tax policies can have on other economies."

ening fiscal strain and excessive taxation on immobile factors (Neumann et al. 2003). In general, the thrust of international efforts has been to encourage the scaling back of preferential tax treatments, thereby permitting tax base broadening and rate reductions. The goal is to minimize distortions on multi-location economic activity and limit negative fiscal spillover effects between jurisdictions.

The question of the positives and negatives of interjurisdictional tax competition has received much attention from researchers, particularly in the field of theoretical welfare economics (Wilson 1999). However, continuing government budget pressures and concerns about maintaining a positive business climate in the face of growing international competition has many states rethinking their tax regimes. The fact is that U.S. state tax codes are riddled with results of decades of favored treatment for selected sectors, attempts to copy the incentives of neighboring states, and a myriad of well-meaning but complex provisions to encourage investments in new technologies, equipment, and research and development. Add to that a growing number of non-tax inducements, from direct cash grants to provision of subsidized training, and interjurisdictional competition for development in the U.S. is probably as intense as ever.

As researchers, it is worth asking why incentives remain so popular if they are ineffective. Part of the answer is surely that many policy makers fail to grasp all of the welfare economics issues involved or are not familiar with the findings in the empirical literature. Also a factor is the political value of the "symbolic" capture of economic activity that would have occurred anyway.[12] But it is also true that research that would aid practical incentives reform is either scarce or uncompelling. Attempts to encourage states to undertake reform cannot simply address whether incentives produce faster growth or truly induce new economic activity, the subject of the bulk of academic research to date. For states and regions to make changes to existing policies, they need to know what incremental adjustments to existing programs make the most sense. The wholesale elimination of incentives is politically infeasible in most cases, and may not even be desirable (Bartik 2004).[13] Three major kinds of information would significantly inform incentives policy.

First, there is a need for consistent and reliable information on the level of financial benefits that given tax and non-tax incentives provide to businesses, acknowledging that many firms are increasingly mobile and multi-locational. Financial benefit is the most basic pre-condition for incentives efficacy, not to mention the evaluation of the welfare implications of a given incentive and the prediction of the fiscal impact of a policy change. Figuring out the relative value of various inducements is no trivial exercise, as work by Fisher and Peters' (1998) convincingly demonstrates. Too much of the existing research uses an effective tax rate approach that is convenient analytically but is of little value to those charged with making adjustments to inducement programs and legislation. Hypo-

---

[12] The political "rationality" of incentives in the face of evidence of their ineffectiveness has received too little attention in the literature (Wolman 1988).

[13] The complete harmonization of tax systems is not contemplated by either the OECD or European Union.

thetical firm models, by contrast, have the potential to significantly inform incentives design, particularly by enabling comparisons across states and regions. They can also lay bare how multi-locational and multinational enterprises are able to exploit the interaction of state tax systems, increasing free rider behavior. Partial interjurisdictional comparisons of incentives and tax burdens almost certainly generate misleading pictures of the benefits enjoyed by firms in various locations.

Second, undoubtedly many incentives are poorly targeted, since states have traditionally focused on encouraging manufacturing activity over other industries.[14] There has also been a bias toward assisting larger companies. While development policy makers are often aware that most new growth will derive from non-manufacturing and technology-based industries, particularly from sectors like advanced business services and health, a legacy of incentives to manufacturing limits states' fiscal capability to assist other industries or smaller businesses. Any serious attempt to improve incentives targeting will necessitate the scaling back of favored treatment for large manufacturing interests, a politically difficult option, especially as manufacturers face stiffening competition from low-cost producers overseas. Without clear and compelling evidence of the net fiscal and economic impacts of redirecting incentives in various ways, few states will have the political will to do anything other than expand inducements when fiscal conditions allow. The default solution to the targeting problem is to pass out more benefits when times are good, further exacerbating the state's fiscal position when times are bad.

Third, a key issue limiting incentives reform is the prisoner's dilemma problem. For obvious reasons, states are loath to unilaterally scale back their incentives programs, even to lower overall tax rates. Research and models that provide estimates of the interregional, interjurisdictional impact of alternative single- and multi-region incentives reforms are therefore essential. Such work is particularly important in light of growing regional agglomerations that span jurisdictional borders as well as evidence that regional economies within nations are becoming increasingly integrated (Hewings et al. 1998; Munroe and Hewings 1999).

Operationalizing workforce transition. One of the biggest challenges that increased international competition is introducing for subnational governments is the need to aid the job transition of workers displaced from declining industries. To build systems capable of doing this on a continuous basis, states and regions are searching for better ways to join analysis of current and projected structural economic change to worker training systems. Intermediaries in this process in the U.S. are local workforce development boards, training providers (e.g., community colleges, private sector training institutes, and to a more limited degree, universi-

---

[14] It also the case that some incentives on the books are simply poorly designed. In 2003, for example, North Carolina found that its existing research and development tax credit had the inadvertent effect of rewarding companies for increasing their R&D activity outside the state. The problem lay in the measurement of research activity, the complexity of the interaction between the state and federal tax codes, and the accounting strategies of taxpayers. The case serves to illustrate how difficult it is to properly implement a tax incentive; it is surely not the only example of a perverse outcome to a well-intentioned policy.

ties), and economic development agencies designing programs to assist the emergence of new sectors or the restructuring of old ones. More rapid structural economic change necessitates a more nimble training system that can adjust relatively quickly to the shifting workforce demands of industry, thereby ensuring a ready pool of skilled workers. Such would be the operational core of the more flexible workforce called for by Rondinelli et al. (1998) and Felbinger and Robey (2001). There is a practical need for better methodologies for continually evaluating and adjusting training curricula to anticipated training demands. Part of the solution is to better mesh economic and workforce development strategies so that efforts to proactively create new job opportunities jive with efforts to supply the workers to fill those opportunities (see Fitzgerald 1998).

Industry occupational employment matrices have long been used to produce occupational projections from industry projections. Such occupational projections have traditionally been central to workforce development planning in the U.S. However, effective training curriculum planning requires extensive information on occupational characteristics, especially the mobility of individuals between occupations (Feser, 2003), as well as emerging skill needs in industry. The development of improved workforce planning tools and data series in the U.S., including the online Occupational Information Network (ONET) and the U.S. Census Bureau's new Local Employment Dynamics Series, offer opportunities to better understand worker transition opportunities between occupations and industries based on workers' basic skills, levels of knowledge and experience, and training. Unfortunately, tools to bring such information to bear effectively in curriculum planning remain scarce.

Location, production fragmentation, and regional cluster opportunities. Presently, conventional wisdom in the regional economic development literature holds that the U.S. and other industrialized economies will capture more knowledge-intensive, higher technology activities while less developed countries will specialize in labor-intensive commodity production and low services (e.g., customer service call centers).[15] But while that may be true in broad aggregate, it may not hold for selected regions and/or clusters of industries within advanced economies, calling into question technology-based economic development as an economic panacea. Globalization trends are significantly altering the set of locational determinants in specific sectors, such that widely-accepted assumptions about the relative footloose nature of "low value," "low skill" activity, versus "high skill" or "technology-intensive" activity may be less and less meaningful. Regulatory issues, market access, and transportation and logistics issues—singly

---

[15] Gresser and West (2001, p. 8) argue that trends in U.S. trade show a clear "shift toward a high-tech, high-skill economy," which raises two basic policy questions: "on one hand, how to expand the winner's circle and give Americans the education, training and other tools necessary to compete in such an economy—and on the other, how to best develop the infrastructure of rules and agreements that will fit in a world economy in which America's greatest opportunities and advantages lie in newly emerging industries." The problem may be not be so simple at the subnational scale, where the pre-conditions for technology-based development vary widely.

or in combination—mean that certain low-wage, low value activities will remain viable in the U.S. Likewise, it is clear that at least some high technology activity can be carried out effectively offshore in environments of modest workforce skill and infrastructure development. These insights are relatively well-established in general terms, but empirical work at the detailed industry level is lacking, limiting their application at the policy level where day-to-day investment and allocation decisions have to be made.

Consider the example of the South, the U.S. region hardest hit by the relaxation of import controls on textiles and apparel with China's entry into the World Trade Organization. U.S. textiles and apparel production is highly concentrated in the South, particularly in the Carolinas and Georgia. One scenario for the South is that its textiles and apparel industries will evolve naturally—or perhaps with a little push—toward the skill-intensive end of what Hanson (1996) calls regional production networks. In regional production network theory, localization economies in skill-intensive activities of established clusters bid wages upward, increasing the incentive for the out-sourcing of low-skill assembly processes to low cost locations. One result is that developed countries with established clusters will gradually adjust to supply design-oriented, R&D, and marketing services while the standardized production of commodity goods will be carried out in developing countries. That might imply that Southern development officials should seek to nurture knowledge-intensive segments of the textiles and apparel sectors as commodity production migrates overseas. The problem for the South, however, is that it has never been a center for higher-value activity in textiles and apparel. Indeed, the South has traditionally specialized in the low-technology segments of most manufacturing industry value chains (Feser and Bergman 2000). Therefore, there may be very little localization economy or "cluster advantage" to one of the largest spatial clusters of textiles and apparel activity in the U.S. Indeed, places like New York and Los Angeles, where clothing design and marketing activities are concentrated, are more likely to benefit from freer trade and cost-saving outsourcing than Southern states.

To take another example, Alabama, Georgia, and North Carolina were among many other states that competed hard in 2003 to attract the planned manufacturing facility for the new Boeing 7E7 Dreamliner (now renamed the 787). Each offered up substantial incentives packages. Key to the attraction efforts was the prospect that the facility would anchor a new technology-based aircraft industry cluster that could replace employment losses in the textiles, apparel, furniture, and tobacco manufacturing industries. In that context, it was hoped that the Boeing facility would be accompanied by a network of locally-based suppliers while also generating lateral linkages from additional aircraft plant locations in proximity of the site. However, in a single company example of vertical specialization, Boeing's plans included an extensive global sourcing strategy that included purchase of wings, fuselage and engine components from overseas partners. It was possible that the 7E7 facility would be little more than a satellite assembly site with little follow-on development, a good project to be sure, but one that might not spur enough additional activity to justify record incentives packages.

Both the Southern textiles/apparel and Boeing recruiting cases are examples of how more research on what Taylor (2001, p. 3) calls the microeconomics of globalization—how "agents directly linked with global markets connect others with the global economy, through their transactions with local and regional markets"—can significantly inform local development planning. In the textiles/apparel case, it is a question of what can and should be salvaged from an industry cluster whose historic competitive advantage is rapidly eroding. In the case of Boeing, it is a matter of what local and regional linkages can be expected to emerge from a multinational firm. Work that is especially promising for uncovering changing location dynamics includes empirical research on vertical production networks and the location of multinational firms (e.g., McCann et al. 2002; Hanson et al. 2003); studies of the relationships between input sourcing, transportation factors and logistics, and location (McCann 1993, 1995; McCann and Fingleton 1996; McCann, 1998); and analysis of whether the international fragmentation production will lead to the regional agglomeration or de-glomeration of economic activity (Jones and Kierzkowski 2004b, 2004a).

Interregional trade modeling and structural analysis. Globalization trends are exposing the high costs of regions' rudimentary approaches for understanding their economies. Regional development agencies still rely heavily on single-region models and economic analysis techniques for impact prediction, forecasting, and studies of structural economic change. This is somewhat paradoxical given the intense level of policy interest in the last decade in industry clusters, a concept which explicitly prioritizes consideration of industrial and interregional interdependence. Surprisingly, much industry cluster analysis remains little more than the application of simple indicators like location quotients to sector-based data; it is often cluster analysis only in name and not technique (Bergman and Feser 1999). The paucity of interregional trade flows data is obviously partly to blame for the dominance of single-region models (Polenske and Hewings 2004), but it is also the case that many development agencies have failed to develop the internal capacity to promote, as well as utilize, more sophisticated approaches (more about this below). That is a shame because the growing extent to which subnational economies are linked to one another and the global economy—with all that implies for development strategy formulation—will never be fully appreciated by policy makers until it can be demonstrated empirically for their own regions.

The promise of interregional models for understanding regional structural economic change and informing practical policy alternatives is demonstrated in convincing fashion in work by the University of Illinois' Regional Economics Applications Laboratory in cooperation with the Federal Reserve Bank of Chicago (Hewings et al. 1998; Hewings et al. 1998; Okuyama et al. 2002; Testa 2002). A range of studies using models constructed with input-output and interregional commodity flows data have shown just how significant trade is to the Chicago, Illinois, and Midwestern economies, and that the vast majority of the flows are domestic rather than international. It is not the case that all the methodological and data challenges of interregional modeling have been solved. As Polenske and

Hewings note (2004, p. 271-2): “For questions about dramatically different structures and locations of production, such as are now occurring in the United States and elsewhere, none of the models developed so far will be able to provide very accurate estimates.” The point is that more progress is needed, and that will only come with continued research and application. It is hard to imagine how, in the face of increased international and interregional integration, we can hope to formulate effective economic development strategies in the absence of interregional analytical techniques. Writing off such methods and models as overly complex or esoteric is precisely the wrong approach. If we did that thirty years ago, we would not even have the single-region planning and analysis models that predominate today.

7.5 Summary

States and regions are making daily investments and spending decisions whose long-range impacts are uncertain (even becoming more uncertain, perhaps) and depend heavily on questionable assumptions about global-regional economic dynamics. I have argued in this paper that those assumptions are not necessarily outmoded at the broad policy-level—as when a state chooses to support technology-based economic development—but rather at the *operational* level, as when programs are designed, sectors are targeted, resources are allocated, and estimates of the impact of a given investment, project or program are calculated and evaluated against estimates derived for other potential interventions. Likewise they are outmoded when decisions about general government spending are made on the basis of revenue projections that fail to reflect regional economic vulnerabilities to international price, supply, and demand shocks. To often it seems we do not know *how* to do *what* we often already know we ought to do, a problem globalization is only making worse. We probably understand better than ever before that regions require a flexible workforce that can adjust more quickly to structural economic change, but figuring out how best to nurture such a workforce through carefully designed interventions remains a major hurdle. Likewise, however much we agree that fostering new competitive industry clusters is a good idea, it is increasingly difficult to make sound judgments about where to target investments in business and technology development, given increasingly mobile capital and labor.

All economic development programs are instituted at the expense of other government programs and therefore the ability to explicitly model trade-offs between investments is essential, especially if the political climate leans toward traditional development interventions that are counter-productive in a global economic environment. In a scarce resource environment, general discussions of global-friendly strategies, without specific elucidation of how to put those strategies into practice and to measure gains and losses, provide only modest substantive policy guidance. They therefore represent little valuable currency in the political marketplace where budgetary and programmatic decisions are made.

Clearly planning effectively in a globalizing economy demands better empirical information (facts) and tools (models, frameworks). But the challenge is not the research community's alone. The time is long overdue for statistical agencies to rethink what information is required to fully understand global economic trends at the regional scale. The absence of reasonably detailed and more timely interregional trade data is a major problem. And regional development agencies must get serious about building the internal capacity to think and act strategically. Most economic development organizations are ill-equipped to effectively utilize the latest knowledge, models and tools because they continue to under-invest in their own applied analytical capability (Feser 2005). Not only is regional economic policy making too reactive as a result, but development officials are poor advocates for needed investments in data collection, modeling and empirical research. This creates a widening and self-reinforcing gulf between researchers and research users, where the former develop increasingly sophisticated models with less data and fewer real-world opportunities to apply them and the latter rely on simplistic tools whose validity is eroding rapidly. Adjusting our regional economic development strategies to the new world economic order will therefore require more robust partnerships between those who would develop new tools and generate facts, those who collect the information necessary to properly understand trends and drive planning tools, and those who would put planning tools and facts to good use.

References

Balassa B (1967) Trade Liberalization Among Industrialized Countries. McGraw-Hill, New York

Bartik TJ (2003) Local economic development policies. Kalamazoo, MI, Upjohn Institute for Employment Research, pp 03-91

Bartik,TJ (2004) Incentive solutions. Kalamazoo, MI, Upjohn Institute for Employment Research, pp 04-99

Bergman EM, Feser EJ (1999) Industrial and Regional Clusters: Concepts and Comparative Applications. Regional Research Institute, West Virginia University, Morgantown, WV

Bhagwati J (2004) In Defense of Globalization. Oxford University Press, New York

Bolton R (1992) 'Place prosperity vs people prosperity' revisited: An old issue with a new angle. Urban Studies 29:185-203

Bordo MD, Eichengreen B, Irwin DA (1999) Is globalization today really different than globalization a hundred years ago? NBER working paper series, Cambridge, MA. 7195.

Bradford SC, Grieco PLE, Hufbauer GC (2005) The payoff to America from global integration. In: Bergsten CF (ed) The United States and the world economy: Foreign economic policy for the next decade. Institute for International Economics, Washington, DC, pp 65-109

Bradshaw TK, Blakely EJ (1999) What are the "third-wave" state economic development efforts? From incentives to industrial policy. Economic Development Quarterly 13:29-44

Clark C, Montjoy RS (2001) Globalization's impact on state-local economic development policy: Introduction to the symposium. Policy Studies Review 18:6-12

Conway C (2003) The impact of globalization on the southern United States. Speech to the Council of Local Authorities, Tokyo, Japan, 7 October 2003

Cooper RN (1995) Comments and discussion on "Growing world trade: causes and consequences." Brookings Papers on Economic Activity 1995:363-68

Council of Economic Advisers (2004) Economic report of the president. US GPO, Washington, DC

Enright MJ (2000) The globalization of competition and the localization of competitive advantage: Policies toward regional clustering. In: Hood N, Young S (eds) The Globalization of Multinational Enterprise Activity and Economic Development. Macmillan, London, pp 303-331

Farrell D (2004) Beyond offshoring: Assess your company's global potential. Harvard Business Review, December 82-90

Feenstra RC (1998) Integration of trade and disintegration of production in the global economy. Journal of Economic Perspectives 12:31-50

Felbinger CL, Robey JE (2001) Globalization's impact on state and local policy: The rise of regional cluster-based economic development strategies. Policy Studies Review (Review of Policy Research) 18:64-79

Feser E (2003) What regions do rather than make: A proposed set of knowledge-based occupation clusters. Urban Studies 40:1937-58

Feser E (2005) Engaging the university in economic development planning and policy. Department of Urban and Regional Planning, University of Illinois at Urbana-Champaign, Champaign, IL

Feser EJ, Bergman EM (2000) National industry cluster templates: A framework for applied regional cluster analysis. Regional Studies 34:1-19

Findlay R (1978) An Austrian model of international trade and interest rate equalization. Journal of Political Economy,86:989-1008

Fisher PS, Peters AH (1998) Industrial Incentives: Competition Among American States and Cities. Upjohn Institute, Kalamazoo, MI

Fitzgerald J (1998) Moving the workforce intermediary agenda forward, Economic Development Quarterly 18:3-9

Fujita M, Hu D (2001) Regional disparity in China 1985-1994: the effects of globalization and economic liberalization. Annals of Regional Science 35:3-37

Glasmeier AK, Conroy ME (1994) Global Squeeze on Rural America: Opportunities, Threats, and Challenges from NAFTA, GATT, and Processes of Globalization. The Institute for Policy Research and Evaluation, The University of Pennsylvania, Philadelphia, PA

Glickman NJ, Woodward DP (1988) The location of foreign direct investment in the United States: Patterns and determinants. International Regional Science Review 11:137-54.

Gresser E, West S (2001) Race to the top II: The facts about American trade and a debate that misses the point. Progressive Policy Institute, Washington, DC

Hanson GH (1996) Localization economies, vertical organization, and trade. American Economic Review 86:1266-78

Hanson GH (1998) Regional adjustment to trade liberalization. Regional Science and Urban Economics 28:419-444

Hanson GH, Mataloni RJ, Slaughter MJ (2003) Vertical production networks in multinational firms. National Bureau of Economic Research. Working Paper 9723, Cambridge, MA

Hanson GH, Slaughter MJ (2003) The role of multinational corporations in international business cycle transmission: Skew lines or arbitrage opportunities? Paper prepared for the Kiel Week Conference 2003. Kiel, Germany.

Head K, Ries J, Swenson D (1995) Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. Journal of International Economics 38:223-247

Helmsing AHJ (2001) Externalities, learning and governance: New perspectives on local economic development. Development and Change 32: 277-308

Hewings GJD, Schindler GR, Israilevich PR, Sonis M (1998) Agglomeration, clustering and structural change: Interpreting changes in the Chicago regional economy. In: Steiner M (ed) Clusters and regional specialisation. Pion, London

Hewings GJD, Sonis M, Guo J, Israilevich PR, Schindler GR (1998) The hollowing out process in the Chicago economy, 1975-2015. Geographical Analysis 30:217-233

Hu D (2002) Trade, rural-urban migration, and regional income disparity in developing countries: a spatial general equilibrium model inspired by the case of China. Regional Science and Urban Economics 32:311-338

Hummels D, Ishii J, Yi K-M (2001) The nature and growth of vertical specialization in world trade. Journal of International Economics 54:75-96

Hummels D, Rapoport D, Yi K-M (1998) Vertical specialization and the changing nature of world trade. FRBNY Economic Policy Review 79-99

Irwin DA (2002) Free trade under fire. Princeton University Press, Princeton, NJ

Ivarsson I, (1999) Competitive industry clusters and inward TNC investments: The case of Sweden. Regional Studies 33:37-49

Jones RW, Kierzkowski H (2004a) International framework and the new economic geography. North American Journal of Economics and Finance, Forthcoming.

Jones RW, Kierzkowski H (2004b) International trade and agglomeration: An alternative framework. Journal of Economics, Forthcoming.

Kletzer L (2002) Imports, exports, and jobs: What does trade mean for employment and job loss? W.E. Upjohn Institute for Employment Research, Kalamazoo, MI

Krugman P (1995) Growing world trade: Causes and consequences. Brookings Papers on Economic Activity 1995:327-377

Krugman P (1996) Pop Internationalism. MIT Press, Cambridge, MA

Malizia EE (1994) A redefinition of economic development. Economic Development Review 83-84.

Mansori KS (2003) The geographic effects of trade liberalization with increasing returns in transportation. Journal of Regional Science 43:249-268

Martin R, Sunley P (2003) Deconstructing clusters: Chaotic concept or policy panacea? Journal of Economic Geography 3:5-35

McCann P (1993) The logistics-cost location-production problem. Journal of Regional Science 33:503-516

McCann P (1995) Rethinking the economics of location and agglomeration. Urban Studies 32:563-577

McCann P (1998) The economics of industrial location: A logistics-costs approach. Springer, Heidelberg

McCann P, Arita T, Gordan IR (2002) Industrial clusters, transactions costs and the institutional determinants of MNE location behavior. International Business Review 11:647-663

McCann P, Fingleton B (1996) The regional agglomeration impact of just-in-time input linkages: evidence from the Scottish electronics industry. Scottish Journal of Political Economy 43:493-518

McCann P, Mudambi R (2004) The location behavior of the multinational enterprise: Some analytical issues. Growth and Change 35:491-524

McKinsey Global Institute (2003) Offshoring: Is it a win-win game? McKinsey Global Institute, San Francisco

Meardon SJ (2001) Modeling agglomeration and dispersion in city and country. American Journal of Economics and Sociology 60:25-57

Munroe DK, Hewings GJD (1999) The role of intraindustry trade in interregional trade in the Midwest of the U.S., REAL 99-T-7. Regional Economics Applications Laboratory, University of Illinois, Champaign, IL

Neumann R, Holman J, Alm J (2003) Globalization and tax policy. Department of Economics, University of Wisconsin, Milwaukee, WI

OECD (1998) Harmful tax competition: An emerging global issue. Organization for Economic Co-operation and Development, Paris

Okuyama Y, Hewings GJD, Sonis M, Israilevich P (2002) Structural changes in the Chicago economy: A field of influence analysis. In: Hewings GJD, Sonis M, Boyce D (eds) Trade, networks and hierarchies. Springer, Heidelberg, pp 201-224

Oneal M (2004) Global economy strains loyalty in company town. Chicago Tribune. Chicago, IL.

Perera-Tallo F (2003) Growth due to globalization. International Economic Review 44:651-676

Pernia EM, Quising PF (2003) Trade openness and regional development in a developing country. Annals of Regional Science 37:391-406

Polenske KR, Hewings GJD (2004) Trade and spatial economic interdependence. Papers in Regional Science 83:269-289

Raines P (2002) The challenge of evaluating cluster behaviour in economic development policy. European Policies Research Center, University of Strathclyde 26, Glasgow, UK

Rondinelli DA, Johnson JH, Kasarda JD (1998) The changing forces of urban economic development: Globalization and city competitiveness in the 21st century. Cityscape 3:71-105

Rosen H (2002) Reforming trade adjustment assistance: Keeping a 40 year promise. Institute for International Economics Conference on Trade Policy in 2002, Washington, DC

Rostow W (1960) The stages of growth: A non-communist manifesto. Cambridge University Press, Cambridge, UK

Rubin HJ (1988) Shoot anything that flies; claim anything that falls; conversations with economic development practitioners. Economic Development Quarterly 2:236-251

Samuelson PA (2004) Where Ricardo and Mill rebut and confirm arguments of mainstream economists supporting globalization. Journal of Economic Perspectives 18:135-146

Sánchez-Reaza J, Rodríguez-Pose A (2002) The impact of trade liberalization on regional disparities in Mexico. Growth and Change 33:72-90

Sanyal KK (1983) Vertical specialization in a Ricardian model with a continuum of stages of production. Economica 50:71-78

Sen A (1999) Development as freedom. Alfred A Knopf, New York

Simmie J, Sennett J (1999) Innovative clusters: Global or local linkages? National Institute Economic Review, 87-98

Southern Growth Strategies Council (2003) Fast forward: Mobilizing the south for prosperity in a global economy. Southern Growth Policies Board, Reseach Triangle Park, NC

Stiglitz JE (2002) Globalization and its discontents. W.W. Norton, New York

Stroomer C, Giles DEA (2003) Income convergence and trade openness: fuzzy clustering and time series evidence. University of Victoria econometrics working paper EWP0304: 29, Victoria Canada

Stumberg R, Schweke W (1999) Could economic development become illegal in the new global policy environment? Corporation for Enterprise Development, Washington, DC

Taylor JE (2001). Microeconomics of globalization: evidence from Mexico, China, El Salvador, and the Galapagos Islands. Center on Rural Economies of the Americas and Pacific Rim, University of California at Davis, Davis, CA

Testa WA (2002) Using the REAL models to understand Midwest restructuring. In: Hewings GJD, Sonis M, Boyce B (eds) Trade, Networks and Hierarchies. Springer, Heidelberg, pp 279-288

Venables AJ (1998) The assessment: Trade and location. Oxford Review of Economic Policy 14:1-6.

Wilson JD (1999) Theories of tax competition. National Tax Journal 52:269-304

Wolman, Harold (1988) Local economic development policy: What explains the divergence between policy analysis and political behavior? Journal of Urban Affairs 10:19-28

Yi K-M (2003) Can vertical specialization explain the growth of world trade? Journal of Political Economy, 111, 52-102

Part B: Methodological Advances—Models of Networks

8 Globalization and Intermodal Transportation: Modeling Terminal Locations Using a Three-Spatial Scales Framework

Jan H.R. van Duin and Gijsbertus P. van Wee

Delft University of Technology, Delft, the Netherlands

8.1 Introduction

During the last decade the influences of globalization have shaped our countries. The meaning of the word globalization has changed from a ‘catch all’ term used in economics to a much broader meaning of describing the impact of increased international integration across a range of fields – from literature to sociology, to technology (Ruane and Sutherland 2002). In the early 1990s some companies in Europe began to integrate operations across national boundaries. The removal of trade barriers increasingly enabled geographical integration and thus operations spanning country borders in creating economies of scale (Bowersox et al. 1996). In order to sustain economic growth, many companies felt the need to develop new markets outside their home countries. New manufacturing plants and distribution centers were established in Asian countries. Not only the market conditions but also manufacturing conditions seemed to be favorable. Low labor wages, good employment and expansion opportunities were the main motives for these companies to invest. As a result relocation and re-allocation of existing manufacturing activities were set-up (Vos 1993). This global shift of manufacturing activities caused a tremendous pressure to change the distribution structures. To guarantee the delivery demands, the organization of distribution channels became of importance. While maintaining the speed of delivery, the price of delivery, and the flexibility of demand, the distribution channels needed to adjust the logistic activities in their structure. Activities of transshipment became important value-adding links to the distribution channel (Porter 1996). As an observation of this tendency, it can be noticed that around 700 European Distribution Centers have been set up in the Netherlands over the last 20 years, almost one third of all the new distribution centers in Europe (Ernst and Young et al. 2002).

The high demands for reliability in distribution structures of companies ask for the presence of well-developed infrastructures. Loyola de Palacio, Vice-president of the European Commission, responsible for energy and transport, states in the foreword of the EU white paper (2001): ‘Transport is crucial for our economic competitiveness and commercial, economic and cultural exchanges. This sector of the economy accounts for some 1000 billion, or over 10% of the EU’s gross domestic product, and employs 10 million people. However, the warning signs are clear. Congestion, resulting in environmental nuisance and accidents, is getting worse day by day, and penalizing both users and the economy. If nothing is done, the cost of congestion will, on its own, account for 1% of the EU’s gross domestic product in 2010 while, paradoxically, the outermost regions remain poorly connected to the central markets. Europe must bring about a real change in the Common Transport Policy. The time has come to set new objectives for it: restoring the balance between modes of transport and developing intermodality, combating congestion and putting safety and the quality of services at the heart of our efforts, while maintaining the right to mobility.’

As for the development of intermodality, a new approach has been developed to identify attractive locations of terminals. In this approach the perceptions of the actors involved at the different levels of policymaking play an important role in the choice of a new terminal. The way actors think and make their decisions is important to understand. Due to this way of thinking and reasoning, studies of decision making in networks have caused a fundamental change in thinking about the role of decision makers, stakeholders, and analysts (Fischer 1993). The image – illusion perhaps – of decision making as a (bounded) rational design effort by an elite group with implementing power is replaced by an emergent perspective on policymaking in which policy is seen as the result of interaction in a network of corporate actors (Marin 1991).

The traditional research on policymaking for transportation and logistics is mainly based on operational research techniques. Mulvey (1994) indicates that these techniques are poorly understood by the general public. A translation of the technical issues like mathematical assumptions and solving-methods into ‘plain English’ is usually not done. According to Nemhauser (1993): ‘Great strides have been made in the use of optimization. Models with thousands of variables and constraints are being solved regularly and the results are being applied routinely. Still we need to do a better job in making optimization easier to use and in making solutions meaningful to all types of stakeholders’.

As Vidal and Goetschalckx (1997) mentioned in their critical review on models, some research opportunities for developing more comprehensive modeling should focus on:

- general simulation of qualitative factors;
- environmental conditions, such as availability of infrastructure, determination of adequate local access capacities;
- modeling of alliances and multi-company network configurations, and
- development of specialized solution methods.

Of course, it is almost impossible to develop a general single model that integrates all these aspects. This conclusion leads to the development of an overall methodological transportation framework, supported by multiple interrelated models being capable of representing qualitative factors and uncertainties. The development of an approach to build these interrelation models must fill the gap between logistic decision processes on the one hand, and the specification of the design contents on the other hand. The new models must be easy to use and understand, with user-friendly capabilities, such as graphical representations of the systems under analysis. Sijbrands (1993) clarifies that the task to support strategic logistic issues seems to be simple; however, the way to process and support is the main research objective to be tackled in the future.

One of the key elements in intermodal transport is the location of the terminals. With this paper we want to demonstrate a first modeling approach to tackle some of the issues mentioned, in order to provide more adequate decision support for the selection of locations for new terminal initiatives.

The remaining part of this paper is organized as follows. In chapter 2 we will discuss the policies and actor perceptions on intermodal transportation at the different levels of policymaking. With insight into the main actor perceptions we have been able to select the main decision variables for each level of policymaking. Using these decision variables, we have made dedicated models to identify potential terminal locations. Chapter 3 includes the descriptions of the models and the results obtained from the models. Chapter 4 gives some conclusions.

8.2 Policies for Intermodal Transportation

In paragraph 2.1 a general introduction on intermodal transportation will be given. In order to get some insight into the perceptions of actors, the policies and actor attitudes for intermodal transportation are discussed on European, national (Dutch) and regional level in the paragraphs 2.2, 2.3, and 2.4, respectively. Paragraph 2.5 briefly describes the operational level. Paragraph 2.6 will give an overview of the main factors to be considered for all actors.

8.2.1 Intermodal Transportation

For the description of intermodal transportation is referred to the description of the European Conference of the Ministers of Transport: 'the movement of goods in one and the same loading unit or vehicle, which uses successively several modes of transport without handling the goods themselves in changing modes' (CEMT 1995). In our research we have only focused on containerized transport, other forms of intermodal transportation (of bulk-goods, fluids and/or gasses) are beyond the scope of this research. For a good understanding the general logistical practice for intermodal transport can be described as follows:

A carrier picks up an empty container from an empty depot by truck, stuffs the container at the shipper's location, and takes the container to the nearest terminal. The terminal operator receives this container, stacks the container temporarily and transships the container on the scheduled train or barge service. On fixed departure-times a train or barge departs for a long-distance trip to another terminal. Then the container is temporarily stacked and a carrier arranges the final delivery to the customer.

'Above all, intermodal transportation is not just the hardware or equipment involved with the freight movement, but the process, which becomes a major component of the systems approach to business' (Muller 1995). For destination to be reached by sea transport, one of the terminals is always located at a port. For the seas-containers this means that one extra container handling is not needed. The implication can be reflected in Table 1 with the break-even distances for intermodal transport in relation to road transport.

Table 8.1. General break-even distances (in kilometers) compared to road transport.

	Barge	Rail
Continental Containers	250	400
Sea Containers	100	200

Source: Ministry of Transport, 1994

In the global logistics supply chains of multinational companies, i.e. the demand-side of transportation, a growth of importance of reliable transportation can be observed due to the rise of more Just-In-Time oriented logistics, i.e. the value-adding functions being at the right time at the right place. From the supply-side of transportation the competition between the transport modes is still focused on price. Table 8.1 shows that road transport still seems to have a competitive market position, especially up to distances of 250-400 kilometers.

8.2.2 The European Level

The first policy level is oriented towards the actors playing a part in the European hinterland. The main borders of this area stretch out from the seaports Le Havre, Zeebrugge, Antwerp, Rotterdam, Bremen, and Hamburg to several European hinterland terminals, such as Milan, Metz, Munich etc. The main actors involved are European Ministers of Transport, shippers, road carriers, intermodal agents and their carriers. Many shippers have changed their distribution structures as a result of the removal of trade barriers. Due to the cost reducing opportunities many stock-holding units have, in their distribution structure, been eliminated or outsourced. In order to maintain the same reliability Just-In-Time deliveries have become more important. This has led to growth in transport, mainly conceived by road carriers. Therefore, this growth has been inevitable and has, in interaction with the increasing passenger car traffic, caused congestion on important highways (Bovy 2001). The congestion negatively influences the total transport time for delivery, the reliability of the delivery and, eventually, the costs. This provides opportunities for intermodal transportation. Also the trend of spatial concentration of inventories increases possibilities to cluster (combine) transport, i.e. the necessary initial condition for intermodal transportation. That's the reason why the shippers are not unwilling towards new intermodal transport initiatives. Apart from this reason, some shippers strive for a 'green' company image by choosing environment-friendly transport like intermodal transport. There are still some obstacles in the eyes of the shippers which have to be solved. Their main concern is the current price of intermodal transport. A large survey showed that the 'cost-oriented' group of actors uses intermodal transport most intensively. With more than 50% of the sample (92 companies) it shows the largest actual demand for intermodal transport (Spin 2002).

A point of concern is the number of intermodal destinations in the European hinterland. The total number of terminals connected by water and/or rail infrastructure is still too limited to allow intermodal transport to compete with road transport on a large scale, due to the high level of flexibility and detailed networks of road transport. The European Ministers of Transport try to draw up plans for the development of Transport European Networks (TEN-T, 2004). These plans are focused on the integration of transport modes and define the transport networks in hubs and spokes. At hubs fast transshipment should be facilitated and subsequently transported by the spokes. The integral management of the transport at a hub consists of physical transport infrastructure, traffic-controlling systems, positioning systems and navigation systems. The European Ministers of Transport try to coordinate and stimulate the infrastructural developments of each country in the direction of the plans for TENs. The Figures 8.1 and 8.2 represent the TEN for the Netherlands, i.e. the railways and waterways, respectively.

They do not have an explicit disposal for steering instruments on national governments, but they can stimulate some developments by providing subsidies. The PACT-program was a first initiative for promoting intermodal services and now the Marco Polo program has been set up as a start-up aid for new non-road transport services (EC-DG/ET, 2003). The intermodal agents and carriers have already established intermodal transport services for long-distance haulage. Thick maritime container flows are transported by these agents from harbors to far locations in the hinterland (for example from Rotterdam to Milan). So far, short-distance intermodal initiatives have not been undertaken since it seems almost impossible to collect thick container flows for short distances. Because of their strong competitiveness, price and flexibility, the road carriers still have the greatest market share of transport in Europe. The internal competition between road carriers is strong and therefore many carriers operate break-even or even less than break-even. As in other sectors mergers, takeovers and new alliances are daily news and should be interpreted as a matter of strategic management to survive in the long run. Some road carriers experience hindrance caused by some measures of national policies which aim at reducing truck traffic. For instance, in Switzerland and Austria truck traffic has strongly been reduced and a lot of trucks are placed on trains. Germany currently tries to solve the remaining technical problems of the announced MAUT-levy (a levy on infrastructure use by trucks).

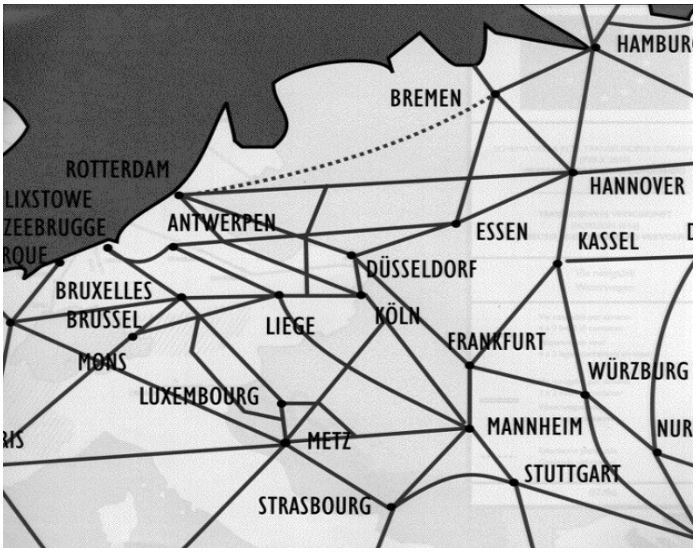


Fig. 8.1. Part of the Trans European network - Railways connected to Holland (horizon 2010)

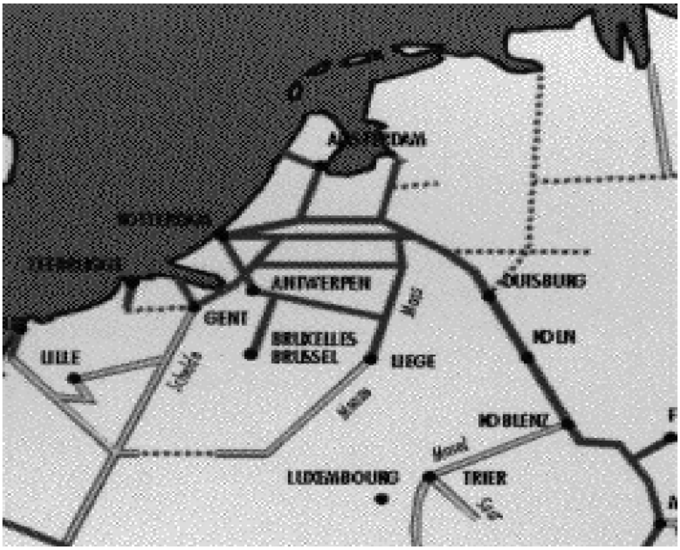


Fig.8.2. Part of the Trans European network - Waterways connected to Holland (horizon 2010)

8.2.3 The National Level (Dutch)

The national government is represented by respectively the Ministry of Transport and Public Works, the Ministry of Economic Affairs and, the Ministry of Environmental Affairs. All these Ministries show a strong compassion for the development of intermodal transport. The Ministry of Transport tries to stimulate intermodal transport initiatives in order to maintain accessibility to important economic centers. The water infrastructure, and to a lesser extent, the rail infrastructure still have enough capacity to accommodate extra traffic. The environmental performances of these modes appeal to the Ministry of Environmental Affairs. The Ministry of Economic Affairs attaches importance to generating economic value at the hub-terminals. At these terminals a lot of transshipment is carried out and these terminals attract companies providing good accessibility to all kinds of transport modes. While the Ministry of Economic Affairs is strongly focused on generating economic value, a governmental policy can be observed being extremely focused on terminal development within the frontiers of the Netherlands (Roermund *et. al.* 1995) (see Figures 8.3 and 8.4). With a minimal break-even distance for intermodal transport of at least 100 kilometers, it is not strange that the national policy is completely 'frontier-oriented' by trying to locate the (new) economic activities around terminals within the Netherlands. The so-called Trans European Inland Terminals (TEIT) and their connected infrastructure should function as the spokes in the European hub-spokes network, having intermodal services on these lines. Therefore we have seen upgrades of the terminals of Twente, Nijmegen, Venlo and Veendam. In the policy plans Valburg was mentioned as a promising new water and railway terminal. The areas within the vicinity of the terminals should develop as logistics service parks for value-adding services. That's the reason why the Ministry of Transport had provided SOIT-subsidies (Subsidies Of Inland Terminals) for seven new terminal initiatives in the regions of Brabant and Overijssel (Modal shift, 2004).

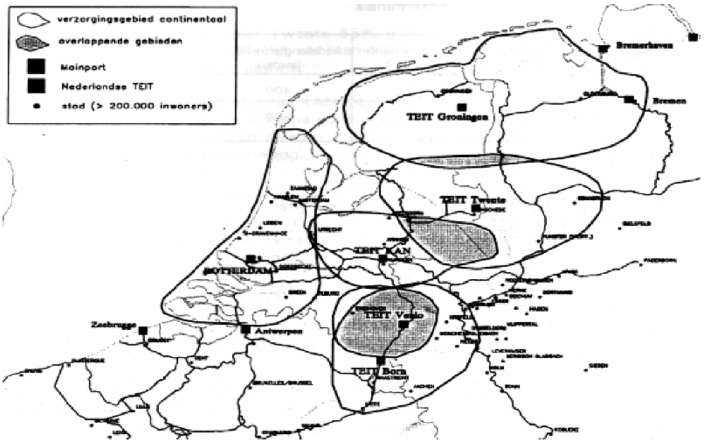


Fig. 8.3. Continental container service areas

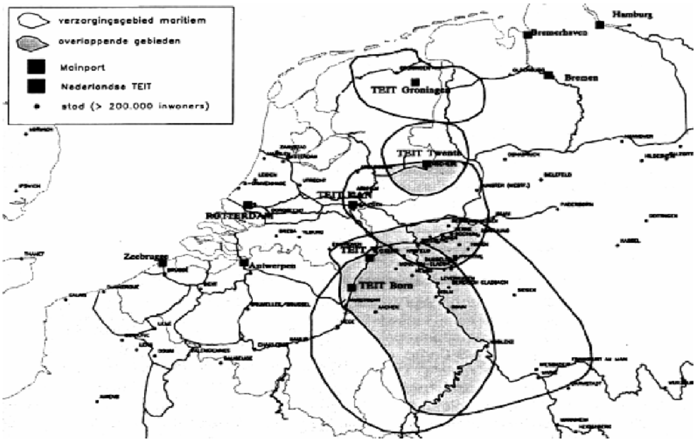


Fig. 8.4. Maritime container service areas[1]

[1] Source: Roermund et. al (1995), page 111

Two branch organizations also playing an important part at national level are the interest groups EVO and TLN. The EVO serves the interests of the shippers with self-organized transport in the Netherlands. Facing these interests, the EVO is quite similar to the attitude of the shippers at European level. The freedom of transport choice seems to be an important issue. TLN represents the carriers, the members of which consist for 90% of traditional road carriers and for 10% of intermodal carriers. This important union strives after fair competition on the transport market without governmental interference regarding any of the transport modes (apart from infrastructure provision). Within TLN, the road carriers want to maintain their current position as well as their total number of members. At this moment the road transport sector has a keen internal competition and many carriers even accept losses. The inland intermodal agents are to a limited extent represented in the TLN. These agents develop long-distance services in the Netherlands. Because the Netherlands is a small country, the number of these services is limited.

Table 8.2. Factors to be considered for each policy level.

Policy Level	Actors	Factors
European	Shippers	Price # connections of intermodal transport services
	Ministers of Transport	# connections of intermodal transport services
	Intermodal Agents	Price Long-distance intermodal transport
	Road carriers	Price Transport for all distances
National	Shippers (EVO)	Price # connections of intermodal transport services Freedom of transport choice
	Ministry of Transport	Hub-and spoke terminal network with frontier orientation Economic value at Dutch terminals
	Ministry of Environmental Affairs	Environment benefits of intermodal transport
	Ministry of Economic Affairs	Hub-and spoke terminal network with frontier orientation Economic value at Dutch terminals
	TLN (carriers and intermodal agents)	Long-distance services Maintain market positions Free market
	Local government	Terminal attraction by low ground taxes
Regional	Shippers	Price Frequency of the transport service Flexibility
	Intermodal Agents	Attraction of transport volumes Frequent schedules
	Local government	Terminal attraction by low ground taxes
Operational	Shippers	Transport information of transport progress
	Carriers	Transport on schedule
	Terminal Operator	Crane utilization Floor utilization

8.2.4 The Regional Level

At regional level local authorities would like to develop their cities in terms of economic growth by providing accessible industrial sites. The attraction of an intermodal terminal could be a serious alternative for the improvement of the accessibility. By providing subsidies or/and raising low ground taxes they try to attract companies to their areas. That's why some shippers are reconsidering locations, but the main motive for changing locations is the accessibility and their location towards their customers' positions. Shippers try to organize their transport with high frequencies (allowing inventory reductions) against low prices and demand a high flexibility towards the ordering times. Therefore, terminal agents have to attract large freight volumes, and transport them with high frequencies on a regular basis.

8.2.5 The Operational Level

At operational level the carriers do their utmost to follow the fixed transport schedule. The terminal operator wants to use its transshipment equipment and its floor space as well as circumstances allow. If the shipper agrees on price and transport schedule, he/she wants to be 'on-line' informed about the transportation progress.

8.2.6 Overview of Actors' Factors

Now all actors' perceptions have been identified for each level of policymaking, it is possible to sum up the main factors to be considered. The identification of these factors is important since they need to be incorporated in the models to be built. Table 8.2 gives an overview of the main factors to be considered.

8.3 Modeling Intermodal Networks

This section describes the modeling approach and presents the results of applying the approach to identify the locations of potential terminals in the Netherlands.

Figure 8.5 represents the modeling approach for the support of selecting good terminal locations at all levels of policymaking. The approach includes three steps:

1. Railway or water infrastructure permitting, the model decides on the basis of regional transportation demands whether it is more cost-efficient to service by intermodal transportation or by road transportation. The focus of the model is to determine the economic feasibility of the connections between national terminals and other European terminals. The formalization is based on a linear programming model minimizing the integral transportation costs. The description of the model and the modeling results are given in paragraph 8.3.1.

2. When the rough locations of terminals are identified in a region, a second model is applied to search for the exact location in the region. The preference locations are determined on the basis of their infrastructural accessi-

bility. A detailed cost model is developed to support this exact location decision. The model and the results are given in paragraph 8.3.2.

3. The last part of this approach is focused on the possible customers in the neighbourhood of the terminals. Analysis of suitability, timeliness, and deliverance reliability of goods are factors of logistic importance. These aspects are both for the terminal and its customers specified by a simulation model. The model is given in paragraph 8.3.3.

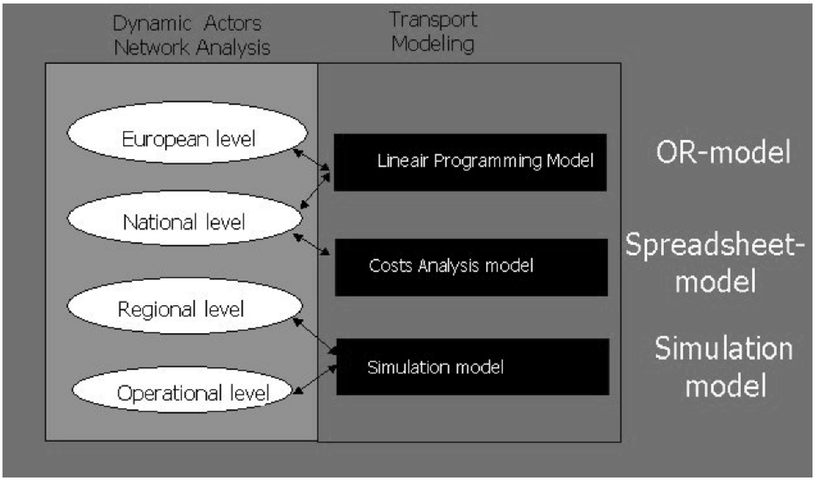


Fig. 8.5. The modeling approach for intermodal transportation.

8.3.1 The Linear Programming Model

Since price seems to be the most important performance indicator at both European and national level (see paragraph 2.5), we developed a linear programming model based on the main factor for selecting transport: 'price'. The model is applied for locating shuttle-services between terminals in Europe and based on the calculated volumes of the shuttles, whether a new terminal location will be attractive or not. A shuttle-service is a general logistic concept for intermodal transportation services. The service operates with a fixed (time-)schedule to fixed terminal locations elsewhere and vica versa. The continental container transport demands (NEA, 1992) between several European regions, with origin or destination Netherlands, are optimally assigned to the transport network. Actually, the model is a special case of location/allocation models (Koopmans & Beckman, 1957) (Daskin, 1995) (Drezner, 1995). The transport network contains direct road transport connections and intermodal transport connections between terminals. The optimization is restricted to one year. The objective function is defined as follows:

$$\text{Minimize}\quad \begin{aligned}&\sum_j\sum_k X_{jk}\,d_{jk}\,C_s(j,k)+\\&\sum_i\sum_j\sum_k\sum_l x_{ijkl}(d_{ij}\,C_p(i,j)+d_{kl}\,C_p(k,l)+C_h(j)+C_h(k))+\\&\sum_i\sum_l x_{il}\,d_{il}\,C_r(i,l)\end{aligned} \tag{8.1}$$

Before describing the model in more detail, we will first elaborate on the objective function which includes the following three cost elements (see Formula (8.1)):

1. Shuttle Costs per TEU[1] between terminals. The costs of a shuttle connection are determined (by the model) at the largest volume in one direction. The volume in the reverse direction should be geared to this volume as much as possible.
2. Costs per TEU for pick-up and delivery, and the transshipment costs per container at a terminal.
3. Costs per TEU for direct road transport between regions.

The linear search algorithm minimizes the objective function subject to the following constraints:

$$X_{jk} \geq \sum_i\sum_l x_{ijkl} \qquad \forall\, j,k \in T \tag{8.2}$$

$$X_{jk} \geq \sum_i\sum_l x_{ikjl} \qquad \forall\, k,j \in T \tag{8.3}$$

The capacity of an intermodal shuttle-service connection is determined by the volume of the largest transport flow in forward or backward direction. This restriction is necessary to model the return of empty containers.

$$x_{il} + \sum_j\sum_k x_{ijkl} \geq D_{il} \qquad \forall i,l \in R \tag{8.4}$$

The sum of transported containers between region i and region l must be met for the specific demands.

x_{il}:	volume TEU of direct road transport from region i to region l
x_{ijkl}:	volume TEU of the intermodal transport from region i via terminal j and terminal k to region l

---

[1] TEU = **T**wenty foot **E**quivalent **U**nit

X_{jk} :	maximum volume TEU of a transport connection between terminal j and terminal k
D_{il} :	transport demand between region i and region l
d_{od} :	distance from location o to location d
$C_r(i,l)$:	costs per TEU-kilometer for road transport between region i and region l
$C_p(j,i)$:	costs for pick-up and delivery per TEU-kilometer from terminal j to region i
$C_s(j,k)$:	shuttle costs per TEU-kilometer between terminal j and terminal k
$C_h(j)$:	costs of transshipment/handling per TEU at a terminal j
T :	collection of selected terminals
R :	collection of all regions.

Assumptions and data. Applying this model formulation to the TEMII-data provides us with a large solution space. The model contains 129 regions in Europe with a transport demand, 21 choices for intermodal rail terminal locations, and 13 choices for intermodal barge terminal locations enabling more than 9 million alternative transport connections. To reduce the solution space, some restrictions have been added to the formulation. For instance, regions having a road distance larger than 100 kilometers to a terminal, have been excluded from the solution space since the chance that these regions will use this terminal in an optimal solution is zero. This LP-formulation allows us to make a network optimization based on costs. The results of this model indicate which terminal locations could have competitiveness towards road transport. Furthermore, based on the incoming and outgoing transport volumes, an evaluation of the geographical location of a terminal could be given. The following tariffs (1995) were applied in the model:

Table 8.3. Overview of cost drivers (Konings, 1993)

Cost drivers	Price per TEU
Road transport direct	€ 0.55 per km
Road transport drayage	€ 0.55 per km
Rail Transport	€ 0.15 per km
Barge Transport	€ 0.11 per km
Transshipment/handling rail	€ 27.27
Transshipment/handling barge	€ 22.73

Considerations about the assumptions and data:

- The cost-prices shown in Table 8.3 are an average representation of regular tariffs at that time. Especially road drayage costs can have a different cost-price structure since in congested areas their prices are based on time instead on traveled kilometers.

- The differences in prices between barge and rail (Table 8.3) are sometimes the reason that a rail terminal will not attract volume in contrast to a barge terminal, especially at origin-destination combinations that have both accessiblity to barge and rail. This situation occurs especially for the terminals Born (part of TEIT Born), Almelo (part of TEIT Twente), and Valburg (part of TEIT KAN).
- The number of TEUs calculated for each terminal is not the real demand to be expected. However, this number should be interpreted as a maximum feasible quantity for this terminal. The model expects to have an optimal market, assuming that any shipper would always choose the cheapest transport alternative available. In reality criteria like quality and deliverance reliability are also important for a shipper (Muilerman, 2001).
- The *TEMII*-data (NEA,1992) omit some precision in the distance matrices. The distances are rounded off to units of 25 kilometers. The calculated potential European shuttle-connections are quite precise. However, the determination of the exact location of a terminal is rather rough due to the scaling of 25 kilometer units. Therefore the results at regional level are only indicative.

This analysis has a special focus on the Dutch TEITs Twente, KAN, Venlo and Veendam. Table 8.4 presents the main results, and also the continental flows for Rotterdam in order to have some reference values.

Table 8.4. Modeling results for the TEITs

Terminal	# TEUs/year	Mode	#Shuttles
Veendam/Leeuwarden	77,500	Rail	5
Venlo	150,500	Rail	5
Born	175,750	Barge	4
Nijmegen	307,000	Barge	5
Almelo	113,750	Barge	5
Leiden	31,326	Barge	5
Utrecht	43,050	Barge	5
Rotterdam	890,000	Rail	8
Rotterdam	811,750	Water	6

The results for the barge terminal in Nijmegen are remarkable. At that time no continental containers were handled at this terminal. Also very remarkable is that the results show no volumes for the new to be built terminal Valburg. The favorable location of the terminal Nijmegen attracts all the possible demand in this region. For the validation of the model it was quite convenient that the total volume transshipped for the TEIT Veendam exactly matched the number of containers 77500 TEUs determined by Buck consultants at the Rail Service Center in (Buck Consultants International, 1996). The results of the model indicate a restricted attraction for the terminals mentioned in Dutch policy plans. Apart from the barge terminal at Nijmegen, the other terminals located nearby the frontiers seem to be less important. The Dutch policy plan for terminal development has forgotten to consider the international competition of terminals situated nearby the Dutch frontiers, such as Duisburg in Germany and Liège in Belgium. The Dutch policy plans do not cover terminals for short distances, i.e. less than 50 kilometers. However, the model results show good opportunities for inland terminal development with terminal distances shorter than 50 kilometers, i.e. Utrecht and Leiden/ Alphen. These results are contrary to current break-even distances, but if enough bundling of volumes can be reached, then the break-even distances can be shorter.

8.3.2 Detailed Cost Model

With the linear programming model we obtain rough indications of possible terminal locations. The next model is applied to identify more specifically where a terminal should be located on a detailed cost specification at regional level. For one of the identified terminal initiatives Leiden/Alphen, a detailed Activity-Based-Cost-model (spreadsheet-model) was developed (van Ham *et al.* 1997). In order to determine the total costs more precisely, the costs for transshipment, the cargo-handling equipment, acreage and personnel requirements are relevant. Each cost component will be discussed.

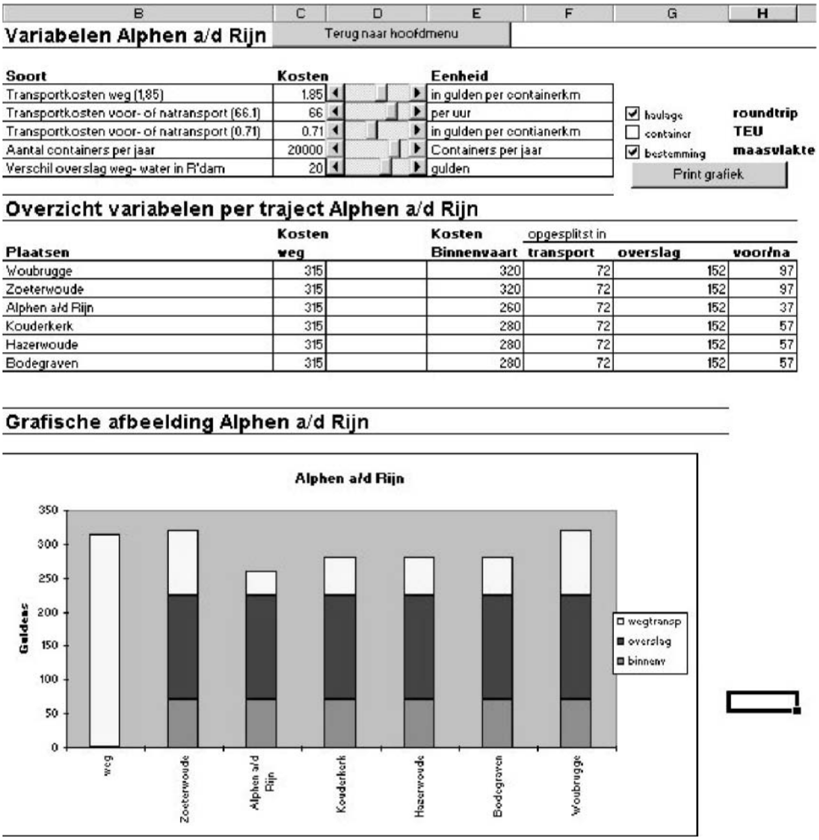
Table 8.5. Typical requirements for inland (barge) terminals (NEA/Haskoning, 1991).

Cost Drivers	<5000 TEU	5000-15000 TEU	>15000 TEU
Equipment:			
-gantry crane		1	1
-mobile crane	1		
-fork-lift truck	1		1
Personnel:			
-crane operator/	1	1	2
fork-lift driver	1	2	2
-gatehouse			
Acreage (hectares)	<0,33	0,33-1,0	>1,0

On this basis the costs of transshipment can be calculated. In this model, due to economies of scale, the costs per container will decrease. But what is even more interesting is that the costs will stabilize when throughput exceeds approximately 10 thousand containers annually.

Transportation by barge. For transportation by barge two situations can be distinguished. A new inland terminal may be located near an inland waterway, where scheduled services by barge already exist. In this case the new terminal is just an extra port of call and a fixed tariff can be arranged with the inland water carrier. Otherwise, a complete new service must be organized. The fast majority of the major Dutch inland waterways can accommodate ECMT IV class-barges with a capacity of 1500 tons or 90 20-foot containers. To charter such a barge amounts 450 thousand euros annually.

Drayage. Local pick-up and delivery of containers is usually carried out by truck. The fee charged for the movement of a container between the terminal and the point of origin/ destination differs from the tariffs in long-distance road haulage. On short distances, costs are determined more by time (€ 30.05 per hour) than by distance (€ 0.32 per TEU kilometer). For long-distance road transport an average tariff of € 0.84 per TEU kilometer was applicable.



Variabelen Alphen a/d Rijn

Soort	Kosten	Eenheid
Transportkosten weg (1,85)	1.85	in gulden per containerkm
Transportkosten voor- of natransport (66.1)	66	per uur
Transportkosten voor- of natransport (0.71)	0.71	in gulden per contianerkm
Aantal containers per jaar	20000	Containers per jaar
Verschil overslag weg- water in R'dam	20	gulden

☑ haulage **roundtrip**
☐ container **TEU**
☑ bestemming **maasvlakte**

Overzicht variabelen per traject Alphen a/d Rijn

	Kosten	Kosten	opgesplitst in		
Plaatsen	**weg**	**Binnenvaart**	**transport**	**overslag**	**voor/na**
Woubrugge	315	320	72	152	97
Zoeterwoude	315	320	72	152	97
Alphen a/d Rijn	315	260	72	152	37
Kouderkerk	315	280	72	152	57
Hazerwoude	315	280	72	152	57
Bodegraven	315	280	72	152	57

Grafische afbeelding Alphen a/d Rijn

Figure 8.6. Overview of cost comparison (model) between different terminal locations and direct road transport. The model shows that Alphen a/d Rijn is the best location. The model was developed for Dutch application.

Results of the cost model for Alphen a/d Rijn. At the time of writing this paper, a new inland terminal in Alphen aan den Rijn is under consideration. In this region several large shippers are to be found. Amongst these shippers the Heineken Brewery and the Swedish Electrolux company are the most important ones. This initiative received a warm welcome. Given the amount of cargo, forecasts indicate a throughput of at least 20,000 containers annually, or 80 containers every day. To get an impression of the terminal operations, the rail-mounted gantry crane needs approximately 5 hours to (un)load the barge. In the vessel, two out of every three slots are in use. The terminal is to be built South of Alphen aan den Rijn, which allows a daily sailing schedule to the Port of Rotterdam with just one barge. Special attention should be paid to minimizing the number of callings at the deep-sea terminals in the port area. Moreover, handling barges at the sea quay incurs extra costs. These extra terminal handling charges (THC) were added as a surplus of € 9.09 on the transshipment costs. The final cost comparison between road transport and intermodal transport showed four favorable locations in the vicinity of Heineken. Figure 8.6 shows that the total costs for transporting one container by barge are lower compared to road transport in case of the use of a terminal at the locations Alphen a/d Rijn, Kouderkerk, Haserwoude, and Bodegraven. The calculated tariffs demonstrated opportunities of cost reductions up to twenty percent of the road tariff.

8.3.3 Simulation Model for Terminal Operations

If shippers have decided to service their transport demands by intermodal transport, the management of the operations becomes an important issue. The transport schedules have to be determined in such a way that from the shippers' perspective the time conditions for delivery have to be met, from the carriers' perspective the vessel/train loads have to be filled up as much as possible, and that from the terminal agents perspective the productivities of the cranes to be guaranteed. From the municipality perspective the usage of space and related truck movements is important with respect to the development of their town-planning plans.

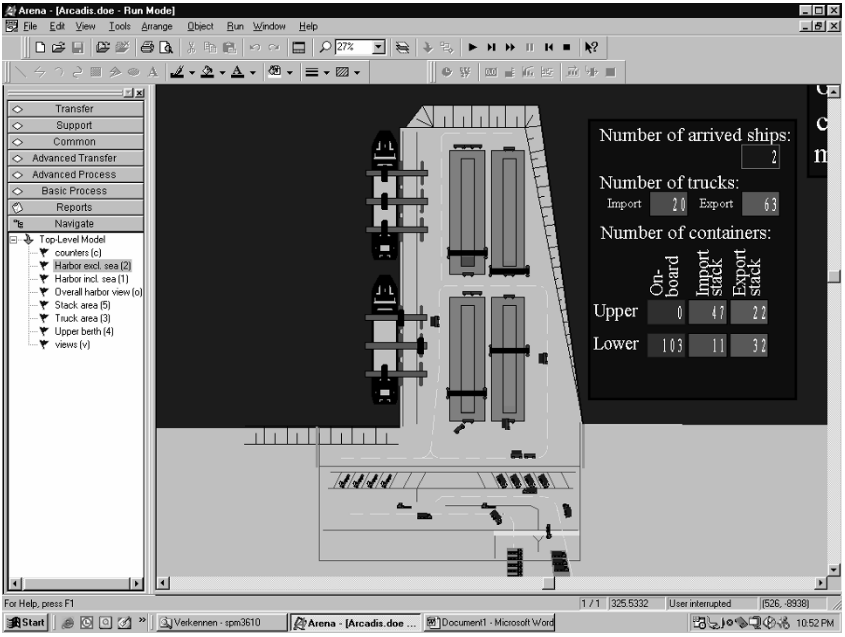
To visualize these individual important factors of the actors, the logistic processes are modeled in a simulation model. The simulation model shows the entering of a ship at a quay. When the ship has arrived, the cranes start to unload the ship. The strattle-carriers pick up the containers and take the containers to the stack. The strattle-carriers also perform the opposite action, taking containers from the stack to the position below the crane. This action is carried out by the strattle-carriers after the ship has been unloaded. Other strattle-carriers work on-demand for the trucks entering the terminal. The trucks bring their containers. If a gate is available, the strattle-carrier will pick up the container from the truck and take it to the stacking area. The opposite handling is done for the containers coming from the stack. (See Figure 8.7.)

The model has not been applied yet for the terminals Alphen a/d Rijn and Utrecht. However, with this model it is possible to determine a feasible terminal configuration with respect to the determination of the number of cranes, strattle-carriers and gates. Besides a visualization of the internal processes, it also provides insight into the space allocation used by the terminal. The seize of stacking area and the number of parking places outside the terminal can be determined. When all the gates are occupied, the trucks should park at parking places.

8.3.4 Models Review

For each scope of analysis a specific model has been developed. First of all, the selection of (new) shuttle-initiatives is determined by a LP cost-minimization model. Summing up all the shuttle-initiatives at a terminal, it is possible to identify whether enough volume will be generated. Secondly, for each identified terminal the exact location of the terminal should be determined. The first model just points to a possible region where a terminal could be located. Several alternative terminal locations are compared to each other. The main costs elements are based on the distances to the main shippers in the vicinity of the terminal, the terminal equipment applied and the number of possible roundtrips per day for a ship or train. At the end all the operational activities at a terminal will be analyzed with the simulation model.

Because of the hierarchy in the modeling approach, the higher level is directive for the level below. The approach starts from rough locations for (new) terminal location identification to more detailed location identification. If no terminal locations can be identified at a more detailed level, other terminal locations at the higher level should be studied over into more detail. This procedure continues until no attractive terminal locations can be identified.



Number of arrived ships: 2

Number of trucks:

Import 20 Export 63

Number of containers:

	On-board	Import stack	Export stack
Upper	0	47	22
Lower	103	11	32

Fig. 8.7. Snapshot of the simulation model.

8.4 Conclusions

Due to actor perception analysis, we have been able to identify critical (success and failure) factors which are felt to be important by the actors at every level. Thanks to the recognition of actors in specific arenas we can build specific decision-support models for adapting specific factors of concern. With these models we can better predict where terminal locations can stay at the long run. With regional demands and national scale it is possible to appoint in which areas terminal locations can be built. The next step is to look over the area for certain possible locations and consider the distances to the most important multi-national companies. With this kind of modeling we can exactly show where the terminal can be built more competitive to the road transport.

Based on our experiences with this approach, we are able to indicate locations for new terminals. Two identified terminal locations have led to serious terminal initiatives in practice.

The gap between the policy plans of the governments and the level of operational processes has grown too far apart. The dynamic behavior of the actors involved cannot be statically stated in policy plans, but as Muller (1995) stated: ‘intermodal transportation is not just the hardware or equipment involved with the freight movement, but the process, which becomes a major component of the systems approach to business’.

References

Bovy P (2001) Traffic flooding the low countries: how the Dutch cope with motorway congestion. Transport Reviews 21(1):89-116

Bowersox DJ, Closs DJ (1996) Logistical Management: The integrated supply chain process. Mc Graw-Hill, New York

Buck Consultants International (1996) Inlandterminals voor intermodaal transport, November. Nijmegen

Daskin MS (1995) Network and discrete location: Models, algorithms and applications. John Wiley and Sons, Inc., New York

Drezner Z (ed) (1995) Facility location: A survey of applications and methods. Springer, New York

European Commission, Directorate-General for Energy and Transport (EC-DG/ET) (2003) http://www.unece.org/trans/main/tem/temdocs/polo.pdf

EC-DG/ET (2001) European transport policy for 2010. EU white paper, Office for official publications of the European communities, http://europa.eu.int/comm/energy_transport/library/lb_texte_complet_en.pdf,

European Conference of Ministers of Transport (ECMT) (1993) Terminology on combined transport. Paris

Ernst & Young ILAS, ECORYS Nederland, Groenewout Consultants & Engineers (2002) Kerncompetentie Nederland: Europees Logistiek service Centrum van Europa. NDL, Zoetermeer

Fischer F, Forester J (eds) The Argumentative Turn in Policy Analysis and Planning. Duke University Press, London, 1993 van Ham JC, van Duin JHR (1997) Logistics: freight transport networks, alliances. In: Managing enterprises—stakeholders, engineering, logistics and achievement. Mechanical Engineering Publications Limited, London, UK, pp 331-335

Konings JW (1993) Tarieven in het containervervoer. Delft

Koopmans T, Beckman M (1957) Assignment problems and the location of economic activities. Econometrica 25:53–76

Marin B, Mayntz R (1991) Policy networks: Empirical evidence and theoretical considerations. Westview, Boulder, CO

Ministry of Transport, Public Works and Watermanagement (1994) Promotion of intermodal transport. Policy paper, The Hague

Modal shift (2004) http://home.hetnet.nl/~hoekmanp40/Modalshift.htm

Muller G (1995) Intermodal freight transportation. Lansdowne, VA

Mulvey JM (1994) Part 3. Models in the public sector: Success, failure and ethical behavior. In Mulvey JM (ed) Ethics in modeling. Pergamon, London, pp 58-73

NEA Transport Research and Training, TEM II (1992) Het goederenvervoer in drievoud to 2015, NEA Rijswijk, Nemhauser, TIMS/ORSA Joint National Meeting Chicago

Porter ME (1996) What is strategy? Harvard Business Review, November/December:61-78

Publicatieblad van de Europese Gemeenschappen (1996) Beschikking nr 1692/96/EG van het Europese Parlement en de Raad van 23 juli 1996 betreffende communautaire richtsnoeren voor de ontwikkeling van een transeuropees vervoersnet. September

van Roermund MWM, Brugge R, ter en Nijhout P (1995) Tijd voor TEIT's, Nederlandse Trans Europese Inland Terminals. Incomaas, December

Ruane F, Sutherland J (2002) Globalization, europeanization and trade in the 1990s: Export responses of foreign and indigenous manufacturing companies. Trinity College, Dublin and University of Wollongong, New South Wales.

Sijbrands MJC (1993) Strategische en logistieke besluitvorming. Tinbergen Research Institute, nr. 53, Amsterdam.

Spin (2002) Deliverable 1 Actors and factors in transport mode decisions in supply chain. GRD2/2000/30187-SI2.324496 SPIN, January.

Vidal CJ, Goetschalckx M (1997) Strategic production-distribution models: A critical review with emphasis on global supply chain models. European Journal of Operational Research 98(1)

Vos GCJM (1993) International manufacturing and logistics; a design method. Enschede

9 The Evolution of OECD ICT Inter-Cluster Networks 1970-2000: An Input-Output Study of Changes in the Interdependencies Between Nine OECD Economies*

Brian Wixted[1] and Russel J. Cooper[2]

[1] Centre for Policy Research on Science and Technology, Simon Fraser University, Vancouver BC, Canada
[2] Department of Economics, Macquarie University, Australia

9.1 Introduction

Cluster studies are now common in economic geography and regional science. These studies have added significantly to our knowledge of the processes behind economic development and industrial location. However, many of these studies have suffered from a tendency to pay little attention to external relations that may exist between particular clusters and other regions. The predominant focus on the endogenous drivers of development and the dynamics of intra-cluster interdependencies has therefore been at the expense of exploring the connection between clustering and the observed fragmentation of economic value chains. However, the alternative of focusing excessively on vague notions of globalization and manufacturing outsourcing are more damaging, distracting attention away from the continuing economic advantages of real places.

The international organization of production and value chains thus needs to integrate an appreciation for the technological advantages of particular innovation systems with the reality of a growing fragmentation of production. These trends need to be viewed as complementary outcomes rather than as contradictory interpretations of current economic processes. Continuing specialization in the divi-

---

* This chapter emerged from the first author's doctoral research program at the University of Western Sydney with preliminary data presented at the Regional Science Association International 2004 World Congress 14-17 April 2004, Port Elizabeth, South Africa. His contribution to this chapter was written while he was a visiting scholar with the Centre for Policy Research on Science and Technolgy, Simon Fraser University, in Vancouver.

sion of labor would suggest that if clusters matter, then production is likely to be specifically spatially structured and connected. A study of the evolution of connections between places should therefore provide us with insights into the strengthening or weakening position of various cluster relationships across time. Such an analysis of *inter-cluster networks* attempts to go beyond the corporate networks approach (global production networks) to adopt a framework which operates at the level of systems of innovation but is developed within a multi-spatial environment. The present chapter therefore develops an analysis of inter-cluster relations between nine countries for the period between 1970 and 2000, utilizing input-output data from the OECD that is used to construct a multi-country model. The chapter focuses on the manufactured segment of the ICT (information and communication technologies) sector because, as the data confirm, these related production systems include some of the most rapidly changing of all industrial systems.

Wixted (2005) showed that *national*[1] clusters have become increasingly dependent on component imports. They have also become relatively more import intensive when compared to the import intensiveness of other production systems. Building on this perspective and concentrating on the ICT sector, the current study pays particular attention to evidence of cross-border inter-cluster relationships that appear to be driven by national ICT clusters. Interestingly, the degree of dependency on particular supply relationships between major OECD industrial economies weakened during the period of study, with the rise of the East Asia producers. Nevertheless, there is a strong hierarchical system within ICT production. The key hubs for components are Germany (for European Countries), the USA (Europe, Asia and North America) and Japan (Asia-Pacific, North America and to a lesser extent Europe). These key hubs retain their status as central suppliers even when using a twenty-country dataset for the year 2000.

The next section will explore the literature on clusters as production network nodes; section 9.3 will present the multi-regional model for inter-cluster interdependencies. The evolution of country requirements for imported components will occupy section 9.4, providing the basis for the presentation of the changing spatial structure of ICT inter-cluster networks over the period, 1970-2000 in section 9.5. Some concluding comments complete the chapter.

---

[1] National clusters for a given ISIC industry category are defined those systems of user-supplier interactions which, for a given increase in activity of a member firm, produce significant additional (indirect) value added within the ISIC industry category. The measure of significance needs to be related to a base-case level of ‘standard’ indirect value added creation. Retention of indirect value added within a sector at a rate above some typical standard is indicative of complementary relationships between firms in that sector. These concepts are given precise formulations below.

9.2 Clusters as Production Network Nodes

Regardless of the perspective or the methodology, the agglomeration of related production (horizontal or vertical) and innovation activities into a limited geographical area appears to be an economic phenomenon at whatever scale one may wish to investigate. Industrial agglomeration occurs across all spatial scales; e.g. Malmberg and Maskell (2001: 443) highlight the existence of micro-agglomerations (Madison Avenue New York has a concentration of advertising companies) as well as the existence of macro multi-country industrial structures (the concentration of industrialization in the European banana). However, the idea of clustering goes beyond production agglomeration. Although there is still a need for a theory of clustering (see Maskell 2001), the term 'cluster' does draw attention to supporting institutions and importantly the interdependencies between the different actors. The predominant focus for analysts has been on intra-cluster dynamics, specifically the role and types of interdependencies. These include:

> supply networks: proximity matters for some industries more than others - see e.g. Steinle and Schiele (2002);
> industrial agglomeration: centripetal forces to co-locate occur at many spatial scales - see e.g. Malmberg and Maskell (2001);
> user-producer relationships: such relationships are critical for developing the trust and knowledge flows required to develop innovations - see e.g. von Hippel (1988) and DeBresson *et. al.* (1998); and
> market milieux: knowledge of contractors and suppliers in a local setting may reduce the risks associated with open bid contracts and promote agglomeration - see e.g. Maskell and Lorenzen (2004).

In the main, most of these are not considered to be important variables in the co-location of industries. Despite the Porterian emphasis on the role of related industries - see Wolfe and Gertler (2004) for a critique[2] - the driving force behind agglomerations is often attributed to untraded interdependencies in tacit knowledge - see e.g., Storper (1995). Tacit knowledge is not codified into scientific journals and as it is often thought to be constructed around solving particular problems it is also thought to be geographically specific (see Audretsch and Feldman 1996) because it is accumulated through experience and passed on from one person to another. Saxenian (1994) attributes some of the success of Silicon Valley in the early years, to 'pub talk'. The richness of the flows of knowledge between organizations is thought to be a critical aspect of clustering (see, for example, the commentary by Breschi and Malerba, 2001) although Breschi and Lissoni (2001) believe that tacit knowledge is not a pure untraded spillover, but is instead regulated by market mechanisms and the organizations involved.

These findings on why clusters develop are useful, but being based on case studies the analyses often lack a consideration of the economic scale or power of

---

[2] 'This suggests that at least two corners of Porter's famous diamond—sophisticated and demanding local customers and strong rivalry between local competitors—are not consistently present in the Canadian context' (2004: 1086).

the cluster being considered or even the relevance of its adjacency to markets. Demand, as Bresnahan et al. (2001) note is often ignored in such studies.

However, the literature is not totally devoid of evidence that clusters might be more appropriately analyzed as nodes on elongated supply chain networks. Wolfe and Gertler (2004) reveal that national and international knowledge links can be as important as local links for successful clusters. Simmie (2002, 2004) and Simmie et al. (2002) have shown that purchaser or supplier relations might extend beyond the boundaries of cities. However, these authors do not emphasize the extra-territorial organization of those relations, only that they are as important as local ones. Amin and Thrift (1992), earlier, suggested that Marshallian agglomerations would be better understood as nodes on value chains. The world cities literature has developed around a framework emphasizing the clustering of business service activities in particular cities that are in turn linked to other cities in hierarchical global networks (see, for example, Sassen 2001; Smith and Timberlake 2001; and Beaverstock et al. 1999). In a similar vein, Bresnahan et al. (2001) have argued that successful emerging ICT clusters in East Asia, North America and Europe are focusing on complementary rather than competing products.

Further, and in contrast to those that argue that clustering relies on tacit knowledge which by its nature is spatially confined, Saxenian and Hsu (2001) argue that technical communities can stretch over vast distances, as they propose happens with the links between Taiwan and Silicon Valley. Indeed, it would appear that connections to global knowledge sources are important and that global connectedness is only likely to increase (see Bathelt et al. 2004). As Leamer and Storper (2001) argue, ‘the Internet … will also probably make possible greater linkages between different localized clusters at very long distances’ (2001, p. 658). Therefore, frameworks that isolate geographic clusters from the wider economic space appear to have important limitations, as Sturgeon has recently commented:

> We need to better understand the various roles that local agglomerations play within spatially extensive value chains and begin to map the activities that tend to concentrate in particular places even as the geographic ‘footprint’ of linked economic activity expands. It is the linkages mechanisms, between firms and between places, that especially deserve more of our research attention (2003, p. 200).

Wixted (2005) has shown that by adopting a methodological tool for measuring cluster networks[3] and utilizing inter-country input-output modeling, it is possible to reveal the different spatial structures of interdependencies (the geographic footprint) for different international production systems. In that study, Wixted reported on the results from an analysis of the interactions of fifteen European Union Member countries (1995 data) and nine OECD countries (1990 data) for transport (*motor vehicles* and *aerospace*) and ICT (*office, accounting and computing machinery* and *radio, TV and communications equipment*) national clusters. The present chapter extends the ICT component of this earlier work by developing the analysis into an inter-temporal context for the nine OECD countries for the period

---

[3] Extending concepts and initial analysis in Wyckoff (1993).

1970-2000. The chapter also takes advantage of the greater number of countries included in the OECD (2002) database to present a snapshot of interdependencies in 2000 between twenty countries.

9.3 Multi-Regional Input-Output Modeling of Inter-Cluster Interdependencies

This chapter adopts the term 'national cluster' to categorize a group of activities centered on supplier and buyer patterns for particular production systems contained within the borders of nation-states. Although it would be preferable to define clusters at a regional level,[4] at the present stage multi-spatial data are rare below that of nation-states. The concept has often been utilized previously - e.g. macro clustering, see OECD (1999 and 2001). The approach is also relevant because it has been shown that user-producer relationships, that are important for generating innovation (see von Hippel 1988), have been related to the interactions apparent in input-output data (DeBresson *et al.* 1998).

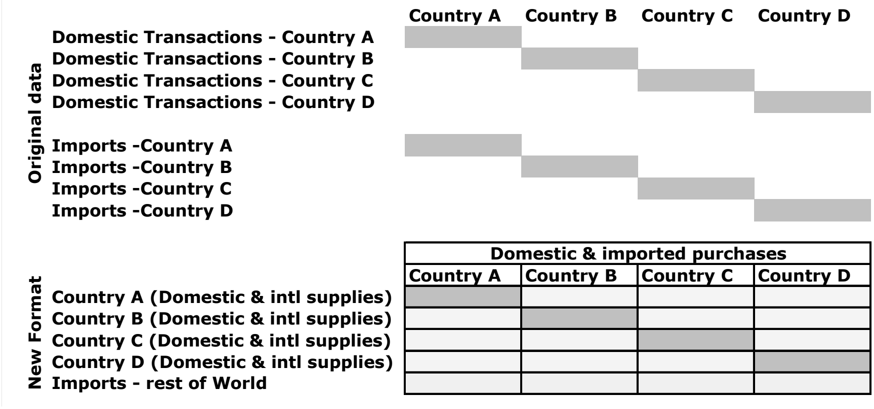
9.3.1 Constructing the Transactions Matrix

The current work utilizes the OECD (1996 and 2002) input-output databases. The 1996 version includes data for nine countries over the approximate time period of 1970 through to 1990. The 2002 version, released in early 2005, contains data for twenty countries with specific tables for particular countries applying to different years ranging from the early to late 1990s.[5]

---

[4] The OECD group of countries includes a wide variety of economic and geographic sizes. Thus in terms of understanding specific cluster dynamics the modelling can be understood on a sliding scale of being strongest for the small geographically concentrated economies and weakest for the larger economic and geographic countries (the USA).

[5] The 1996 version contains data for a tenth country (Italy) but for only one time period (1985). The actual years for each country in the 2002 version and further details on both versions including a product concordance are contained in an Appendix available on request. Specific problems and issues with use of the OECD data and with construction of a single multi-regional matrix from individual country input-output tables are detailed in Wixted (2005). The discussion here focuses on general principles behind construction of the multi-regional transactions tables.



Note: Each block represents a sub-matrix of 36 * 36 cells in the 1996 database and 42 *42 cells in the 2002 database.

Fig. 9.1. Transaction matrix construction

The procedure for developing the multi-regional transactions matrix is straightforward, theoretically. The domestic input-output tables from the databases are used to form the diagonal of the new transaction matrix. Trade tables consisting of industry (rows) and country sources (columns) are created from bilateral trade data (OECD 2004b) for each country in the model together with a rest of the world category. These trade ratios are then used to distribute the imports supply and use tables to create hypothetical[6] inter-country sub-matrices for each partner country for every country in the model. Fig. 9.1 illustrates this process diagrammatically.

Since the national data corresponds to different years for various countries, the choice of date for the calculation of trade ratios required a compromise between a useful policy relevant date and a date that would be appropriate for the original data. In the early dataset (OECD 1996), this was not a major issue as many of the national tables were for a year that was close to the ‘useful’ dates of 1970 and 1990. However, in the latter dataset the dates of the data years ranged from 1992 to 1998. As many of the national tables were for the late 1990s and because input-output ratios do not often change quickly, it was considered appropriate, given our interest in the trade pattern (spatial structure), to develop the transactions table with trade data for the year 2000. The more serious problems occur with the calculation of appropriate trade ratios for service sector activities. Bilateral trade data for services activities is only available on an industry-to-country basis for a few industries and countries. Although these data exist for the last few years, they cannot be reliably backcast to create historical series. Fortunately, services trade has been *relatively* small (although growing rapidly). As a proxy for the geo-

---

[6] Trade data are not collected on a supplying country domestic user basis and thus these approximations are necessary.

graphic preference for trade by countries, we have used the aggregate manufacturing trade pattern where other data are not available.

9.3.2 The Modeling

The methodology that follows is significant because apart from Wixted (2005) this is the first use of Cooper's (2000) suggested approach to calculating net value added input-output flows with empirical data sets. To articulate the notation of the modeling, we use here the OECD 1970 and 1990 modeling year structure.

Following assembly of the multi-regional multi-sectoral transactions table for the OECD economies, a mirror coefficients matrix was constructed by dividing elements in each column by the total value of the sectoral output. The input-output model itself is standard. However, in analyzing results from the model, emphasis is placed here upon the flow of value added effects rather than the multipliers. This chapter also presents specific non-standard measures related to indirect value added which are designed first of all to identify national clusters and secondly to help in the search for linkages between national clusters. These specialized measures need to be discussed. To do this efficiently, it will be useful to summarize the analytical model and introduce relevant notation.

Each of the 9 regions in this model distinguishes 36 sectors. In addition, there is a Rest of the World region providing imports, also disaggregated into 36 sectors. Exports from the OECD economies to the Rest of the World are aggregated into one sector and combined with other final demand. To describe the basis of measurement of indirect value added, it is initially useful to consider this model as one huge (9 times 36)-sector model, without necessarily distinguishing the different regions. Following this, the multi-regional structure will be discussed in order to highlight the ability of the model to trace the regional distribution of value added. A general transactions table for the (9 times 36)-sector overall model can be represented as

$$\begin{array}{cc|c} T & f & x \\ T_R & 0 & \\ w' & 0 & \\ \cline{1-3} x' & & \end{array} \qquad (9.1)$$

where:

T is a [(9 times 36) x (9 times 36)] multi-sectoral multi-regional transactions table of intermediate flows;

T_R is a [36 x (9 times 36)] matrix of import flows—intermediate goods imported by the 36 sectors and 9 regions in the model—originating from 36 sectors distinguished in the Rest of the World;

f is a [(9 times 36) x 1] column vector of combined exports and other final demand;

x is a similarly dimensioned - i.e. [(9 times 36) x 1] - column vector of sectoral outputs; and
w' is a[1 x (9 times 36)] row vector of gross value added for the various sectors.
Dividing the 9 times 36 inter-industry flow columns in (9.1) by their column totals (and concentrating just on these), the transformation to coefficients can be depicted as

$$\begin{array}{l|} T \\ T_R \\ w' \\ \hline x' \end{array} \quad \Rightarrow \quad \begin{array}{l|} A \\ A_R \\ v' \\ \hline i' \end{array} \tag{9.2}$$

where:
A is a [(9 times 36) x (9 times 36)] matrix of multi-sectoral multi-regional input-output coefficients;
A_R is a [36 x (9 times 36)] matrix of import coefficients - originating from the 36 sectors distinguished in the Rest of the World;
v' is a [1 x (9 times 36)] row vector of value added coefficients; and
i' is a unit row vector depicting the vertical adding up identity in (9.2), viz..

$$i'A + i'A_R + v' = i' \tag{9.3}$$

Traditional application of the input-output model involves combination of the horizontal adding up identity in (9.1), viz.

$$Ti + f = x \tag{9.4}$$

with the Leontief technology assumption

$$T = A\hat{x} \tag{9.5}$$

where $\hat{x}$ denotes that the output vector is laid out as a diagonal matrix of column multipliers. The combination suggests:

$$\begin{aligned} & x = Ti + f \\ & \Rightarrow x = A\hat{x}i + f \\ & \Rightarrow x = Ax + f \\ & \Rightarrow x = (I - A)^{-1} f \end{aligned} \tag{9.6}$$

thereby introducing the Leontief inverse $(I - A)^{-1}$ as a multiplier matrix. In addition to its ‘multiplier’ interpretation in (9.6), however, the Leontief inverse has an interesting distributional implication when applied to (9.3). This combination yields:

$$\begin{aligned}
& i'A_R + v' = i' - i'A \\
& \Rightarrow i'A_R + v' = i'(I-A) \\
& \Rightarrow i'A_R(I-A)^{-1} + v'(I-A)^{-1} = i' \\
& \Rightarrow i'A_R(I-A)^{-1} + v'\left[(I-A)^{-1} - I\right] + v' = i' \\
& \Rightarrow i'A_R(I-A)^{-1} + i'\hat{v}\left[(I-A)^{-1} - I\right] + v' = i'
\end{aligned} \tag{9.7}$$

and this suggests a further rearrangement of (9.2) to highlight an indirect adding-up identity:

$$\begin{array}{c|} A \\ A_R \\ v' \\ \hline i' \end{array} \quad \Rightarrow \quad \begin{array}{c|} \hat{v}\left[(I-A)^{-1} - I\right] \\ A_R(I-A)^{-1} \\ v' \\ \hline i' \end{array} \tag{9.8}$$

The LHS of (9.8) indicates that a dollar's worth of product in any sector of any region goes partly as direct value added to that region (the relevant element of v'), partly as payment for imports (the relevant column of A_R) and partly as payments for the intermediate inputs purchased from any of the sectors and regions distinguished in the model (the 9 times 36 components of the relevant column in A). Of course, intermediate input payments end up themselves as value added for some sector/region or as additional import leakages after further rounds of inter-industry transactions. The final direct and indirect distribution of value added is indicated on the RHS of (9.8). A dollar's worth of production in a sector ultimately leads to a reallocation of the value of intermediate inputs to indirect value added. A portion of this dollar's worth of production leaks from the nine OECD countries as indirect value added to imports. This portion is contained in the relevant column of $A_R(I-A)^{-1}$. The portion that is retained within the OECD, and is redistributed among the various sectors and regions distinguished in the model, is given by the components of the relevant 9 times 36 element column of $\hat{v}\left[(I-A)^{-1} - I\right]$.

The importance of the viewpoint indicated by RHS (9.8) is especially apparent when the various regions are explicitly distinguished. However, before developing the implications of (9.8) for interpretation of cross-border inter-cluster relationships, we pause to point out that the structure made explicit in (9.8) also applies for the single country input-output model and at that level is capable of providing a methodology for identifying national clusters. To develop this perspective briefly, interpret A on LHS (9.8) as the input-output coefficient matrix for one country. Then A_R contains import coefficients from the rest of the world. We can assume for discussion that A_R is at the same level of sectoral disaggregation as is distinguished in A. Now, consideration of RHS (9.8) shows where such a national model would imply that indirect value added goes. Notice that some indirect value added must leak to the rest of the world.[7] Given the overall equality of column sums on RHS and LHS (9.8), this means that any column sum in $\hat{v}\left[(I-A)^{-1}-I\right]$ can be no greater than the corresponding column sum in A. On an element by element basis, the ij^{th} element of $\hat{v}\left[(I-A)^{-1}-I\right]$ must typically be less than the ij^{th} element of A. Cases where this inequality is reversed will stand out. They would imply that, once activity in sector j generates first round intermediate demands for the products of sector i, further rounds of intermediate purchases to support the activity must be intensifying the flow of value added to sector i. This can only be consistent with the existence of complementarities between firms in sector i and thus it could be used as an indicator of the presence of a national cluster associated with sector i.[8]

Suppose now a method for identifying national clusters has been applied. This could, for example, involve use of that a national variant of (9.8) in the manner described above. We now wish to describe how a multi-regional variant of (9.8) can be used to further identify cross-border links with such national clusters – effectively identifying trans-national inter-cluster linkages. In the case of the OECD model, the number of distinct regions is n = 9 countries. Thus, the (9 times 36) x (9 times 36) coefficient matrix A is represented in detail by:

$$A \Rightarrow \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} \tag{9.9}$$

where A_{kj} denotes a 36 x 36 inter-country coefficient matrix depicting sectoral specific flows from country k to country j. Using this notation, the multi-regional variant of (9.8) can be written out in country by country detail as

---

[7] This is clear because $A_R(I-A)^{-1} > A_R$ (each element is 'greater than or equal to' in this matrix inequality).

[8] At high levels of sectoral aggregation this approach needs to be treated with caution, since it may be that some elements of the i^{th} row of $\hat{v}\left[(I-A)^{-1}-I\right]$ are greater than corresponding elements of A, while the reverse is true of other components of the i^{th} row. On face value, such contradictory evidence probably should be taken as an indicator of heterogeneity within the sub-industries aggregated into sector i. A more cautious approach would look for consistent corroborative evidence across row i before concluding that sector i firms make up a national cluster.

$$\begin{array}{c|}\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn}\end{bmatrix} \\ \begin{matrix} A_{R1} & \cdots & A_{Rn} \\ v_1' & \cdots & v_n' \end{matrix} \\ \hline \\ i' \end{array} \quad \Rightarrow \quad \begin{array}{c|}\begin{bmatrix} \hat{v}_1 A_{11}^{(n)} & \cdots & \hat{v}_1 A_{1n}^{(n)} \\ \vdots & & \vdots \\ \hat{v}_n A_{n1}^{(n)} & \cdots & \hat{v}_n A_{nn}^{(n)}\end{bmatrix} \\ \begin{matrix} A_{R1}^{(n)} & \cdots & A_{Rn}^{(n)} \\ v_1' & \cdots & v_n' \end{matrix} \\ \hline \\ i' \end{array} \tag{9.10}$$

where $A_{kj}^{(n)} = A_{kj} + A_{k\bullet} A_{\bullet j}^{(n)}$, $k = 1,\ldots,n$ and $A_{Rj}^{(n)} = A_{Rj} + A_{R\bullet} A_{\bullet j}^{(n)}$ and the notation on RHS (9.10), defined in (9.11) and (9.12) below, is designed for consistency with results on partitioned inversion of the Leontief inverse developed in Cooper (2000). Specifically, for an *n*-block partition, Cooper shows that the Leontief inverse can be represented as:

$$(I - A)^{-1} = I + A^{(n)} \tag{9.11}$$

where $A^{(n)}$ is constructed recursively by successive addition of blocks, viz.:

$$A^{(i)} = A^{(i-1)} + A_{\bullet i}^{(i-1)} (I - A_{ii}^{(i-1)})^{-1} A_{i\bullet}^{(i-1)}, \quad i = 1,\ldots,n \tag{9.12}$$

starting from $A^{(0)} = A$, where $A_{\bullet i} \equiv \begin{bmatrix} A_{1i} \\ \vdots \\ A_{ni} \end{bmatrix}$ and $A_{i\bullet} \equiv \begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix}$.

Apart from computational advantages in multi-country/sector cases and the analytical advantage of being able to compute the Leontief inverse for sub-groups of countries along the way, the main advantage of this formulation for present purposes is an interpretational one. To develop this aspect, note that (9.11) and (9.7) imply:

$$v' A^{(n)} + i' A_R \left[I + A^{(n)} \right] + v' = i' \tag{9.13}$$

Now $A^{(n)}$ is by definition the matrix of sectoral multipliers abstracting from an initial unit injection to final demand. Moreover, the double entry accounting ensures that a unit of final demand either leaks out of the nine OECD countries via payment for imports or eventually finds its way to a unit of value added somewhere within the various regions distinguished in the model. Thus $v' A^{(n)}$ measures EU-wide indirect value added. Equivalently, by comparison of (9.13) with (9.3) we have:

$$i' A = v' A^{(n)} + i' A_R A^{(n)} \tag{9.14}$$

so that the total value of within OECD intermediate inputs is fully accounted for either by contributing to indirect value added within the nine OECD or by further leakages via imports (in successive rounds of inter-industry transactions).

Utilizing the detail in (9.9), we may define a 9 x (9 times 36) matrix of partial intermediate sums:

$$B = \begin{bmatrix} i'A_{11} & \cdots & i'A_{1n} \\ \vdots & & \\ i'A_{n1} & \cdots & i'A_{nn} \end{bmatrix} \tag{9.15}$$

where the unit vectors in (9.15) are each 36 element. Thus the matrix B summarizes the direct intermediate input value flows from sectors (in columns) to whole regions/countries in rows and is simply a partial aggregation of the A matrix with the same column totals. From (9.9) and (9.15) it follows that $i'A = i'B$. On the other hand, given that $A^{(n)}$ is constructed from an *n* = 9 block partitioned system, we can represent it in block detail as:

$$A^{(n)} = \begin{bmatrix} A_{11}^{(n)} & \cdots & A_{1n}^{(n)} \\ \vdots & & \vdots \\ A_{n1}^{(n)} & \cdots & A_{nn}^{(n)} \end{bmatrix} \tag{9.16}$$

and we now wish to use this to define an 9 x (9 times 36) matrix of partial indirect value added sums:

$$B^{(n)} = \begin{bmatrix} v_1'A_{11}^{(n)} & \cdots & v_1'A_{1n}^{(n)} \\ \vdots & & \vdots \\ v_n'A_{n1}^{(n)} & \cdots & v_n'A_{nn}^{(n)} \end{bmatrix} \tag{9.17}$$

Thus the matrix $B^{(n)}$ is a partially aggregated variant of $A^{(n)}$ with value added coefficient weights employed in the aggregation across sectors in any given region. By construction, $i'B^{(n)} = v'A^{(n)}$. Given (9.14) and the definitions (9.15) and (9.17), it is also obvious that:

$$i'B = i'B^{(n)} + i'A_R A^{(n)} \tag{9.18}$$

which expresses in an aggregated form the identity that the value of within-OECD intermediate inputs is equal to within-OECD indirect value added plus indirect import leakages. What is equally clear, however, is that there is no necessary reason why the individual (region/country specific) elements within (9.15) and (9.17) should be related in any obvious way. In fact, it is the differences between these elements with which we are especially concerned in this analysis. Consider a typical (row vector) element in both matrices, say the 1 x 36 row vector $i'A_{ij}$ in the B matrix (9.15) compared to $v_i'A_{ij}^{(n)}$ in the $B^{(n)}$ matrix (9.17). A particular element of these vectors, say the k^{th}, relates to the effect of activity within sector k in country j. The k^{th} element in $i'A_{ij}$ measures the extent to which sector k in country j has business links (in the form of purchasing relationships) with country i. The value of these within-OECD intermediate purchases ultimately becomes value added to some sectors in some countries within the OECD or leaves the nine OECD countries subsequently as indirect imports. However, subject to the overall restriction (9.18), there is no reason for the indirect value added flow resulting from the chain of interactions set off by these intermediate purchases to necessarily favor country i. Of particular interest, in fact, are cases where the k^{th} element in $v_i' A_{ij}^{(n)}$ is substantially greater than the corresponding element in $i' A_{ij}$. This can only arise from a further chain of activity that has some degree of focus within country i, which then results in country i accumulating additional indirect value added as a result of a chain of activity which has begun in sector k of country j but has now switched to interactions within country i.

To investigate these types of effects we can construct the difference matrix:

$$D = B^{(n)} - B \tag{9.19}$$

Of course, the restriction (9.18) means that there is a negative sum game with respect to the columns of D - viz.

$$i' D = -i' A_R A^{(n)} < 0' \tag{9.20}$$

Result (9.20) expresses the fact that final demand for the output of sector k in country j does have to ultimately translate to a combination of value added somewhere within the OECD and to leakages via imports. The elements of any column of $B^{(n)}$ represent a reallocation of the values in the corresponding column of B less indirect leakages. While any element in D may be either positive or negative, they are (by construction) on balance negative since the column sums are negative (at least, non-positive).

Given the above, our particular interest is in elements of D which are atypically large in size (and positive in sign). To pre-empt the results of our analysis somewhat, we should note that typically the row vector elements that make up the block diagonal in (9.19) will be negative (and typically relatively large - this is invariably true for our results and would commonly be expected although it need not necessarily be so). To see why this result is likely and to investigate its implications it is useful to exhibit a detailed variant of (9.19), viz.:

$$D = \begin{bmatrix} v_1' A_{11}^{(n)} - i' A_{11} & \cdots & v_1' A_{1n}^{(n)} - i' A_{1n} \\ \vdots & & \vdots \\ v_n' A_{n1}^{(n)} - i' A_{n1} & \cdots & v_n' A_{nn}^{(n)} - i' A_{nn} \end{bmatrix} \tag{9.21}$$

The typical block diagonal row vector in D is $v_j' A_{jj}^{(n)} - i' A_{jj}$. The reason why this is likely to be a row vector of negative numbers is as follows. The second term $i' A_{jj}$ represents the sum of within country direct intermediate inputs for each sector. For the most part, these are large compared to cross country intermediate input coefficients (which represent trade in intermediate inputs). Then, as second and third round effects are calculated to determine the eventual location of the value added, it is clear with a 9 country model that there will be dispersion of value added from any country to the 8 other countries with only a small flow back to the originating country in general. Thus, $v_j{}'A_{jj}^{(n)}$ will tend to be smaller than $i'A_{jj}$, unless there are very unusual cross-border inter-sectoral relationships. This, of course, should not be seen as a problem for country j. After all, sectoral final demand increases that originate in country j provide direct value added per unit of output as indicated in the row vector $v_j{}'$. As pointed out, the value of within-country-j direct intermediate inputs $i'A_{jj}$ is likely to be quite large relative to cross border intermediate inputs such as $i'A_{ij}$ and in subsequent rounds of inter-industry relationships indirect value added is likely to be transferred from country j to other countries such as i through a process of depletion of the value implied by $i'A_{jj}$ and either addition to the value implied by $i'A_{ij}$ (for $i \neq j$) or further OECD import leakages.

This line of reasoning also suggests that if there are positive elements in D, these are to be found predominantly in the block off-diagonal row vectors. It is the relative sizes of these that will be of particular interest. We normalize each column of D by calculating the indirect value added flows as percentage changes relative to a base level of potential indirect value added. For the base level we take the value of intermediate inputs (that is, based on direct business purchasing agreements), since this is the source of indirect value added in subsequent rounds of interactions, although it may be depleted by import leakages. Let $D*$ denote the 9 x (36 times 9) table of these results. Then:

$$D* = 100*\left(B^{(n)} - B\right)*\widehat{(i'B)}^{-1} \tag{9.22}$$

where $\widehat{i'B}$ denotes a diagonal matrix formed from the vector $i'B$.

The information contained in the matrix $D*$ will be useful for identifying particular countries which need to be examined in greater sectoral detail when attempting to determine whether a particular sector, wherever located, is linked through value added flows to other sectors to such an extent as to be suggestive of a cluster relationship. To support the development of this interpretation consider the detailed representation:

$$D* = 100*\begin{bmatrix} v_1{}'A_{11}^{(n)} - i'A_{11} & \cdots & v_1{}'A_{1n}^{(n)} - i'A_{1n} \\ \vdots & & \vdots \\ v_n{}'A_{n1}^{(n)} - i'A_{n1} & \cdots & v_n{}'A_{nn}^{(n)} - i'A_{nn} \end{bmatrix} \begin{bmatrix} \widehat{i'A_{\bullet 1}}^{-1} & & \\ & \ddots & \\ & & \widehat{i'A_{\bullet n}}^{-1} \end{bmatrix} \tag{9.23}$$

In the j^{th} column block, the k^{th} element of the j^{th} row vector block, $100*\left(v_j{}'A_{jj}^{(n)} - i'A_{jj}\right)_k / \left(i'A_{\bullet j}\right)_k$ say, is likely to be negative. However, elements of the remainder of this column may be positive, although the entire column must be non-positive. A typical such element (the k^{th} element of the j^{th} column block in row i of $D*$, for example) is $100*\left(v_i{}'A_{ij}^{(n)} - i'A_{ij}\right)_k / \left(i'A_{\bullet j}\right)_k$. In fact, the interpretation of these elements off the block diagonal is that they represent the indirect value added that spills over into other OECD countries, relative to the indirect value added potentially available for creation, as a result of the fact that activity in sector k of country j creates flow-on activity in those other OECD countries. Since there are 8 such 'other' countries in this model, plus another region representing the rest of the world, and since, once indirect leakages are also accounted for, the sum of these effects exactly balances the (typically) negative term $100*\left(v_j{}'A_{jj}^{(n)} - i'A_{jj}\right)_k / \left(i'A_{\bullet j}\right)_k$, the total of these within-OECD effects can be at most $-100*\left(v_j{}'A_{jj}^{(n)} - i'A_{jj}\right)_k / \left(i'A_{\bullet j}\right)_k$, although this potential maximum will clearly be depleted by indirect imports. Given the potential indirect import depletion, the average value of these off diagonal elements is no more than $-(1/n)*100*\left(v_j{}'A_{jj}^{(n)} - i'A_{jj}\right)_k / \left(i'A_{\bullet j}\right)_k$.

From (9.20), it also follows that, within the j^{th} column block, the k^{th} column of D sums to the non-positive value $-(i'A_R A_{\bullet j}^{(n)})_k \leq 0$, and the corresponding sum for this column of $D*$ is $-100*(i'A_R A_{\bullet j}^{(n)})_k /(i'A_{\bullet j})_k \leq 0$. As argued above for the matrix D, so also for the normalized matrix $D*$, the block diagonal row vector will typically be negative. It will only be in exceptional circumstances that such a term is positive. Given these two points, the column sum $-100*(i'A_R A_{\bullet j}^{(n)})_k /(i'A_{\bullet j})_k \leq 0$ also serves as a reasonable lower bound for the sum of the off diagonal elements in the k^{th} column within the j^{th} column block of $D*$. Let $d_{ijk}*$ indicate the expected value for the i^{th} element ($i \neq j$) in the k^{th} column within the j^{th} column block of $D*$. Since there are n–1 such off-diagonal terms, for any given off-row-diagonal elements in the k^{th} column contained in the j^{th} column block, a reasonable lower bound on the average value $d_{ijk}*$ is $-(1/(n-1))*100*(i'A_R A_{\bullet j}^{(n)})_k /(i'A_{\bullet j})_k \leq 0$. Combining the upper and lower bound information on the typical value to expect in off-row-diagonal terms in $D*$, we could expect to find that

$$\begin{aligned} &-(1/(n-1))*100*(i'A_R A_{\bullet j}^{(n)})_k /(i'A_{\bullet j})_k \leq d_{ijk}* \leq \\ &\qquad -(1/n)*100*\left(v_j{}'A_{jj}^{(n)} - i'A_{jj}\right)_k / \left(i'A_{\bullet j}\right)_k \end{aligned} \tag{9.24}$$

Our particular interest is in abnormally large individual effects, which are likely to be balanced by a dispersion of lower values for other elements in the same column of $D*$. Since countries will naturally have developed trade links with particular partners, it is not likely that the values of $100*\left(v_i{}'A_{ij}^{(n)} - i'A_{ij}\right)_k / \left(i'A_{\bullet j}\right)_k$ for $i = 1,...,n,\ i \neq j$ will be spread evenly along the continuum indicated by the upper and lower bounds for the typical value $d_{ijk}*$. It is likely that some (major trading partners) will be located above the average upper bound while others (less closely linked) will be positioned nearer the average lower bound. However, with successive rounds of inter-industry relationships modifying the primary trading relationships, it would be expected that there would be some tendency for these "trading partner" effects to be ameliorated and for a greater spread in indirect value added adjustments to be evident along the continuum. What is much less likely, without the need for special explanation, would be to observe elements of $D*$ that are substantially above the average upper bound. Recalling that the average upper bound will be the operative average under the extreme assumption that there are no indirect import leakages, under this extreme assumption together with the null hypothesis that all regions are equally attractive as accumulators of indirect value added within the nine OECD economies, and that the distribution of indirect flows is uniform across the continuum of possible linkages, the coefficients on the off-row-diagonal elements of $D*$ will lie between zero and twice the average upper bound. We therefore use twice the average upper bound as a critical value to assess violations of the null hypothesis and hence indicate the presence of special linkages suggestive of clustering activity.

In the normal situation, indirect value added will be less than the value of intermediate coefficients because of import leakages. Calculation of cases where indirect value added is more than the intermediate coefficients and then more than twice the average of available indirect value added would require the complementary existence of balancing lower values for other observations. Moreover, they would imply that country i has captured an abnormally large proportion of the available indirect value added arising from the chain of activities originating in sector k of country j. This would suggest that sector k of country j may well be connected an above average degree to a sector or sectors within country i. We would take this as evidence of transborder clustering. There could in fact be several *such supra-critical-value* entries for sector k of country j, suggesting the existence of multi-country transborder clustering.

9.4 The Evolution of Country Requirements for Imported Components

It was indicated at the beginning of this chapter that the manufactured goods component of the ICT sector was of interest to us because of the rapid evolution in its position as a major set of traded commodities. Multi-country input-output modeling can shed new light on the development trajectory.

Trade in ICT related goods (which includes electronics) is growing faster than overall merchandise trade, growing by an annual compound rate of 4 per cent between 1996 and 2002 (OECD 2004a: 66). Further, intra-industry trade is also on the rise (OECD 2004a, p. 71) indicating that there is both growing international specialization and integration occurring within the goods section of the ICT sector. However, measures of integration based on direct measures of intra-industry trade metrics are not measures of value chain interdependency because they do not include the full range of supplies necessary for production, while calculations based on multipliers tend to double count contributions. These limitations of the input-output methodology have arguably contributed to a lack of comparative analysis of the industries that require the highest level of imports to produce each additional unit of value added. The multi-regional input-output modeling reported on in this chapter provides a methodology for just such analysis.

Results are presented in the form of percentage changes in indirect value added as a result of imports from partner countries within the multi-regional model in successive rounds of intermediate purchases. The percentage changes are calculated relative to what could have been expected under a base case assumption in which no modeling of further inter-industry relationships were to occur, viz. calculations are consistent with the description of the matrix D^* in the previous section. Although individual bilateral import dependencies typically represent only a small share of value added, the relationships are highly concentrated. To understand the degree of specialization the percentage of indirect value added flowing to each trade partner is assessed for whether it is more than twice the average if all trade was spread evenly. This measure of significance is designed to highlight important links. The network charts also include those relationships which could be described as ‘second tier’ – representing between $1^1/_2$ and two times the average under the base case.

Based on a prior identification of national clusters (see previous discussion of identification of these), the top five clusters ranked by their trans-border linkages (contributions to indirect value added in trade partners) are highlighted in table 9.1.

The data in the table indicate the number of countries (out of a possible maximum of 9 for the first three columns of results, and out of a possible maximum of 20 countries for the final column). Table 9.1 has two sections; the first reveals the evolution of internationalized industries for the standard set of 9 OECD economies, while the last column provides the results from the full twenty country modeling using the year 2000 data. The table indicates, for example, that in 1970 all nine countries exhibited major cross-border linkages for the sector ‘Petroleum and Coal products’, indicating that significant above average transfer of indirect value added through imports to each other occurred for each of the nine OECD countries distinguished in the model.

Table 9.1. Occurrence of Cross-border Cluster Linkages in Top 5 national Clusters for Each Country – Ordered by 2000 List

Top Importing Clusters	**1970**	**1990**	**2000 (9) ‡**	**2000 (20)***
Petroleum & coal products	9	8	8	14
Office equip; Radio TV & comms	2 (of 14 series)	7 (of 14 series)	11 (of 17 series)	27 (of 38 series)
Motor vehicles	3 (of 8 series)	5 (of 8 series)	5	12
Aircraft	4	4	5	9
Non-ferrous metals	6	6	4	9
Iron and steel	4		3	8
Shipbuilding & repairing ships		1	3	3
TCF	5	5	2	2
Industrial chemicals	2	4		2
Other transport	3	2	2	2
Paper, paper prods & printing		1		
Wood products & furniture	2	1	1	2
Rubber & plastic products		1		4
Other manufacturing	2			
Mining & quarrying			1	1
Professional goods	1			
Food, beverages & tobacco	1			
Transport & storage	1			
Pharmaceuticals				2
Electrical Machinery				2
Computer services				1
Total	45 (9 * 5)	45 (9 * 5)	45 (9 * 5)	100 (20 * 5)

Note 1: ‡ modeling on the same basis as the 1970 and 1990 models. * Modeling utilizing the entire OECD 2002 Input-Output database.

Note 2: Not all countries provide data for all I-O classifications and thus there were a number of missing series – particularly in the earlier data (OECD 1996). In such cases the missing sectoral detail is typically included in another sector. To deal with data classification changes, the sector denoted 'Office equip and Radio TV & comms' is actually an amalgamation of two sectors for many countries, so that the number of series or 'country counts' for these combined sectors is larger than the individual number of countries.

Three key findings emerge from the evidence in table 9.1:

The demand by industrialized countries for imported petroleum is clear, in 2000 eight of the OECD 9 group have it within their top five intermediate imports (by the indirect value added indicator) but of the additional 11 countries in the expanded database only six have it within their top 5;

The rapid trend towards an increasingly internationalized ICT production system is very visible, with ICT represented in 50 per cent of its available linkage opportunities (1990) and 65 per cent of the available opportunities in 2000 (71 per cent in the 20 country series); and

The list includes a variety of high technology and low technology / resource processing industries.

The trend towards greater trade and growing international interdependence for the manufactured ICT goods is thus established in comparison to other sectors of the world economy. Table 9.2 digs deeper into the results for the ICT sector to report on the actual share of indirect value added flowing to trade partners in 1970, 1990 and 2000.

Table 9.2 ICT sector indirect value added flows to trade partners from the originating indicated importing country for 1970, 1990 and 2000.

Countries	Office 1970	Office 1990	Office 2000	Radio 1970	Radio 1990	Radio 2000
Australia	N/A	N/A	0.356	0.196	0.318	N/A
Canada	0.370	0.508	0.674	0.202	0.386	0.447
Denmark	N/A	N/A	0.317	0.335	0.366	0.385
France	0.418	0.308	0.453	0.091	0.180	0.242
Germany	0.152	0.184	0.297	N/A	N/A	0.276
Japan	0.101	0.097	0.173	0.092	0.093	0.102
Netherlands	0.386	0.438	0.524	N/A	N/A	0.466
UK	0.108	0.367	0.458	0.161	0.285	0.328
USA	0.050	0.172	0.253	0.052	0.126	0.137
Other countries in 2000 series			Office & computing 2000			Radio, TV & Comms Equip 2000
Brazil			0.264			0.000
China			0.350			0.295
Czech Rep			0.585			0.584
Finland			0.728			0.429
Greece			0.106			0.234
Hungary			0.769			0.710
Italy			0.475			0.196
Korea			0.510			0.372
Norway			0.375			0.333
Poland			0.417			0.301
Spain			0.375			0.351

It is universally the case that linkages measured by indirect value added flows were stronger for each country in 2000 than was the case in 1970. In general the production system for Radio, TV and communications equipment requires a smaller share of value added be transferred to imports than for office equipment. A number of countries in the year 2000 modeling are shown to have an especially strong linkage to imports and by inference to have substantial trans-national relationships in that the percentage of indirect value added that flows to imports is greater than 50 per cent. Of course, these countries may also have strong export linkages, for example, Finland, which known to be a major exporter of ICT goods (OECD 2004a). Typically, the larger economies naturally require fewer internationally traded imports. However, the spatial fragmentation of value chains can be ongoing *within* these countries[9].

---

[9] To better understand such processes it would be necessary to have access to regionalised input-output models for the larger geographic nations of the USA, France and Germany and in the future China.

9.5 The Changing Spatial Structure of ICT Inter-Cluster Networks 1970-2000

It is apparent from authors such as Bresnahan *et al.* (2001), Sturgeon (2003) and Ernst (2001), that computing and electronic products require an increasingly complex array of components. From WTO data (2005) on world export shares by regions and Wixted (2005) on the changing export specializations of East Asian economies, it is apparent that the source of those components is increasingly diverse. Table 9.3 reveals the extent of current intra and extra (multinational) regional trade in Office and Telecom equip. In this section we provide network flow diagrams based on direct import requirements.[10]

Table 9.3. Major regional flows in world exports of office and telecom equipment, 2004 (billion dollars).

Region	Flows
Intra-Asia	310.3
Intra-Europe	239.3
Asia to North America	156.2
Asia to Europe	126.0
Intra-North America	78.6
North America to Asia	52.1

Source: WTO 2005 Table IV.42.

Unfortunately, due to data limitations, this chapter cannot explore the evolution of the spatial structure of interdependencies with and between the East Asian economies.

9.5.1 Cluster Linkages 1970 – Nine Countries

Fig. 9.2 presents the national cluster linkages for the *Office, Accounting and Computing Machines* and the *Radio, TV and Communications Equipment* categories. The flows indicated are the flows of payments (goods move in the reverse direction) and principally represent the statistically significant flows.

---

[10] While the methodology outlined in Section 3 could be extended to the production of network flow diagrams based on indirect value added flows, this section concentrates on network analysis of more traditional measures of direct import dependence in order to provide independent support for previous conclusions.

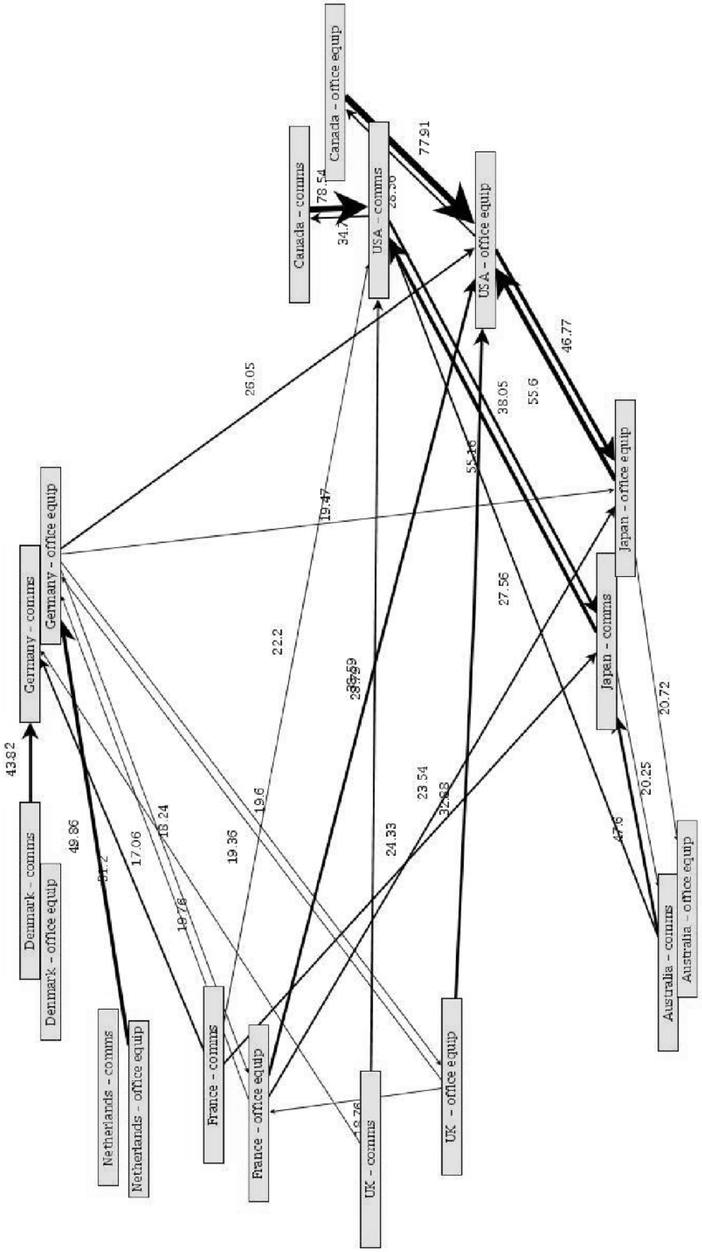


Fig. 9.2. Networks based on the share of OECD 9 imports (1970): ICT related manufacturing categories

In fig. 9.2, the arrows point in the direction of value added flows – goods flow in the opposite direction. The dotted lines represent 1.5-2 times the average of imports and solid lines are statistically significant at twice the average or above. The numbers represent the share of imports from the nine countries in the model. Australia and Denmark are not shown as purchasers of office equipment whilst the Germany and Netherlands is not shown as purchasing radio, TV and communications equipment - in these cases the purchases are included in the alternate category. This reflects the construction of the original OECD tables. In 1970 European countries were primarily dependent on Germany, the USA and Japan. Non-European countries (Australia, Canada, Japan and the USA) were not strongly dependent on any European countries.

In fig. 9.2, it can be clearly seen that the USA and Japan as early as 1970 had established a strong bilateral flow of intermediate goods for their respective office and computing machinery as well as their radio, TV and communications equipment clusters. Japan, perhaps surprisingly was moderately dependent on goods from Australia in both categories. Australia[11] principally imported from Japan and to a lesser extent the USA. Similarly, it is apparent that the USA relied on products from Canada and Canada on those from the USA (to a very major extent). Denmark (Communications) was primarily dependent on Germany for component imports as was the Netherlands for intermediate goods for office computing machines (recall the earlier remarks about the data limitations). Neither country was a major source of supplies. France bought from the USA, Japan and Germany whilst being a moderate supplier to Germany and the UK in office and computing products. Germany purchased mainly off the USA and to a lesser extent the UK and Japan for office machines. The United Kingdom purchased from Germany (moderate link) and France and the USA (stronger links) but notably not Japan.

The configuration in 1970 of the spatial organization of the *Office and computing machines* and the *Radio, TV and communications* production systems is interesting. Already, Japan is an important player with the strength of the USA clearly evident. The US-Canadian network is economically important to both and because of their geographic proximity is no surprise. The real surprise in the data is that Japan was dependent on Australia for supplies. Also of note is the point that whilst European countries require imports from the USA and Japan this is not reciprocated with purchases of European products (from Germany for example).

---

[11] The data for Australia was only reported in one industrial classification. Office and computing machines were amalgamated with Radio, TV and Communications Equipment for confidentiality reasons.

9.5.2 Cluster Linkages 2000 – Nine Countries

From the analysis of the 2000 dataset, the picture is in many ways similar, but there are important differences (see fig. 9.3). Some of the important conclusions from the 2000 model of nine countries include:

> Australia purchases supplies of components in the Office and Computing Machines and Radio, TV and Communications Equipment categories from the USA and Japan;
> Canada's Office and Computing Machine national cluster relies on the USA and Japan whilst Radio, TV and Communications Equipment is predominantly sourced from the USA;
> Denmark's Computer industry is dependent on Germany, the UK and the USA, whilst its Radio, TV and Communications industry relies on Germany and the UK for supplies;
> France's Computing activity is dependent on the USA, Japan and the UK (and interestingly not Germany), whilst its Radio, TV and Communications industry is purchased from the USA and Germany;
> For both its Computing and Radio, TV and Communications industries, Germany purchases from the USA and Japan;
> Japan requires industrial ingredients from the USA in both categories and no other countries in the 9 country component of the study;
> the Netherlands buys from the same three countries, the USA, Germany and the UK for both categories;
> the United Kingdom for Office and Computing Machines primarily imports from the USA and Germany whilst for Communications Equipment it uses USA, Germany and Japan.
> the USA buys from Canada and Japan.

It is worthwhile noting, again, that whilst European countries are dependent on Germany they also purchase from the USA and Japan. However, Australia, Canada, Japan and the USA still did not purchase from European countries at a level beyond the 1.5 * the average trade cut off.

9.5.3 Changes in the Spatial Structure of Interdependencies 1970 – 2000

Visually, two features that seem to be apparent from a comparison of the 1970 network structure and the 2000 network structure is that the lines are perhaps thinner and there looks to be more lines. This is partly because there are more series in the latter dataset as Denmark, Germany and the Netherlands do not have missing data. However, it is also because there are more links between national clusters. In both production systems there is one more link, even though Germany and Japan both *reduced* the number of their important trade dependencies.

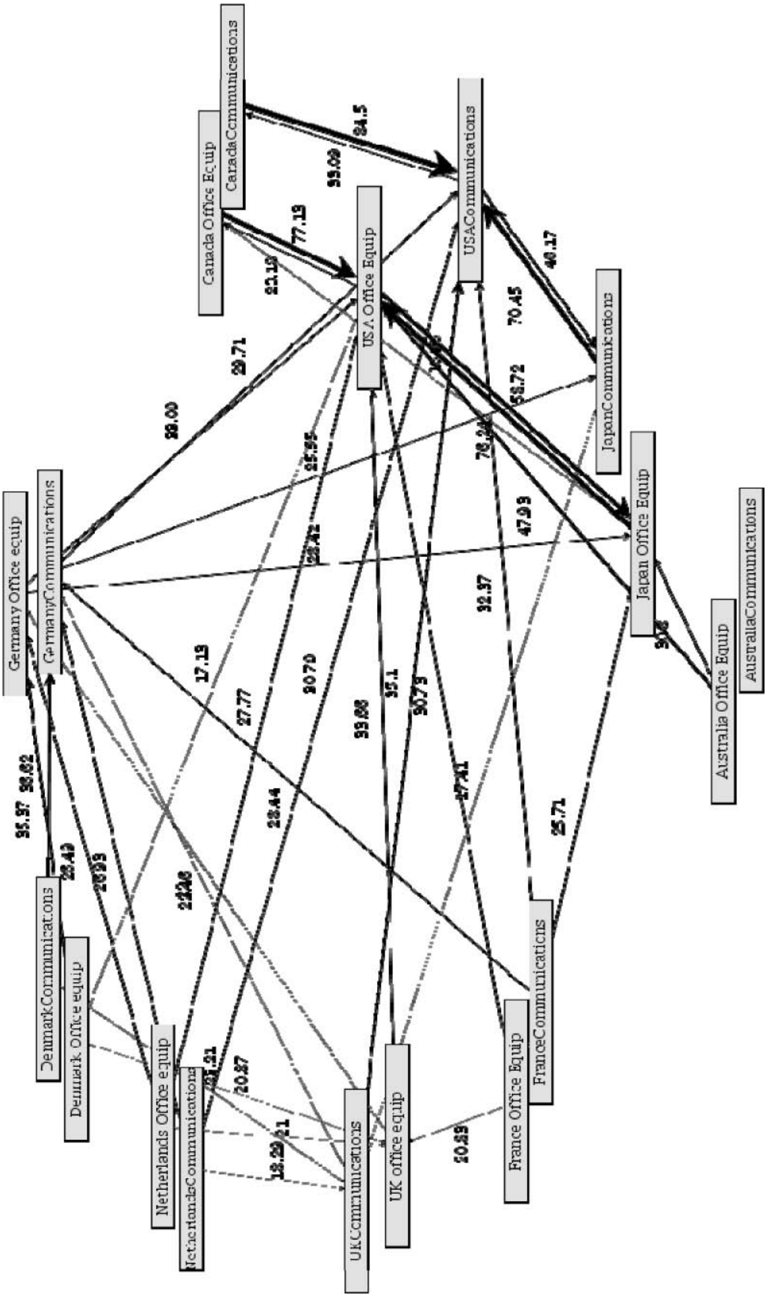


Fig. 9.3. Networks based on the share of OECD 9 imports (2000): ICT related manufacturing categories

Table 9.4. Import network (de)specialization: office & computing machines 1970-2000.

					Purchasers				
Suppliers	Aust	Can	Den	Fra	Deu	Jap	Neth	UK	USA
Australia		-0.2		-0.4	-0.9	-19.7	-0.2	-4.6	-3.2
Canada				-0.8	-1.6	-9.5	-0.2	-6.1	-22.7
Denmark		0		-0.2	-1.8	-0.2	-0.7	-1.3	-0.2
France		-0.3			-14.5	-2.1	-9.8	-14.3	-3.6
Germany		-0.9		-10.9		-2.8	-37.8	-9.6	-5.8
Japan		-6.5		-12.2	-7.9		-0.6	2.3	-32
Netherlands		-0.3		-0.2	-6.8	-0.6		-4.7	-1.5
UK		-1		-6.1	-13.3	-3.8	-4.3		-5.8
USA		-34.4		-25	-12.6	-36.8	-0.9	-18.2	

The other visual impression is also accurate. The strength of dependency on individual trade relationships has fallen, almost without exception in Office and Computing Machines and generally also for Radio, TV and Communications. Tables 9.4 and 9.5 present the change in trade dependency percentages between 1970 and 2000.

Table 9.5. Import network (de)specialization: radio, TV & communications 1970-2000.

	Aust	Can	Den	Fra	Deu	Jpn	Neth	UK	USA
Australia		-0.3	0	-0.2		-4.8		-1	-1.8
Canada			0.1	0.1		-2.4		-0.9	-6.5
Denmark		0.1		0		0		0	0
France		-0.4	-1.6			-0.2		-1.8	-1.2
Germany		0.8	-6.2	-3.3		0.3		0.9	-1.3
Japan		-6	-5.6	-5.9				0.5	-4.1
Netherlands		-0.3	1.7	0		0.1		-1.1	-0.5
UK		0	4.8	-0.4		-0.4			-1.8
USA		-2.9	0.7	2.9		0.5		2.8	

Note: The analysis for Australia is not available because in the 1970 data the Office Equipment category was excluded from the original database for confidentiality reasons and in the 2000 data the Radio, TV and Communications equipment category was excluded.

The numbers in tables 9.4 and 9.5 represent the change in the value of imports as a percentage of value added sourced from each country. Therefore, although the network structure remains largely intact, the degree of the reliance on many of those bilateral dependencies is diminishing.

Another approach to understanding the dependency of particular national clusters on others was outlined in section 9.3.2. That methodology functions beyond the calculation of the value of imports as a share of value added to assess whether particular clusters receive supra-normal gain from trade. This methodology proposes that by subtracting the original trade coefficient from the final I-O trade link and then assessing that difference against a nominal average for statistical significance, it is possible to better understand gains from trade. What we have called supra-critical links can emerge because input-output modeling calculates the flow of value added through an iterative approach to ensure that as each additional input is produced for the original increase of 1 unit of value added then those also need inputs etc. If inputs have a high import quotient it is possible to see that a particular country may gain more than what would be apparent from the intra industry trade coefficient. As it turns out, only one country indicated a supra-critical link in ICT related activities (table 9.6).

Table 9.6. Supra-critical values 1970-2000.

Supplier	Canada (Office & computing)	Canada (Radio, TV and Comms)
1970 USA	2.59	2.94
1970 Critical values	1.55	1.76
2000 Japan	1.30	
2000 Critical values	.88	

In 1970 the USA gained more from its exports in both ICT categories to Canada than would be suggested by the trade data alone. However, by 2000 Canada was only indicating one supra-critical link and this time to Japan. Interestingly, imports from Japan were much less important in absolute terms than those of the USA.

9.5.3 The Year 2000 – 20 Country Model of Interdependencies

Fig. 9.4, in a similar fashion to the earlier diagrams, maps the network structure of the expanded 20 country model. Because the 20 country dataset represents a larger share of total global economic activity within its system of bilateral relations, fig. 9.4 is calculated on a slightly different basis to figs. 9.2 and 9.3. In the earlier diagrams, the rest of the world was excluded from the calculation of the statistically significant trigger points (which doesn't change the shape but increases the number of visible linkages), whilst in fig. 9.4 it is included (minimizing the number of linkages).

Once again, the dotted lines represent import shares that are between 1.5 and two times the nominal average (5% of imports). The network analysis, whilst giving some support for the triad perspective on world trade reveals that, for ICT related products it is more accurately described as tri-polar. European countries particularly the EU expansion countries of the Czech Republic, Hungary and Poland are dependent on Germany. Beyond this they are dependent on the USA and to a lesser extent Japan (as are France Germany and Norway). Within the EU, France and the United Kingdom also seem to represent something of a second tier with a number of countries having a moderate (between 1.5 and 2 times the nominal average) system of customers.

The USA is dependent on Japan and to a moderate extent, China (but not Canada). Japan, as well as its connection to the USA is also moderately dependent on China (but not Korea). However, Korea is dependent on imports from Japan, the USA and China. The analysis of the twenty country model confirms and strengthens the understanding developed from the more limited OECD 9 analysis that the Asia-Pacific and North American economies are co-dependent but are not heavily reliant on European economies. Conversely Europe is dependent on Asia and the USA.

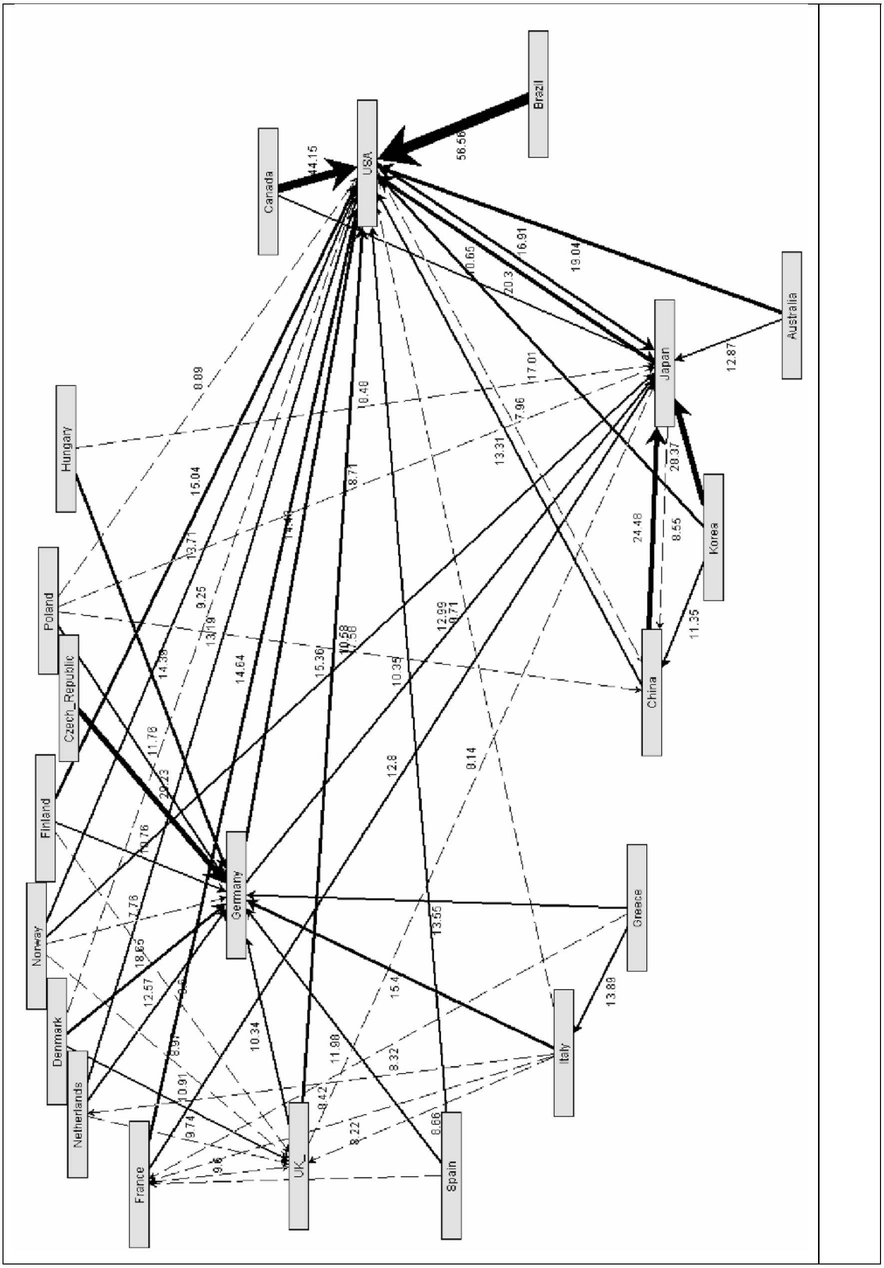


Fig. 9.4. Year 2000 twenty country model of office and computing equipment network interdependencies

At this point it is worthwhile reminding the reader that the special feature of these network presentations is that they are from the perspective of the importer (not the exporter). Thus Korea may be specialized in exports to the USA but its value in the USA is lower than that of China's.

9.6 Conclusions

In this chapter we have called the domestic national interdependencies around a particular industry a *national cluster*. We extended the analysis of interdependencies beyond national borders to determine the extent to which national clusters are linked to national clusters in other countries.

The modeling reveals that most national clusters are typically strongly connected (as a purchaser) to one or more other clusters. There were no cases where a cluster did not have enough of a specialized (concentrated) trading pattern to report at least one cluster-to-cluster supply relationship.

A number of important conclusions from this study of the evolution of cluster networking patterns over the period 1970-2000 for ICT related manufacturing emerge from the analysis. These include the conclusions that:

the core relationships for imports, even in rapidly progressing industries such as computing and electronics, remain through an extended period of time;

trade relationships are typically diversified, with new relationships typically added to existing relationships rather than replacing previous links. However, Germany and Japan both specialized their trade pattern within the nine country analysis (Japan shifted emphasis from Australia to China);

there is a tri-polar structure within ICT production centering on Germany, the USA and Japan, with the latter two important suppliers to Europe, but with no European country an important supplier outside of Europe;

in terms of import dependency (trade weight), for the OECD 9 countries between 1970 and 2000 all intra-9 bilateral relations in computing and many in electronics weakened across the period;

as the bilateral dependencies were weakening and diversifying, the value of imports as a share of value added rose significantly during the period;

dependency on the Rest of the World has increased from an average of 9 cents in the dollar of value added in office equip and seven cents for radio and TV equip in 1970 to 17.7 cents for Office equip and 15.6 cents for radio and TV equip in 2000); and

compared to other industries this rise in the dependency on imports is making ICT related manufacturing very import intensive.

The modeling approach adopted here is well suited to this form of analysis because it facilitates direct analysis of the percentage of value added that is required for imports. The methodology increases the value of input-output analysis as it provides results that are easily related to the trade and globalization literature.

This analysis provides new insights into the dynamics of both production fragmentation and clustering. The literature that addresses fragmentation to date has done so in a neo-classical welfare framework that emphasizes the role of improvement and cost reductions in service activities (e.g telecommunications) for reducing the transaction costs of splitting up production processes (see e.g. Jones and Kierzkowski (2001). It has not extensively considered the growing component modularity and product complexity, as seen in ICT merchandise.

Our results point to the need to extend this approach to work towards a more hierarchical understanding of the structure of clusters. As the concept of clustering is more typically applied to regional agglomerations rather than the national activities reported here, a key future requirement will be to apply this methodology on a within-country multi-regional basis.

References

Amin A, Thrift N (1992) Neo-Marshallian nodes in global networks. International Journal of Urban and Regional Research 16:571-587

Audretsch D, Feldman M (1996) R&D spillovers and the geography of innovation and production. American Economic Review 86:630-640

Bathelt H, Malmberg A, Maskell P (2004) Clusters and knowledge, local buzz, global pipelines and the process of knowledge creation. Progress in Human Geography 28:31–56

Beaverstock JV, Smith RG, Taylor PG (1999) A roster of world cities. Cities 16:445–458

Breschi S, Malerba F (2001) The geography of innovation and economic clustering: Some introductory notes. Industrial and Corporate Change 10:817-833

Breschi S, Lissoni F (2001) Knowledge Spillovers and Local innovation Systems, A Critical Survey. Industrial and Corporate Change 10:975-1005

Bresnahan T, Gambardella A, and Saxenian A (2001) Old economy inputs for the “new economy” outcomes, cluster formation in the new silicon valleys. Industrial and Corporate Change 10:835-860

Cooper RJ (2000) An extension of the block spatial path approach to analysis of the influence of intra- and interregional trade on multiplier effects in general multiregional input-output models. In: Reggiani A (ed) Spatial economic science, new frontiers in theory and methodology. Springer, Berlin

DeBresson C, Hu X, Drejer I, Lundvall B-Å (1998) Innovative activity in the learning economy - a comparison of systems in 10 OECD countries (draft report). OECD Paris

Ernst D, Linsu K (2002) Global production networks, knowledge diffusion and local capability formation. *Research Policy*, special issue in honour of Richard Nelson and Sydney Winter 31:1417-1429

Feenstra R, Lipsey T, Deng H, Ma A, Mo H (2005) NBER-United Nations trade data, 1962-2000 downloaded from http,//cid.econ.ucdavis.edu/data/undata/undata.html

Heston A, Summers R, Aten B (2002) Penn world table version 6.1. Center for International Comparisons (CICUP). at the University of Pennsylvania

Jones RW, Kierzkowski H (2001) A framework for fragmentation. In: S. Arndt and H. Kierzkowski (eds) Fragmentation: New production patterns in the world economy. Oxford University Press, New York

Leamer E, Storper M (2001) The economic geography of the internet age. Journal of International Business Studies 32:641-665

Malmberg A, Maskell P (2002) The elusive concept of localization economies, towards a knowledge-based theory of spatial clustering. Environment and Planning A 34:429-449

Maskell P (2001) Towards a knowledge-based theory of the geographical cluster. Industrial and Corporate Change 10:919-941

Maskell P, Lorenzen M (2004) The cluster as market organization. Urban Studies 41:991-1009

OECD (1996) Input-output database. OECD, Paris

OECD (1999) Boosting innovation, The cluster approach. OECD, Paris

OECD (2001c) Innovative clusters, drivers of national innovation systems. OECD, Paris

OECD (2002) [publicly released 2005] OECD Input-Output Database. OECD, Paris

OECD (2004a). OECD Information Technology Outlook. OECD, Paris

OECD (2004b). OECD Bilateral trade database 2004. OECD, Paris

OECD (2005). PPPs for GDP - Historical series. OECD, Paris http,//www.oecd.org/department/0,2688,en_2649_34347_1_1_1_1_1,00.html

Sassen S (ed) (2002) Global Networks / Linked Cities. Routledge, New York and London

Saxenian A (1994) Regional Advantage, Culture and Competition in Silicon Valley and Route 128. Harvard University Press, Cambridge MA

Saxenian A, Hsu J-Y (2001) The Silicon Valley–Hsinchu Connection, Technical Communities and Industrial Upgrading. Industrial and Corporate Change 10:893-920

Simmie J (2002) Trading places, competitive cities in the global economy. European Planning Studies 10:201-214

Simmie J. (2004) Innovation and clustering in the globalised international economy. Urban Studies 41:1095-1112

Simmie J, Sennet J, Wood P, Hart D (2002) Innovation in europe, a tale of networks, knowledge and trade in five cities. Regional Studies 36:47–64

Smith D, Timberlake M (2001) World city networks and hierarchies, 1977-1997. An empirical analysis of global air travel links. American Behavioral Scientist 44:1656-1678

Steinle C, Schiele H (2002) When do industries cluster? A proposal on how to assess an industry's propensity to concentrate in a single region or nation. Research Policy 31:849-858

Storper M (1995) The resurgence of regional economies, ten years later, the region as a nexus of untraded interdependencies. European Urban and Regional Studies 2:191-221

Storper M, Venables AJ (2004). Buzz, Face-to-Face Contact and the Urban Economy. Journal of Economic Geography 4:351-370

UNCTAD International Trade Centre (2003) International Trade Statistics http,//www.intracen.org/tradstat/welcome.htm

von Hippel E (1988) The sources of innovation. Oxford University Press, New York

Wixted B (2005) Systems of innovation beyond borders, linked clusters and the role, scale and spatial structure of extra-territorial interdependencies. Unpublished Ph.D. thesis – URL http,//adt.caul.edu.au/

Wolfe DA, Gertler MS (2004) Clusters from the inside and out, local dynamics and global linkages. Urban Studies 41:1071-1093

World Bank (2001) Trade and Production Database 1976-1999. World Bank, Washington http,//web.worldbank.org/WBSITE/EXTERNAL/TOPICS/TRADE/0,,contentMDK,20103741~menuPK,167374~pagePK,148956~piPK,216618~theSitePK,239071,00.html

WTO (2005) International trade statistics 2005. World Trade Organization, Geneva

Wyckoff A (1993) The Extension of Networks of Production across Borders. OECD STI Review 13:61-87

10 The Co-Evolution and Emergence of Integrated International Financial Networks and Social Networks: Theory, Analysis, and Computations

Anna Nagurney, Jose Cruz, and Tina Wakolbinger

Isenberg School of Management, University of Massachusetts, Amherst MA USA

10.1 Introduction

Globalization and technological advances have made major impacts on financial services in recent years and have allowed for the emergence of electronic finance. Indeed, the financial landscape has been transformed through increased financial integration, increased cross border mergers, and lower barriers between markets. In addition, boundaries between different financial intermediaries have become less clear (cf. Claessens, Glaessner, and Klingebiel (2000, 2001), Claessens (2003), Claessens et al. (2003), G-10 (2001)).

During the period 1980-1990, global capital transactions tripled with telecommunication networks and financial instrument innovation being two of the empirically identified major causes of globalization with regards to international financial markets (Kim (1999)). The growing importance of networks in financial services and their effects on competition have been also addressed by Claessens et al. (2003). Kim (1999), in particular, has argued for the necessity of integrating various theories, including portfolio theory with risk management, and flow theory in order to capture the underlying complexity of the financial flows over space and time.

At the same time that globalization and technological advances have transformed financial services, researchers have identified the importance of social networks in a plethora of financial transactions (cf. Sharpe (1990), Uzzi (1997, 1999), Anthony (1997), Arrow (1998), DiMaggio and Louch (1998), Ghatak (2002)), notably, in the context of personal relationships. Nevertheless, the relevance of social networks within an international financial context has yet to be examined theoretically (or empirically). Clearly, the existence of appropriate social networks can affect not only the risk associated with financial transactions but also transaction costs.

Given the prevalence of networks (be they in the form of telecommunication networks, social networks, as well as the foundational financial networks; cf. Nagurney and Siokos (1997), Nagurney (2003), and the references therein) in the discussions of globalization and international financial flows, it seems natural that any theory for the illumination of the behavior of the decision-makers involved in this context as well as the impacts of their decisions on the financial product flows, prices, appreciation rates, etc., should be network-based. In this paper, hence, we take on a network perspective for the theoretical modeling, analysis, and computation of solutions to international financial networks with intermediation in which we explicitly integrate the social network component. We also capture electronic transactions within our framework since this aspect is critical in the modeling of international financial flows today.

In particular, in this paper, we focus on the development of a *supernetwork* framework for the integration of social networks with international financial networks with intermediation and electronic transactions. Supernetwork theory (cf. Nagurney and Dong (2002) and the references therein) has been used, to-date, to study a variety of network-based applications in which humans interact on two or more networks (very often transportation and telecommunication networks). Applications that have been formulated and solved using this approach include: telecommuting versus commuting decision-making, supply chains with electronic commerce, power/energy networks, as well as knowledge networks.

In addition, in this paper, we build upon the recent work of Nagurney and Cruz (2003, 2004) in the development of international financial network models (static and dynamic) and that of Nagurney, Wakolbinger, and Zhao (2004) and Wakolbinger and Nagurney (2004) in the integration of social networks with other economic networks.

This paper is organized as follows. In Section 10.2, we develop the multilevel supernetwork model consisting of multiple tiers of decision-makers acting on the international financial network with intermediation and the social network. We describe the decision-makers' optimizing behavior, and establish the governing equilibrium conditions along with the corresponding variational inequality formulation.

In Section 10.3, we describe the disequilibrium dynamics of the international financial flows, the prices, and the relationship levels as they co-evolve over time and formulate the dynamics as a projected dynamical system (cf. Nagurney and Zhang (1996a, b), Nagurney and Ke (2003), Nagurney and Cruz (2004), and Nagurney, Wakolbinger, and Zhao (2004)). We establish that the set of stationary points of the projected dynamical system coincides with the set of solutions to the derived variational inequality problem.

In Section 10.4, we present a discrete-time algorithm to approximate (and track) the international financial flow, price, and relationship level trajectories over time until the equilibrium values are reached. We then apply the discrete-time algorithm in Section 10.5 to several numerical examples to further illustrate the supernetwork model. We conclude with Section 10.6, in which we summarize our results and suggest possibilities for future research.

10.2 The Supernetwork Model Integrating International Financial Networks with Intermediation and Social Networks

In this Section, we develop the supernetwork model consisting of the integration of the international financial network with intermediation and the social network in which the decision-makers are those with sources of funds, the financial intermediaries, as well as the consumers associated with the demand markets. Here we describe the model in an equilibrium context, whereas in Section 3, we provide the disequilibrium dynamics and the co-evolution of the international financial flows, the prices, as well as the relationship levels between tiers of decision-makers over time. This model generalizes the model of Nagurney and Cruz (2003) to explicitly include social networks. In addition, it broadens the framework proposed in Nagurney, Wakolbinger, and Zhao (2004) to the international dimension.

As in the model of Nagurney and Cruz (2003), the model consists of L countries, with a typical country denoted by l or $\hat{l}$; I "source" agents in each country with sources of funds, with a typical source agent denoted by i, and J financial intermediaries with a typical financial intermediary denoted by j. Examples of source agents are households and businesses, whereas examples of financial intermediaries include banks, insurance companies, investment companies, brokers, including electronic brokers, etc. Intermediaries in our framework need not be country-specific but, rather, may be virtual.

We assume that each source agent can transact directly electronically with the consumers through the Internet and can also conduct his financial transactions with the intermediaries either physically or electronically in different currencies. There are H currencies in the international economy, with a typical currency being denoted by h. Also, we assume that there are K financial products which can be in distinct currencies and in different countries with a typical financial product (and associated with a demand market) being denoted by k. Hence, the financial intermediaries in the model, in addition to transacting with the source agents, also determine how to allocate the incoming financial resources among distinct uses, which are represented by the demand markets with a demand market corresponding to, for example, the market for real estate loans, household loans, or business loans, etc., which, as mentioned, can be associated with a distinct country and a distinct currency combination. We let m refer to a mode of transaction with $m = 1$ denoting a physical transaction and $m = 2$ denoting an electronic transaction via the Internet.

The depiction of the supernetwork is given in Figure 10.1. As this figure illustrates, the supernetwork is comprised of the social network, which is the bottom level network, and the international financial network, which is the top level network. Internet links to denote the possibility of electronic financial transactions are denoted in the figure by dotted arcs. In addition, dotted arcs/links are used to depict the integration of the two networks into a supernetwork. Examples of other supernetworks can be found in Nagurney and Dong (2002). Subsequently, we describe the interrelationships between the financial network and the social network through the functional forms and the flows on the links.

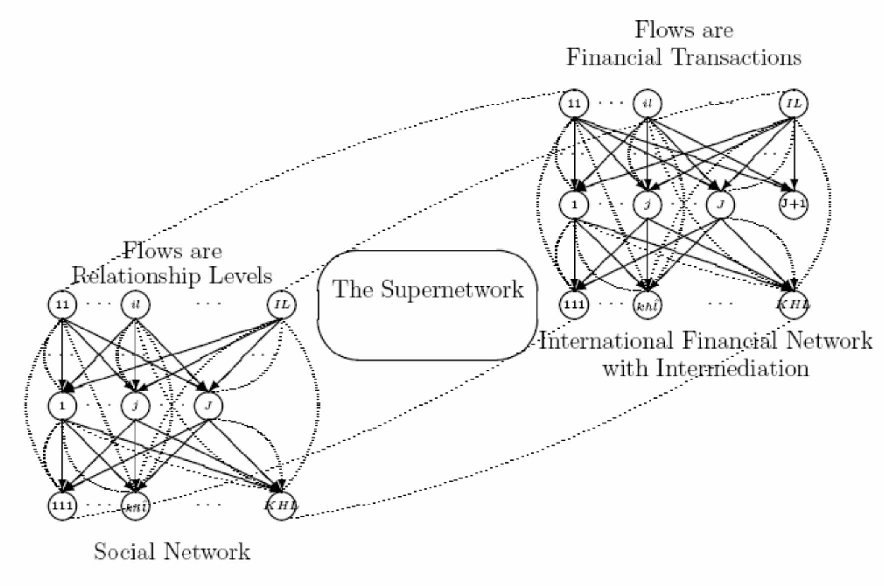


Fig. 10.1. The Multilevel Supernetwork Structure of the Integrated International Financial Network / Social Network System

The supernetwork in Figure 10.1 consists of a social and an international financial network with intermediation. Both networks consist of three tiers of decision-makers. The top tier of nodes consists of the agents in the different countries with sources of funds, with agent *i* in country *l* being referred to as agent *il* and associated with node *il*. There are, hence, *IL* top-tiered nodes in the network. The middle tier of nodes in each of the two networks consists of the intermediaries (which need not be country-specific), with a typical intermediary *j* associated with node *j* in this (second) tier of nodes in the networks. The bottom tier of nodes in both the social network and in the financial network consists of the demand markets, with a typical demand market for product *k* in currency *h* and country $\hat{l}$ associated with node $kh\hat{l}$. There are, as depicted in Figure 10.1, *J* middle (or second) tiered nodes corresponding to the intermediaries and *KHL* bottom (or third) tiered nodes in the international financial network. In addition, we add a node *J* + 1 to the middle tier of nodes in the financial network only in order to represent the possible non-investment (of a portion or all of the funds) by one or more of the source agents (see also Nagurney and Ke (2003)).

The supernetwork in Figure 10.1 includes classical physical links as well as Internet links to allow for electronic financial transactions. Electronic transactions are possible between the source agents and the intermediaries, the source agents and the demand markets as well as the intermediaries and the demand markets. Physical transactions can occur between the source agents and the intermediaries and between the intermediaries and the demand markets. We now turn to the description of the behavior of the various economic decision-makers, i.e., the source agents, the financial intermediaries, and the demand markets.

10.2.1 The Behavior of the Agents with Sources of Funds and Their Optimality Conditions

As described in Nagurney and Cruz (2003), we assume that each agent *i* in country *l* has an amount of funds S^{il} available in the base currency. Since there are assumed to be *H* currencies and 2 modes of transaction (physical or electronic), there are *2H* links joining each top tier node *il* with each middle tier node *j; j=1,...,J*, with the first *H* links representing physical transactions between a source and intermediary, and with the corresponding flow on such a link given, respectively, by x^{il}_{jh1}, and the subsequent *H* links representing electronic transactions with the corresponding flow given, respectively, by x^{il}_{jh2}. Hence, x^{il}_{jh1} denotes the nonnegative amount invested (across all financial instruments) by source agent *i* in country *l* in currency *h* transacted through intermediary *j* using the physical mode whereas x^{il}_{jh2} denotes the analogue but for an electronic transaction. We group the financial flows for all source agents/countries/intermediaries/currencies/modes into the column vector $x^1 \in R^{2ILJH}_{+}$. In addition, a source agent *i* in country *l* may transact directly with the consumers at demand market *k* in currency *h* and country $\hat{l}$ via an Internet link. The nonnegative flow on such a link joining node *il* with node $kh\hat{l}$ is denoted by $x^{il}_{kh\hat{l}}$. We group all such financial flows, in turn, into the column vector $x^2 \in R^{ILKHL}_{+}$. Also, we let x^{il} denote the *2JH+KHL*-dimensional column vector associated with source agent *il* with components: $\{x^{il}_{jhm}, x^{il}_{kh\hat{l}}; j=1,\dots,J; h=1,\dots,H; m=1,2; k=1,\dots,K; \hat{l}=1,\dots,L\}$. Furthermore, we construct a link from each top tiered node to the second tiered node *J+1* and associate a flow s^{il} on such a link emanating from node *il* to represent the possible nonnegative amount not invested by agent *i* in country *l*.

For each source agent *il* the amount of funds transacted either electronically and/or physically cannot exceed his financial holdings. Hence, the following conservation of flow equation must hold:

$$\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2} x^{il}_{jhm} + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} x^{il}_{kh\hat{l}} \leq S^{il}, \quad \forall i,l. \tag{10.1}$$

In Figure 10.1 the slack associated with constraint (10.1) for source agent *i* in country *l* is represented as the flow on the link joining node *il* with the non-investment node *J+1*.

Furthermore, let η^{il}_{jhm} denote the nonnegative level of the relationship between source agent *il* and intermediary *j* associated with currency *h* and mode of transaction *m* and let $\eta^{il}_{kh\hat{l}}$ denote the nonnegative relationship level associated with the virtual mode of transaction between source agent *il* and "demand market" $kh\hat{l}$. Each source agent *il* may actively try to achieve a certain relationship level with an intermediary and/or a demand market. We group the η^{il}_{jhm}s for all source agent/country/intermediary/currency/mode combinations into the column vector $\eta^1 \in R^{2ILJH}_{+}$ and the $\eta^{il}_{kh\hat{l}}$s for all the source agent/country/demand market/currency/country combinations into the column vector $\eta^2 \in R^{ILKHL}_{+}$. Moreover, we assume that these relationship levels take on a value that lies in the range [0, 1]. No relationship is indicated by a relationship level of zero and the strongest possible relationship is indicated by a relationship level of one. In the supernetwork depicted in Figure 10.1 the relationship flows are associated with the links in the social network component of the supernetwork. Specifically, the vector of flows η^1 corresponds to flows on the links between the source agents in the countries and the intermediaries, whereas the vector of flows η^2 corresponds to the flows on the links between the source agents and the demand markets in the various currencies and countries on the social network. The relationship levels, along with the financial flows, are endogenously determined in the model.

The source agent may spend money, for example, in the form of gifts and/or additional time/service in order to achieve a particular relationship level. The production cost functions for relationship levels are denoted by b^{il}_{jhm} and $b^{il}_{kh\hat{l}}$ and represent, respectively, how much money a source agent *il* has to spend in order to achieve a certain relationship level with intermediary *j* transacting through mode *m* and currency *h* or in order to achieve a certain relationship level with demand market $kh\hat{l}$. These relationship production cost functions are distinct for each such combination. Their specific functional forms may be influenced by such factors as the willingness of intermediaries or demand markets to establish/maintain a relationship and the level of previous business relationships and private relationships that exist. In an international setting, in particular, cultural differences, difficulties with languages, and distances, may also play a role in making it more costly to establish (and maintain) a relationship level.

The relationship production cost function is assumed, hence, to be a function of the relationship level between the source agent *il* and intermediary *j* transacting via mode *m* and currency *h* or with the consumers at demand market $kh\hat{l}$, that is,

$$b^{il}_{jhm} = b^{il}_{jhm}(\eta^{il}_{jhm}), \quad \forall i,l,j,h,m, \tag{10.2}$$

$$b^{il}_{kh\hat{l}} = b^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}}), \quad \forall i,l,k,h,\hat{l}. \tag{10.3}$$

We assume that these functions are convex and continuously differentiable. We denote the transaction cost associated with source agent *il* transacting with intermediary j in currency h via mode m by c^{il}_{jhm} (and measured in the base currency) and assume that:

$$c^{il}_{jhm} = c^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}), \quad \forall i, l, j, h, m, \tag{10.4}$$

that is, the cost associated with source agent i in country l transacting with intermediary j in currency h depends on the volume of financial transactions between the particular pair via the particular mode and on the relationship level between them. If the relationship level increases, the transaction cost may be expected to decrease since higher relationship levels may lead to higher levels of trust, which can affect transactions. This is especially important in international exchanges in which transaction costs may be significant.

We denote the transaction cost associated with source agent il transacting with demand market k in country $\hat{l}$ in currency h via the Internet link by $c^{il}_{kh\hat{l}}$ (and also measured in the base currency) and assume that:

$$c^{il}_{kh\hat{l}} = c^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}}), \quad \forall i, l, k, h, \hat{l}, \tag{10.5}$$

that is, the cost associated with source agent i in country l transacting with the consumers for financial product k in currency h and country $\hat{l}$. The transaction cost functions are assumed to be convex and continuously differentiable and depend on the volume of flow of the transaction and on the relationship level.

The source agent il faces total costs that equal the sum of the total transaction costs plus the costs that he incurs for establishing relationship levels. His revenue, in turn, is equal to the sum of the price (rate of return plus the rate of appreciation) that the agent can obtain times the total quantity obtained/purchased. Let now ρ^{il*}_{1jhm} denote the actual price charged agent il for the instrument in currency h by intermediary j transacting via mode m and let $\rho^{il*}_{1kh\hat{l}}$, in turn, denote the actual price associated with source agent il transacting electronically with demand market $kh\hat{l}$. Let e^*_h denote the actual rate of appreciation of currency h against the base currency, which can be interpreted as the rate of return earned due to exchange rate fluctuations (see Nagurney and Siokos (1997)). These ``exchange" rates are grouped into the column vector $e^* \in R^H_+$. We later discuss how such prices are recovered.

We assume that each source agent in each country seeks to maximize his net return with the net revenue maximization problem for source agent il being given by:

$$\text{Maximize} \sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}(\rho^{il*}_{1jhm} + e^*_h)x^{il}_{jhm} + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}(\rho^{il*}_{1khl} + e^*_h)x^{il}_{kh\hat{l}}$$

$$-\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}c_{jhm}^{il}(x_{jhm}^{il},\eta_{jhm}^{il})-\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}c_{kh\hat{l}}^{il}(x_{kh\hat{l}}^{il},\eta_{kh\hat{l}}^{il})$$

$$-\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}b_{jhm}^{il}(\eta_{jhm}^{il})-\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}b_{kh\hat{l}}^{il}(\eta_{kh\hat{l}}^{il}) \tag{10.6}$$

subject to:

$$x_{jhm}^{il}\geq 0,\quad x_{kh\hat{l}}^{il}\geq 0,\quad \forall j,h,m,k,\hat{l}, \tag{10.7}$$

$$0\leq \eta_{jhm}^{il}\leq 1,\quad 0\leq \eta_{kh\hat{l}}^{il}\leq 1,\quad \forall j,h,m,k,\hat{l}, \tag{10.8}$$

and the constraint (10.1) for source agent *il*.

The first two terms in (10.6) represent the revenue. The next four terms represent the various costs. The constraints are comprised of the conservation of flow equation (10.1), the nonnegativity assumptions on the financial flows (cf. (10.7)) and on the relationship levels with the latter being bounded from above by the value one (cf. (10.8)).

Furthermore, it is reasonable to assume that each source agent tries to minimize risk and has been noted in empirical studies by Kim (1999). Here, for the sake of generality, and as in the papers by Nagurney and Cruz (2003, 2004) and Nagurney, Wakolbinger, and Zhao (2004), we assume, as given, a risk function for source agent *il*, dealing with intermediary *j* via mode *m* and currency *h*, denoted by r_{jhm}^{il}, and a risk function for source agent *il* dealing with demand market $kh\hat{l}$ denoted by $r_{kh\hat{l}}^{il}$. These functions depend not only on the quantity of the financial flow transacted between the pair of nodes (and via a particular currency and mode) but also on the corresponding relationship level. If the relationship level increases, the risk is likely to decrease because trust reduces transaction uncertainty. Since international financial transactions are potentially riskier, high relationship levels can be of utmost significance and can, hence, create competitive advantages.

These risk functions are assumed to be as follows:

$$r_{jhm}^{il}=r_{jhm}^{il}(x_{jhm}^{il},\eta_{jhm}^{il}),\quad \forall i,l,j,h,m, \tag{10.9}$$

$$r_{kh\hat{l}}^{il}=r_{kh\hat{l}}^{il}(x_{kh\hat{l}}^{il},\eta_{kh\hat{l}}^{il}),\quad \forall i,l,k,h,\hat{l}, \tag{10.10}$$

where r_{jhm}^{il} and $r_{kh\hat{l}}^{il}$ are assumed to be convex and continuously differentiable.

Hence, source agent *il* also faces an optimization problem associated with his desire to minimize the total risk and corresponding to:

$$\text{Minimize}\ \sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}r_{jhm}^{il}(x_{jhm}^{il},\eta_{jhm}^{il})+\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}r_{kh\hat{l}}^{il}(x_{kh\hat{l}}^{il},\eta_{kh\hat{l}}^{il}) \tag{10.11}$$

subject to:

$$x^{il}_{jhm} \geq 0, \quad x^{il}_{kh\hat{l}} \geq 0, \quad \forall j,h,m,k,\hat{l}, \tag{10.12}$$

$$0 \leq \eta^{il}_{jhm} \leq 1, \quad 0 \leq \eta^{il}_{kh\hat{l}} \leq 1, \quad \forall j,h,m,k,\hat{l}. \tag{10.13}$$

In addition, the source agent also tries to maximize the relationship value generated by interacting with other decision-makers in the network. Here, v^{il}_{jhm} denotes the relationship value function for source agent *il*, intermediary *j*, mode *m* and currency *h*, and v^{il}_{jhm} is assumed to be a function of the relationship level of *il* with intermediary *j* transacting via mode *m* and currency *h*. Similarly, $v^{il}_{kh\hat{l}}$ denotes the relationship value function for source agent *il* and demand market $kh\hat{l}$. It is assumed to be a function of the relationship level with the particular demand market $kh\hat{l}$ such that

$$v^{il}_{jhm} = v^{il}_{jhm}(\eta^{il}_{jhm}), \quad \forall i,l,j,h,m, \tag{10.14}$$

$$v^{il}_{kh\hat{l}} = v^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}}), \quad \forall i,l,k,h,\hat{l}. \tag{10.15}$$

We assume that the value functions are continuously differentiable and concave.

Hence, source agent *il* is also faced with the optimization problem representing the maximization of the total value of his relationships expressed mathematically as:

$$\text{Maximize} \quad \sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2} v^{il}_{jhm}(\eta^{il}_{jhm}) + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} v^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}}) \tag{10.16}$$

subject to:

$$0 \leq \eta^{il}_{jhm} \leq 1, \quad 0 \leq \eta^{il}_{kh\hat{l}} \leq 1, \quad \forall j,h,m,k,\hat{l}. \tag{10.17}$$

10.2.2 The Multicriteria Decision-Making Problem Faced by a Source Agent

We can now construct the multicriteria decision-making problem facing a source agent, which allows him to weight the criteria of net revenue maximization (cf. (10.6)), risk minimization (cf. (10.11)), and total relationship value maximization (see (10.16)) in an individual manner. Source agent *il*'s multicriteria decision-making objective function is denoted by U^{il}. Assume that source agent *il* assigns a nonnegative weight α^{il} to the risk generated and a nonnegative weight β^{il} to the relationship value. The weight associated with net revenue maximization serves as the numeraire and is set equal to 1. The nonnegative weights measure the importance of risk and the total relationship value and, in addition, transform these values into monetary units. We can now construct a value function for each source agent (cf. Keeney and Raiffa (1993), Dong, Zhang, and Nagurney (2002), Nagurney, Wakolbinger, and Zhao (2004), and the references therein) using a constant additive weight value function. Therefore, the multicriteria decision-making problem of source agent *il* can be expressed as:

$$\begin{aligned}
\text{Maximize } U^{il} = {} & \sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}(\rho^{il*}_{1jhm} + e^{*}_{h})x^{il}_{jhm} + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}(\rho^{il*}_{1khl} + e^{*}_{h})x^{il}_{kh\hat{l}} \\
& - \sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2} c^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}) - \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} c^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}}) \\
& - \sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2} b^{il}_{jhm}(\eta^{il}_{jhm}) - \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} b^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}}) \\
& - \alpha^{il}\Big(\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2} r^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}) + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} r^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}})\Big) \\
& + \beta^{il}\Big(\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2} v^{il}_{jhm}(\eta^{il}_{jhm}) + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} v^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}})\Big)
\end{aligned} \tag{10.18}$$

subject to:

$$x^{il}_{jhm} \geq 0, \quad x^{il}_{kh\hat{l}} \geq 0, \quad \forall j, h, m, k, \hat{l}, \tag{10.19}$$

$$0 \leq \eta^{il}_{jhm} \leq 1, \quad 0 \leq \eta^{il}_{kh\hat{l}} \leq 1, \quad \forall j, h, m, k, \hat{l}, \tag{10.20}$$

and the constraint (10.1) for source agent *il*.

The first six terms on the right-hand side of the equal sign in (10.18) represent the net revenue which is to be maximized, the next two terms represent the weighted total risk which is to be minimized and the last two terms represent the weighted total relationship value, which is to be maximized. We can observe that such an objective function is in concert with those used in classical portfolio optimization (see Markowitz (1952, 1959)) but substantially more general to reflect specifically the additional criteria, notably, that of total relationship value maximization.

Under the above assumed and imposed assumptions on the underlying functions, the optimality conditions for all source agents simultaneously can be expressed as the following inequality (cf. Bazaraa, Sherali, and Shetty (1993), Gabay and Moulin (1980); see also Nagurney (1999)):

determine $(x^{1*}, x^{2*}, \eta^{1*}, \eta^{2*}) \in \mathbf{K}^{1}$, satisfying

$$\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}\left[\alpha^{il}\frac{\partial r^{il}_{jhm}(x^{il*}_{jhm}, \eta^{il*}_{jhm})}{\partial x^{il}_{jhm}} + \frac{\partial c^{il}_{jhm}(x^{il*}_{jhm}, \eta^{il*}_{jhm})}{\partial x^{il}_{jhm}} - \rho^{il*}_{1jhm} - e^{*}_{h}\right]$$

$$\times \left[x^{il}_{jhm} - x^{il*}_{jhm} \right]$$

$$+ \sum_{i=1}^{I} \sum_{l=1}^{L} \sum_{j=1}^{J} \sum_{h=1}^{H} \sum_{\hat{l}=1}^{L} \left[\alpha^{il} \frac{\partial r^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}}, \eta^{il*}_{kh\hat{l}})}{\partial x^{il}_{kh\hat{l}}} + \frac{\partial c^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}}, \eta^{il*}_{kh\hat{l}})}{\partial x^{il}_{kh\hat{l}}} - \rho^{il*}_{1kh\hat{l}} - e^{*}_{h} \right]$$

$$\times \left[x^{il}_{kh\hat{l}} - x^{il*}_{kh\hat{l}} \right]$$

$$+ \sum_{i=1}^{I} \sum_{l=1}^{L} \sum_{j=1}^{J} \sum_{h=1}^{H} \sum_{m=1}^{2} \left[\frac{\partial c^{il}_{jhm}(x^{il*}_{jhm}, \eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}} + \frac{\partial b^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}} - \beta^{il} \frac{\partial v^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}} \right.$$

$$\left. + \alpha^{il} \frac{\partial r^{il}_{jhm}(x^{il*}_{jhm}, \eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}} \right] \times \left[\eta^{il}_{jhm} - \eta^{il*}_{jhm} \right]$$

$$+ \sum_{i=1}^{I} \sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{h=1}^{H} \sum_{\hat{l}=1}^{L} \left[\frac{\partial c^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}}, \eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} + \frac{\partial b^{il}_{kh\hat{l}}(\eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} - \beta^{il} \frac{\partial v^{il}_{kh\hat{l}}(\eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} \right.$$

$$\left. + \alpha^{il} \frac{\partial r^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}}, \eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} \right] \times \left[\eta^{il}_{kh\hat{l}} - \eta^{il*}_{kh\hat{l}} \right] \geq 0,$$

$$\forall (x^1, x^2, \eta^1, \eta^2) \in \mathbf{K}^1, \qquad (10.21)$$

where

$$\mathbf{K}^1 \equiv \left[(x^1, x^2, \eta^1, \eta^2) \mid x^{il}_{jhm} \geq 0,\ x^{il}_{kh\hat{l}} \geq 0,\ 0 \leq \eta^{il}_{jhm} \leq 1, \right.$$

$$\left. 0 \leq \eta^{il}_{kh\hat{l}} \leq 1, \forall i, l, j, h, m, k, \hat{l},\ and\, (12.1)\, holds \right] \qquad (10.22)$$

Inequality (10.21) is actually a variational inequality (cf. Nagurney (1999) and the references therein).

10.2.3 The Behavior of the Intermediaries and Their Optimality Conditions

The intermediaries (cf. Figure 10.1), in turn, are involved in transactions both with the source agents in the different countries, as well as with the users of the funds, that is, with the ultimate consumers associated with the markets for the distinct types of loans/products in different currencies and countries and represented by the bottom tier of nodes of the network. Each intermediary node *j; j=1,...,J,* may transact with a demand market via a physical link, and/or electronically via an Internet link. Hence, from each intermediary node *j*, we construct two links to each node $kh\hat{l}$, with the first such link denoting a physical transaction and the second such link denoting an electronic transaction. The corresponding flow, in turn, which is nonnegative, is denoted by $y^{j}_{kh\hat{l}m}$; *m* = 1, 2, and corresponds to the amount of the financial product k in currency h and country $\hat{l}$ transacted from intermediary j via mode m. We group the financial flows between node j and the bottom tier nodes into the column vector $y^j \in R_+^{2KHL}$. All such financial flows for all the intermediaries are then further grouped into the column vector $y \in R_+^{2JKHL}$.

As in the case of source agents, the intermediaries have to bear some costs to establish and maintain relationship levels with source agents and with the consumers. We denote the relationship level between intermediary j and demand market $kh\hat{l}$ transacting through mode m by $\eta^j_{kh\hat{l}m}$. We group the relationship levels for all intermediary/demand market pairs into the column vector $\eta^3 \in R_+^{2JKHL}$. We assume that the relationship levels are nonnegative and that they may assume a value from 0 through 1. These relationship levels represent the flows between the intermediaries and the demand market nodes in the social network level of the supernetwork in Figure 10.1.

Let $\hat{b}^{il}_{jhm}$ denote the cost function associated with the relationship between intermediary j and source agent il transacting in currency h and via mode m and let $b^j_{kh\hat{l}m}$ denote the analogous cost function but associated with intermediary j, demand market $kh\hat{l}$, and mode m. Note that these functions are from the perspective of the intermediary (whereas (10.2) and (10.3) are from the perspective of the source agents). These cost functions are a function of the relationship levels (as in the case of the source agents) and are given by:

$$\hat{b}^{il}_{jhm} = \hat{b}^{il}_{jhm}(\eta^{il}_{jhm}), \quad \forall i,l,j,h,m, \tag{10.23}$$

$$b^j_{kh\hat{l}m} = b^j_{kh\hat{l}m}(\eta^j_{kh\hat{l}m}), \quad \forall i,l,k,h,\hat{l},m. \tag{10.24}$$

The intermediaries also have associated transaction costs in regards to transacting with the source agents, which can depend on the type of currency as well as the source agent. We denote the transaction cost associated with intermediary j transacting with source agent il associated with currency h via mode m by $\hat{c}^{il}_{jhm}$ and we assume that it is of the form

$$\hat{c}^{il}_{jhm} = \hat{c}^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}), \quad \forall i,l,j,h,m, \tag{10.25}$$

that is, such a transaction cost is allowed to depend on the amount allocated by the particular agent in a currency and transacted with the particular intermediary via the particular mode as well as the relationship level between them. In addition, we assume that an intermediary j also incurs a transaction cost $c^j_{kh\hat{l}m}$ associated with transacting with demand market $kh\hat{l}$ through mode m, where

$$c^j_{kh\hat{l}m} = c^j_{kh\hat{l}m}(y^j_{kh\hat{l}m}, \eta^j_{kh\hat{l}m}), \quad \forall j,k,h,\hat{l},m. \tag{10.26}$$

Hence, the transaction costs given in (10.26) can vary according to the intermediary/product/currency/country/mode combination and are a function of the volume of the product transacted and the relationship level.

In addition, an intermediary j is faced with what we term a *handling/conversion* cost, which may include, for example, the cost of converting the incoming financial flows into the financial loans/products associated with the demand markets. We denote such a cost faced by intermediary j by c_j and, in the simplest case, c_j would be a function of $\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2} x_{jhm}^{il}$, that is, the holding/conversion cost of an intermediary is a function of how much he has obtained in the different currencies from the various source agents in the different countries. For the sake of generality, however, we allow the function to depend also on the amounts held by other intermediaries and, therefore, we may write:

$$c_j = c_j(x^1), \quad \forall j. \tag{10.27}$$

We assume that the cost functions (10.23)-(10.27) are convex and continuously differentiable and that the costs are measured in the base currency.

The actual price charged for the financial product k associated with intermediary j transacting with the consumers in currency h via mode m and country $\hat{l}$ is denoted by $\rho_{2kh\hat{l}m}^{j*}$, for intermediary j. Similarly, as in the case of source agents, e_h^* denote the actual rate of appreciation in currency h. Later, we discuss how such prices are arrived at.

We assume that each intermediary seeks to maximize his net revenue with the net revenue criterion for intermediary j being given by:

$$\begin{aligned}
\text{Maximize } & \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}(\rho_{2kh\hat{l}m}^{j*}+e_h^*)y_{kh\hat{l}m}^{j} - c_j(x^1) \\
& -\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{c}_{jhm}^{il}(x_{jhm}^{il},\eta_{jhm}^{il}) - \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}c_{kh\hat{l}m}^{j}(y_{kh\hat{l}m}^{j},\eta_{kh\hat{l}m}^{j}) \\
& -\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{b}_{jhm}^{il}(\eta_{jhm}^{il}) - \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}b_{kh\hat{l}m}^{j}(\eta_{kh\hat{l}m}^{j}) \\
& -\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}(\rho_{1jhm}^{il*}+e_h^*)x_{jhm}^{il}
\end{aligned} \tag{10.28}$$

subject to:

$$\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}y_{kh\hat{l}m}^{j} \le \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}x_{jhm}^{il} \tag{10.29}$$

$$x_{jhm}^{il} \ge 0, \quad y_{kh\hat{l}m}^{j} \ge 0, \quad \forall i,l,h,\hat{l},m. \tag{10.30}$$

$$0 \le \eta^{il}_{jhm} \le 1, \quad 0 \le \eta^{j}_{kh\hat{l}m} \le 1, \quad \forall i,l,h,m,k,\hat{l}. \tag{10.31}$$

Constraint (10.29) guarantees that each intermediary does not reallocate more financial flows than he has available. Constraints (10.30) and (10.31) guarantee that the financial flows and relationship levels are nonnegative (from the perspective of the intermediary) and that the levels of the relationships do not exceed one.

In addition, we assume that each intermediary is also concerned with risk minimization. For the sake of generality, we assume, as given, a risk function $\hat{r}^{il}_{jhm}$, for intermediary j in transacting with source agent il in currency h through mode m and a risk function $r^{j}_{kh\hat{l}m}$ for intermediary j associated with his transacting with consumers at demand market $kh\hat{l}$ through mode m. The risk functions are assumed to be continuous and convex and a function of the amount transacted with the particular source agent or demand market and the relationship level with this source agent or demand market. A higher relationship level can be expected to reduce risk since trust reduces transactional uncertainty. The risk functions may be distinct for each source agent/country/intermediary/currency/mode and intermediary/demand market/currency/country/mode combination and are given, respectively, by:

$$\hat{r}^{il}_{jhm} = \hat{r}^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}), \quad \forall i,l,j,h,m, \tag{10.32}$$

$$r^{j}_{kh\hat{l}m} = r^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m}, \eta^{j}_{kh\hat{l}m}), \quad \forall j,k,h,\hat{l},m. \tag{10.33}$$

Since a financial intermediary j is assumed to minimize his total risk, he is also faced with the optimization problem given by:

$$\text{Minimize} \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2} \hat{r}^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}) + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2} r^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m}, \eta^{j}_{kh\hat{l}m}) \tag{10.34}$$

subject to:

$$x^{il}_{jhm} \ge 0, \quad y^{j}_{kh\hat{l}m} \ge 0, \quad \forall i,l,h,m,k,\hat{l}, \tag{10.35}$$

$$0 \le \eta^{il}_{jhm} \le 1, \quad 0 \le \eta^{j}_{kh\hat{l}m} \le 1, \quad \forall i,l,h,m,k,\hat{l}. \tag{10.36}$$

As in the case of the source agents, intermediary j also tries to maximize his relationship values associated with the source agents and with the demand markets. We assume, as given, a relationship value function $\hat{v}^{il}_{jhm}$ for intermediary j in dealing with source agent il in currency h through transaction mode m and a relationship value function $v^{j}_{kh\hat{l}m}$ for intermediary j associated with his transacting with consumers at demand market $kh\hat{l}$ through mode m. The relationship value functions are assumed to be continuously differentiable and concave. They are assumed to be functions of the corresponding relationship levels and given, respectively, by

$$\hat{v}^{il}_{jhm} = \hat{v}^{il}_{jhm}(\eta^{il}_{jhm}), \quad \forall i,l,j,h,m, \tag{10.37}$$

$$v^{j}_{kh\hat{l}m} = v^{j}_{kh\hat{l}m}(\eta^{j}_{kh\hat{l}m}), \quad \forall j,k,h,\hat{l},m. \tag{10.38}$$

Finally, financial intermediary j tries to maximize his total relationship value, given mathematically by the optimization problem:

$$\text{Maximize } \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{v}^{il}_{jhm}(\eta^{il}_{jhm}) + \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2} v^{j}_{kh\hat{l}m}(\eta^{j}_{kh\hat{l}m}) \tag{10.39}$$

subject to:

$$0 \leq \eta^{il}_{jhm} \leq 1, \quad 0 \leq \eta^{j}_{kh\hat{l}m} \leq 1, \quad \forall i,l,j,h,m,k,\hat{l}. \tag{10.40}$$

10.2.4 The Multicriteria Decision-Making Problem Faced by a Financial Intermediary

We are now ready to construct the multicriteria decision-making problem faced by an intermediary which combines with appropriate individual weights the criteria of net revenue maximization given by (10.28); risk minimization, given by (10.34), and total relationship value maximization, given by (10.39). In particular, we let intermediary j assign a nonnegative weight δ^j to the total risk and a nonnegative weight γ^j to the total relationship value. The weight associated with net revenue maximization is set equal to 1 and serves as the numeraire (as in the case of the source agents). Let U^j denote the multicriteria objective function associated with intermediary j with his multicriteria decision-making problem expressed as:

$$\begin{aligned}\text{Maximize } U^j = &\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}(\rho^{j*}_{2kh\hat{l}m} + e^*_h)y^{j}_{kh\hat{l}m} - c_j(x^1)\\ &-\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{c}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})\\ -\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}c^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m},\eta^{j}_{kh\hat{l}m}) &- \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{b}^{il}_{jhm}(\eta^{il}_{jhm})\\ -\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}b^{j}_{kh\hat{l}m}(\eta^{j}_{kh\hat{l}m}) &- \sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}(\rho^{il*}_{1jhm} + e^*_h)x^{il}_{jhm}\end{aligned}$$

$$-\delta^j(\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{r}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})+\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}r^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m},\eta^{j}_{kh\hat{l}m}))$$

$$+\gamma^j(\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\hat{v}^{il}_{jhm}(\eta^{il}_{jhm})+\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}v^{j}_{kh\hat{l}m}(\eta^{j}_{kh\hat{l}m})) \qquad (10.41)$$

subject to:

$$\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}y^{j}_{kh\hat{l}m}\le\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}x^{il}_{jhm}, \qquad (10.42)$$

$$x^{il}_{jhm}\ge 0,\quad y^{j}_{kh\hat{l}m}\ge 0,\quad \forall i,l,k,h,\hat{l},m, \qquad (10.43)$$

$$0\le\eta^{il}_{jhm}\le 1,\quad 0\le\eta^{j}_{kh\hat{l}m}\le 1,\quad \forall i,l,h,m,k,\hat{l}. \qquad (10.44)$$

Here we assume that the financial intermediaries can compete, with the governing optimality/equilibrium concept underlying noncooperative behavior being that of Nash (1950, 1951), which states that each decision-maker (intermediary) will determine his optimal strategies, given the optimal ones of his competitors. The optimality conditions for all financial intermediaries simultaneously, under the above stated assumptions, can be compactly expressed as (cf. Gabay and Moulin (1980), Dafermos and Nagurney (1987), and Nagurney and Ke (2001, 2003)): determine $(x^{1*},y^{*},\eta^{1*},\eta^{3*},\lambda^{*})\in\mathbf{K}^2$, such that

$$\sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\left[\delta^j\,\frac{\partial\hat{r}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial x^{il}_{jhm}}+\frac{\partial c_j(x^{1*})}{\partial x^{il}_{jhm}}+\rho^{il*}_{1jhm}\right.$$

$$\left.+e^{*}_{h}+\frac{\partial\hat{c}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial x^{il}_{jhm}}-\lambda^{*}_{j}\right]\times\left[x^{il}_{jhm}-x^{il*}_{jhm}\right]$$

$$+\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}\left[\delta^j\,\frac{\partial r^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}\right.$$

$$\left.-\rho^{j*}_{2kh\hat{l}m}-e^{*}_{h}+\lambda^{*}_{j}\right]\times\left[y^{j}_{kh\hat{l}m}-y^{j*}_{kh\hat{l}m}\right]$$

$$+\sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{l=}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}\left[\delta^j\,\frac{\partial\hat{r}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial\eta^{il}_{jhm}}+\frac{\partial\hat{c}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial\eta^{il}_{jhm}}\right.$$

$$\left.-\gamma^j\,\frac{\partial\hat{v}^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial\eta^{il}_{jhm}}+\frac{\partial\hat{b}^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial\eta^{il}_{jhm}}\right]\times\left[\eta^{il}_{jhm}-\eta^{il*}_{jhm}\right]$$

$$+\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}\left[\delta^{j}\frac{\partial r^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}\right.$$

$$\left.-\gamma^{j}\frac{\partial v^{j}_{kh\hat{l}m}(\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}+\frac{\partial b^{j}_{khl}(\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}\right]\times\left[\eta^{j}_{kh\hat{l}m}-\eta^{j*}_{kh\hat{l}m}\right]$$

$$+\sum_{j=1}^{J}\left[\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}x^{il*}_{jhm}-\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}y^{j*}_{kh\hat{l}m}\right]$$

$$\times\left[\lambda_{j}-\lambda^{*}_{j}\right]\geq 0,\quad \forall(x^{1},y,\eta^{1},\eta^{3},\lambda)\in \mathbf{K}^{2}, \tag{10.45}$$

where

$$\mathbf{K}^{2}\equiv\left[(x^{1},y,\eta^{1},\eta^{3},\lambda)\,|\,x^{il}_{jhm}\geq 0,\ y^{j}_{kh\hat{l}m}\geq 0,\ 0\leq\eta^{il}_{jhm}\leq 1,\right.$$

$$\left.0\leq\eta^{j}_{kh\hat{l}m}\leq 1,\ \lambda_{j}\geq 0,\forall i,l,j,h,m,k,\hat{l}\right] \tag{10.46}$$

Here λ_j denotes the Lagrange multiplier associated with constraint (10.42) and λ is the column vector of all the intermediaries' Lagrange multipliers. These Lagrange multipliers can also be interpreted as shadow prices. Indeed, according to the fifth term in (10.45), λ^{*}_{j} serves as the price to ``clear the market" at intermediary *j*.

Inequality (10.45) provides us with conditions under which optimal virtual and/or physical financial transactions between intermediaries and source agents occur and optimal conditions under which virtual transactions between source agents and demand markets occur. Furthermore, it formulates the optimality conditions under which the relationship levels associated with intermediaries interacting with either the source agents or the demand markets will take on positive values; in other words, a relationship exists.

10.2.5 The Consumers at the Demand Markets and the Equilibrium Conditions

We now describe the consumers located at the demand markets. The consumers take into account in making their consumption decisions not only the price charged for the financial product by the agents with source of funds and intermediaries but also their transaction costs associated with obtaining the product.

Let $\hat{c}^{j}_{kh\hat{l}m}$ denote the transaction cost associated with obtaining product *k* in currency *h* in country $\hat{l}$ via mode *m* from intermediary *j* and recall that $y^{j}_{kh\hat{l}m}$ is the amount of the financial product *k* in currency *h* flowing between intermediary *j* and consumers in country $\hat{l}$ via mode *m*. We assume that the transaction cost is measured in the base currency, is continuous, and of the general form:

$$\hat{c}^{j}_{kh\hat{l}m} = \hat{c}^{j}_{kh\hat{l}m}(x^2, y, \eta^2, \eta^3), \quad \forall j,k,h,\hat{l},m. \tag{10.47}$$

Hence, the cost of transacting between an intermediary and a demand market via a specific mode, from the perspective of the consumers, can depend upon the volume of financial flows transacted either physically and/or electronically from intermediaries as well as from source agents and the associated relationship levels. As in the case of the source agents and the financial intermediaries, higher relationship levels potentially reduce transaction costs, which means that they can lead to quantifiable cost reductions. The generality of this cost function structure enables the modeling of competition on the demand side. Moreover, it allows for information exchange between the consumers at the demand markets who may inform one another as to their relationship levels, which, in turn, can affect the transaction costs.

In addition, let $\hat{c}^{il}_{kh\hat{l}}$ denote the transaction cost associated with obtaining the financial product *k* in currency *h* in country $\hat{l}$ electronically from source agent *il*, where we assume that the transaction cost is continuous and of the general form:

$$\hat{c}^{il}_{kh\hat{l}} = \hat{c}^{il}_{kh\hat{l}}(x^2, y, \eta^2, \eta^3), \quad \forall i,l,k,h,\hat{l}. \tag{10.48}$$

Hence, the transaction cost associated with transacting directly with source agents is of a form of the same level of generality as the transaction costs associated with transacting with the financial intermediaries.

Let $\rho_{3kh\hat{l}}$ denote the price of the financial product *k* in currency *h* and in country $\hat{l}$, and defined in the base currency, and group all such prices into the column vector $\rho_3 \in R^{KHL}_{+}$. Denote the demand for product *k* in currency *h* in country $\hat{l}$ by $d_{kh\hat{l}}$ and assume, as given, the continuous demand functions:

$$d_{kh\hat{l}} = d_{kh\hat{l}}(\rho_3), \quad \forall k,h,\hat{l}. \tag{10.49}$$

Thus, according to (10.49), the demand of consumers for the financial product in a currency and country depends, in general, not only on the price of the product at that demand market (and currency and country) but also on the prices of the other products at the other demand markets (and in other countries and currencies). Consequently, consumers at a demand market, in a sense, also compete with consumers at other demand markets.

The consumers take the price charged by the intermediary, which was denoted by $\rho^{j*}_{2kh\hat{l}m}$ for intermediary *j*, product *k*, currency *h*, and country $\hat{l}$ via mode *m*, the price charged by source agent *il*, which was denoted by $\rho^{il*}_{1kh\hat{l}}$, and the rate of appreciation in the currency, plus the transaction costs, in making their consumption decisions. The equilibrium conditions for the consumers at demand market $kh\hat{l}$, thus, take the form: for all intermediaries: *j*; *j*=1,…,*J* and all modes *m*; *m* = 1, 2:

$$\rho^{j*}_{2kh\hat{l}m} + e^*_h + \hat{c}^j_{kh\hat{l}m}(x^{2*}, y^*, \eta^{2*}, \eta^{3*}) \begin{cases} = \rho^*_{3kh\hat{l}}, & \text{if} \quad y^{j*}_{kh\hat{l}m} > 0 \\ \geq \rho^*_{3kh\hat{l}}, & \text{if} \quad y^{j*}_{kh\hat{l}m} = 0, \end{cases} \tag{10.50}$$

and for all source agents $il;\ i = 1,\ldots,I$ and $l = 1,\ldots,L$:

$$\rho^{il*}_{1kh\hat{l}} + e^*_h + \hat{c}^{il}_{kh\hat{l}}(x^{2*}, y^*, \eta^{2*}, \eta^{3*}) \begin{cases} = \rho^*_{3kh\hat{l}}, & \text{if} \quad x^{il*}_{kh\hat{l}} > 0 \\ \geq \rho^*_{3kh\hat{l}}, & \text{if} \quad x^{il*}_{kh\hat{l}} = 0. \end{cases} \tag{10.51}$$

In addition, we must have that

$$d_{kh\hat{l}}(\rho^*_3) \begin{cases} = \sum_{j=1}^{J} \sum_{m=1}^{2} y^{j*}_{kh\hat{l}m} + \sum_{i=1}^{I} \sum_{l=1}^{L} x^{il*}_{kh\hat{l}}, & \text{if} \quad \rho^*_{3kh\hat{l}} > 0 \\ \leq \sum_{j=1}^{J} \sum_{m=1}^{2} y^{j*}_{kh\hat{l}m} + \sum_{i=1}^{I} \sum_{l=1}^{L} x^{il*}_{kh\hat{l}}, & \text{if} \quad \rho^*_{3kh\hat{l}} = 0. \end{cases} \tag{10.52}$$

Conditions (10.50) state that consumers at demand market $kh\hat{l}$ will purchase the product from intermediary j, if the price charged by the intermediary for the product and the appreciation rate for the currency plus the transaction cost (from the perspective of the consumer) does not exceed the price that the consumers are willing to pay for the product in that currency and country, i.e., $\rho^*_{3kh\hat{l}}$. Note that, according to (10.50), if the transaction costs are identically equal to zero, then the price faced by the consumers for a given product is the price charged by the intermediary for the particular product and currency in the country plus the rate of appreciation in the currency. Condition (10.51) states the analogue, but for the case of electronic transactions with the source agents.

Condition (10.52), on the other hand, states that, if the price the consumers are willing to pay for the financial product at a demand market is positive, then the quantity of at the demand market is precisely equal to the demand.

In equilibrium, conditions (10.50), (10.51), and (10.52) will have to hold for all demand markets and these, in turn, can be expressed also as an inequality analogous to those in (10.21) and (10.45) and given by: determine $(x^{2*}, y^*, \rho^*_3) \in R_+^{(IL+2J+1)KHL}$, such that

$$\sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{h=1}^{H} \sum_{\hat{l}=1}^{L} \sum_{m=1}^{2} \left[\rho^{j*}_{2kh\hat{l}m} + e^*_h + \hat{c}^j_{kh\hat{l}m}(x^{2*}, y^*, \eta^{2*}, \eta^{3*}) - \rho^*_{3kh\hat{l}}\right] \times \left[y^j_{kh\hat{l}m} - y^{j*}_{kh\hat{l}m}\right]$$

$$+ \sum_{i=1}^{I} \sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{h=1}^{H} \sum_{\hat{l}=1}^{L} \left[\rho^{il*}_{1kh\hat{l}} + e^*_h + \hat{c}^{il}_{kh\hat{l}}(x^{2*}, y^*, \eta^{2*}, \eta^{3*}) - \rho^*_{3kh\hat{l}}\right] \times \left[x^{il}_{kh\hat{l}} - x^{il*}_{kh\hat{l}}\right]$$

$$+\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\left[\sum_{j=1}^{J}\sum_{m=1}^{2} y_{kh\hat{l}m}^{j*} + \sum_{i=1}^{I}\sum_{l=1}^{L} x_{kh\hat{l}}^{il*} - d_{kh\hat{l}}(\rho_3^*)\right]$$

$$\times\left[\rho_{3kh\hat{l}} - \rho_{3kh\hat{l}}^*\right] \geq 0, \forall (x^2, y, \rho_3) \in R_+^{(IL+2J+1)KHL}. \qquad (10.53)$$

For further background, see Nagurney and Dong (2002).

10.2.6 The Equilibrium Conditions of the Supernetwork Integrating the International Financial Network and the Social Network

In equilibrium, the financial flows that the source agents in different countries transact with the intermediaries must coincide with those that the intermediaries actually accept from them. In addition, the amounts of the financial products that are obtained by the consumers in the different countries and currencies must be equal to the amounts that both the source agents and the intermediaries actually provide. Hence, although there may be competition between decision-makers at the same level of tier of nodes of the financial network there must be, in a sense, cooperation between decision-makers associated with pairs of nodes (through positive flows on the links joining them). Thus, in equilibrium, the prices and financial flows must satisfy the sum of the optimality conditions (10.21) and (10.45) and the equilibrium conditions (10.53). We make these relationships rigorous through the subsequent definition and variational inequality derivation below.

Definition 10.1: Supernetwork Integrating the International Financial Network and the Social Network
The equilibrium state of the supernetwork integrating the international financial network with the social network is one where the financial flows and relationship levels between the tiers of the network coincide and the financial flows, relationship levels, and prices satisfy the sum of conditions (10.21), (10.45), and (10.53).

The equilibrium state is equivalent to the following:

Theorem 10.1: Variational Inequality Formulation
The equilibrium conditions governing the supernetwork integrating the international financial network with the social network according to Definition 10.1 are equivalent to the solution of the variational inequality given by: determine $(x^{1*}, x^{2*}, y^*, \eta^{1*}, \eta^{2*}, \eta^{3*}, \lambda^*, \rho_3^*) \in \mathbf{K}$, *satisfying:*

$$\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}\left[\alpha^{il}\,\frac{\partial r_{jhm}^{il}(x_{jhm}^{il*}, \eta_{jhm}^{il*})}{\partial x_{jhm}^{il}} + \frac{\partial c_{jhm}^{il}(x_{jhm}^{il*}, \eta_{jhm}^{il*})}{\partial x_{jhm}^{il}}\right.$$

$$+\delta^{j}\,\frac{\partial \hat{r}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial x^{il}_{jhm}}+\frac{\partial c_{j}(x^{1*})}{\partial x^{il}_{jhm}}+\frac{\partial \hat{c}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial x^{il}_{jhm}}-\lambda^{*}_{j}\Bigg]$$

$$\times\left[x^{il}_{jhm}-x^{il*}_{jhm}\right]$$

$$+\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\Bigg[\alpha^{il}\,\frac{\partial r^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}},\eta^{il*}_{kh\hat{l}})}{\partial x^{il}_{kh\hat{l}}}+\frac{\partial c^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}},\eta^{il*}_{kh\hat{l}})}{\partial x^{il}_{kh\hat{l}}}$$

$$+\hat{c}^{il}_{kh\hat{l}}(x^{2*},y^{*},\eta^{2*},\eta^{3*})-\rho^{*}_{3kh\hat{l}}\Bigg]\times\left[x^{il}_{kh\hat{l}}-x^{il*}_{kh\hat{l}}\right]$$

$$+\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}\Bigg[\delta^{j}\,\frac{\partial r^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}$$

$$+\hat{c}^{j}_{kh\hat{l}m}(x^{2*},y^{*},\eta^{2*},\eta^{3*})+\lambda^{*}_{j}-\rho^{*}_{3kh\hat{l}}\Bigg]\times\left[y^{j}_{kh\hat{l}m}-y^{j*}_{kh\hat{l}m}\right]$$

$$+\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}\Bigg[\frac{\partial c^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}+\frac{\partial \hat{c}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}$$

$$-\beta^{il}\,\frac{\partial v^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}-\gamma^{j}\,\frac{\partial \hat{v}^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}+\alpha^{il}\,\frac{\partial r^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}$$

$$+\delta^{j}\,\frac{\partial \hat{r}^{il}_{jhm}(x^{il*}_{jhm},\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}+\frac{\partial b^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}+\frac{\partial \hat{b}^{il}_{jhm}(\eta^{il*}_{jhm})}{\partial \eta^{il}_{jhm}}\Bigg]$$

$$\times\left[\eta^{il}_{jhm}-\eta^{il*}_{jhm}\right]$$

$$+\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\Bigg[\frac{\partial c^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}},\eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}+\frac{\partial b^{il}_{kh\hat{l}}(\eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}$$

$$-\beta^{il}\,\frac{\partial v^{il}_{kh\hat{l}}(\eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}+\alpha^{il}\,\frac{\partial r^{il}_{kh\hat{l}}(x^{il*}_{kh\hat{l}},\eta^{il*}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}\Bigg]\times\left[\eta^{il}_{kh\hat{l}}-\eta^{il*}_{kh\hat{l}}\right]$$

$$+\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}\Bigg[\delta^{j}\,\frac{\partial r^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m},\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}$$

$$-\gamma^{j}\,\frac{\partial v^{j}_{kh\hat{l}m}(\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}+\frac{\partial b^{j}_{khl}(\eta^{j*}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}\Bigg]\times\left[\eta^{j}_{kh\hat{l}m}-\eta^{j*}_{kh\hat{l}m}\right]$$

$$+\sum_{j=1}^{J}\left[\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}\sum_{m=1}^{2}x^{il*}_{jhm}-\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}y^{j*}_{kh\hat{l}m}\right]\times\left[\lambda_{j}-\lambda^{*}_{j}\right]$$

$$+\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\left[\sum_{j=1}^{J}\sum_{m=1}^{2}y^{j*}_{kh\hat{l}m}+\sum_{i=1}^{I}\sum_{l=1}^{L}x^{il*}_{kh\hat{l}}-d_{kh\hat{l}}(\rho^{*}_{3})\right]$$

$$\times\left[\rho_{3kh\hat{l}} - \rho^*_{3kh\hat{l}}\right] \geq 0, \quad \forall (x^1, x^2, y, \eta^1, \eta^2, \eta^3, \lambda, \rho_3) \in \mathbf{K}, \tag{10.54}$$

where

$$\mathbf{K} \equiv \left[(x^1, x^2, y, \eta^1, \eta^2, \eta^3, \lambda, \rho_3) \mid x^{il}_{jhm} \geq 0,\ x^{il}_{kh\hat{l}} \geq 0,\ y^{j}_{kh\hat{l}m} \geq 0,\ 0 \leq \eta^{il}_{jhm} \leq 1,\right.$$
$$\left. 0 \leq \eta^{il}_{kh\hat{l}} \leq 1, 0 \leq \eta^{j}_{kh\hat{l}m} \leq 1,\ \rho_{3kh\hat{l}} \geq 0,\ \lambda_j \geq 0, \forall i, l, j, h, m, k, \hat{l},\ \mathit{and}\, (12.1)\, \mathit{holds}\right] \tag{10.55}$$

Proof: Summation of inequalities (10.21), (10.45), and (10.53), yields, after algebraic simplification, the variational inequality (10.54).

We now put variational inequality (10.54) into standard form which will be utilized in the subsequent sections. For additional background on variational inequalities and their applications, see the book by Nagurney (1999). In particular, we have that variational inequality (10.54) can be expressed as:

$$\langle F(X^*), X - X^* \rangle \geq 0, \quad \forall X \in \mathbf{K}, \tag{10.56}$$

Where
$X \equiv (x^1, x^2, y, \eta^1, \eta^2, \eta^3, \lambda, \rho_3)$ and
$F(X) \equiv (F_{iljhm}, F_{ilkh\hat{l}}, F_{jkh\hat{l}m}, \hat{F}_{iljhm}, \hat{F}_{ilkh\hat{l}}, \hat{F}_{jkh\hat{l}m}, F_j, F_{kh\hat{l}})$ with indices: $i = 1, \ldots, I$; $l = 1, \ldots, L; j = 1, \ldots, J$; $h = 1, \ldots, H$; $\hat{l} = 1, \ldots, L; m = 1, 2$, and the specific components of F given by the functional terms preceding the multiplication signs in (10.56), respectively. The term $\langle \cdot, \cdot \rangle$ denotes the inner product in N-dimensional Euclidean space.

We now describe how to recover the prices associated with the first two tiers of nodes in the international financial network. Clearly, the components of the vector ρ^*_3 are obtained directly from the solution of variational inequality (10.56) as will be demonstrated explicitly through several numerical examples in Section 5. In order to recover the second tier prices associated with the intermediaries and the exchange rates one can (after solving variational inequality (10.56) for the particular numerical problem) *either* (cf. (10.50)) set

$$\rho^{j*}_{2kh\hat{l}m} + e^*_h = [\rho^*_{3kh\hat{l}} - \hat{c}^{j}_{kh\hat{l}m}(x^{2*}, y^*, \eta^{2*}, \eta^{3*})] \ ,$$

for any $j, k, h, \hat{l}, m$ such that $y^{j*}_{kh\hat{l}m} > 0$, *or* (cf. (10.45)) for any $y^{j*}_{kh\hat{l}m} > 0$, set

$$\rho^{j*}_{2kh\hat{l}m} + e^*_h = \left[\delta^j \frac{\partial r^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m}, \eta^{j*}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}} + \frac{\partial c^{j}_{kh\hat{l}m}(y^{j*}_{kh\hat{l}m}, \eta^{j*}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}} + \lambda^*_j \right].$$

Similarly, from (10.21) we can infer that the top tier prices comprising the vector ρ_1^* can be recovered (once the variational inequality (10.56) is solved with particular data) thus: for any i,l,j,h,m, such that $x_{jhm}^{il*} > 0$, set

$$\rho_{1jhm}^{il*} + e_h^* = \left[\alpha^{il} \frac{\partial r_{jhm}^{il}(x_{jhm}^{il*}, \eta_{jhm}^{il*})}{\partial x_{jhm}^{il}} + \frac{\partial c_{jhm}^{il}(x_{jhm}^{il*}, \eta_{jhm}^{il*})}{\partial x_{jhm}^{il}} \right],$$

or, equivalently, (cf. (10.45)), to

$$\left[\lambda_j^* - \delta^j \frac{\partial \hat{r}_{jhm}^{il}(x_{jhm}^{il*}, \eta_{jhm}^{il*})}{\partial x_{jhm}^{il}} - \frac{\partial c_j(x^{1*})}{\partial x_{jhm}^{il}} - \frac{\partial \hat{c}_{jhm}^{il}(x_{jhm}^{il*}, \eta_{jhm}^{il*})}{\partial x_{jhm}^{il}} \right].$$

In addition, in order to recover the first tier prices associated with the demand market and the exchange rates one can (after solving variational inequality (10.56) for the particular numerical problem) *either* (cf. (10.21)) set

$$\rho_{1kh\hat{l}}^{il*} + e_h^* = \left[\alpha^{il} \frac{\partial r_{kh\hat{l}}^{il}(x_{kh\hat{l}}^{il*}, \eta_{kh\hat{l}}^{il*})}{\partial x_{kh\hat{l}}^{il}} + \frac{\partial c_{kh\hat{l}}^{il}(x_{kh\hat{l}}^{il*}, \eta_{kh\hat{l}}^{il*})}{\partial x_{kh\hat{l}}^{il}} \right],$$

for any $i,l,k,h,\hat{l}$ such that $x_{kh\hat{l}}^{il*} > 0$, *or* (cf. (10.51)) for any $x_{kh\hat{l}}^{il*} > 0$, set

$$\rho_{1kh\hat{l}}^{il*} + e_h^* = [\rho_{3kh\hat{l}}^* - \hat{c}_{kh\hat{l}}^{il}(x^{2*}, y^*, \eta^{2*}, \eta^{3*})].$$

Under the above pricing mechanism, the optimality conditions (10.21) and (10.45) as well as the equilibrium conditions (10.53) also hold separately (as well as for each individual decision-maker).

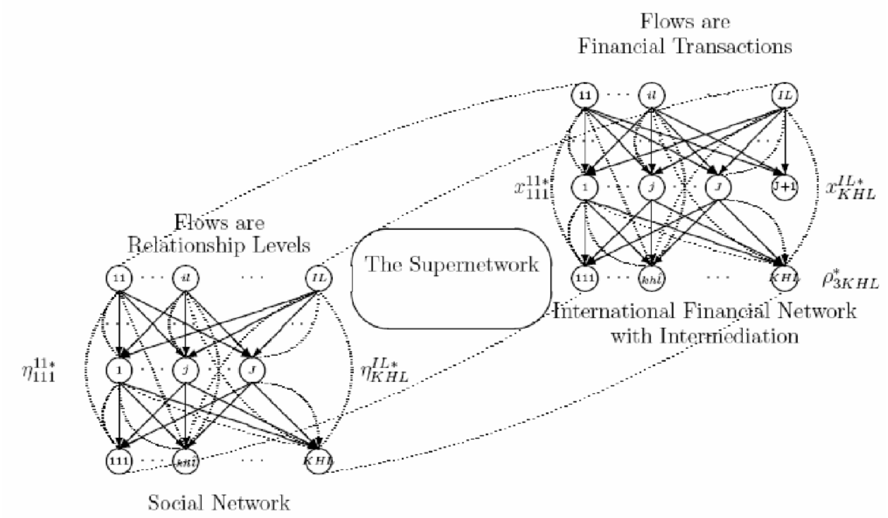


Fig.10. 2. The supernetwork at equilibrium

In Figure 10.2, we display the supernetwork in equilibrium in which the equilibrium financial flows, relationship levels, and prices now appear. Note that, if the equilibrium values of the flows (be they financial or relationship levels) on links are identically equal to zero, then those links can effectively be removed from the supernetwork (in equilibrium). Moreover, the size of the equilibrium flows represent the ``strength" of respective links (as discussed also in the social network/supply chain network equilibrium model of Wakolbinger and Nagurney (2004)). Thus, the supernetwork model developed here also provides us with the emergent integrated social and financial network structures. In the next section, we discuss the dynamic evolution of the financial flows, relationship levels, and prices until this equilibrium is achieved.

10.3 The Dynamic Adjustment Process

In this section, we describe the dynamics associated with the supernetwork model developed in Section 10.2 and formulate the corresponding dynamic model as a projected dynamical system (cf. Nagurney and Zhang (1996a), Nagurney and Ke (2003), Nagurney and Cruz (2004), and Nagurney, Wakolbinger, and Zhao (2004)). Importantly, the set of stationary points of the projected dynamical system which formulates the dynamic adjustment process will coincide with the set of solutions to the variational inequality problem (10.54). In particular, we describe the disequilibrium dynamics of the international financial flows, the relationship levels, as well as the prices.

10.3.1 The Dynamics of the Financial Flows From the Source Agents

Note that, unlike the financial flows (as well as the prices associated with the distinct nodal tiers of the network) between the intermediaries and the demand markets, the financial flows from the source agents are subject not only to nonnegativity constraints but also to budget constraints (cf. (10.1)). Hence, in order to guarantee that these constraints are not violated we need to introduce some additional machinery based on projected dynamical systems theory in order to describe the dynamics of these financial flows (see also, e.g., Nagurney and Siokos (1997), Nagurney and Zhang (1996a), Nagurney and Cruz (2004), and Nagurney, Cruz, and Matsypura (2003)).

In particular, we denote the rate of change of the vector of financial flows from source agent *il* by $\dot{x}^{il}$ and noting that the *best realizable direction* for the financial flows from source agent *il* must include the constraints, we have that:

$$\dot{x}^{il} = \Pi_{\mathbf{K}^{il}}(x^{il}, -F^{il}), \tag{10.57}$$

where Π_K is defined as (see also Nagurney and Zhang (1996a)):

$$\Pi_K(x, v) = \lim_{\delta \to 0} \frac{P_K(x + \delta v) - x}{\delta}, \tag{10.58}$$

and P_K is the norm projection defined by

$$P_K(x) = \operatorname{argmin}_{x' \in K} \|x' - x\|. \tag{10.59}$$

The feasible set $\mathbf{K}^{il}$ is defined as: $\mathbf{K}^{il} \equiv \{x^{il} \mid x^{il} \in R_+^{2JH+KHL} \text{ and satisfies (12.1)}\}$, and F^{il} is the vector (see following (10.56)) with components: $F_{iljhm}, F_{ilkh\hat{l}}$ and with indices: $j = 1, \ldots, J$; $h = 1, \ldots, H$; $m = 1, 2$, and $k = 1, \ldots, K$. Hence, expression (10.57) reflects that the financial flow on a link emanating from a source agent will increase if the price (be it the market-clearing price associated with an intermediary or a demand market price) exceeds the various costs and weighted marginal risk; it will decrease if the latter exceeds the former.

10.3.2 The Dynamics of the Financial Products Between the Intermediaries and the Demand Markets

The rate of change of the financial flow $y^j_{kh\hat{l}m}$, denoted by $\dot{y}^j_{kh\hat{l}m}$, is assumed to be equal to the difference between the price the consumers are willing to pay for the financial product at the demand market minus the price charged and the various transaction costs and the weighted marginal risk associated with the transaction.

Here we also guarantee that the financial flows do not become negative. Hence, we may write: for every $j,k,h,\hat{l},m$:

$$\dot{y}^{j}_{kh\hat{l}m}=\begin{cases}\rho_{3kh\hat{l}}-\delta^{j}\dfrac{\partial r^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m},\eta^{j}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}-\dfrac{\partial c^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m},\eta^{j}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}} & \\ -\hat{c}^{j}_{kh\hat{l}m}(x^{2},y,\eta^{2},\eta^{3})-\lambda_{j}, & \text{if } y^{j}_{kh\hat{l}m}>0 \\ & \\ \max\{0,\rho_{3kh\hat{l}}-\delta^{j}\dfrac{\partial r^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m},\eta^{j}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}-\dfrac{\partial c^{j}_{kh\hat{l}m}(y^{j}_{kh\hat{l}m},\eta^{j}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}} & \\ -\hat{c}^{j}_{kh\hat{l}m}(x^{2},y,\eta^{2},\eta^{3})-\lambda_{j}\}, & \text{if } y^{j}_{kh\hat{l}m}=0.\end{cases}$$

(10.60)

Hence, according, to (10.60), if the price that the consumers are willing to pay for the product (in the currency and country) exceeds the price that the intermediary charges and the various transaction costs and weighted marginal risk, then the volume of flow of the product to that demand market will increase; otherwise, it will decrease (or remain unchanged).

10.3.3 The Dynamics of the Relationship Levels Between the Source Agents and the Financial Intermediaries

Now the dynamics of the relationship levels between the source agents in the various countries and the intermediaries are described. The rate of change of the relationship level η^{il}_{jhm}, denoted by $\dot{\eta}^{il}_{jhm}$, is assumed to be equal to the difference between the weighted relationship value for source agent *il*, intermediary *j*, currency *h* and mode *m*, and the sum of the marginal costs and the weighted marginal risks. Again, one must also guarantee that the relationship levels do not become negative. Moreover, they may not exceed the level equal to one. Hence, we can immediately write:

$$\dot{\eta}^{il}_{jhm}=\begin{cases}\beta^{il}\dfrac{\partial v^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}+\gamma^{j}\dfrac{\partial\hat{v}^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\alpha^{il}\dfrac{\partial r^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}\\ -\dfrac{\partial c^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial\hat{c}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}\\ -\delta^{j}\dfrac{\partial\hat{r}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial b^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial\hat{b}^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}, & \text{if } 0<\eta^{il}_{jhm}<1\\[2ex] \min\{1,\max\{0,\beta^{il}\dfrac{\partial v^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}+\gamma^{j}\dfrac{\partial\hat{v}^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}\\ -\alpha^{il}\dfrac{\partial r^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial c^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial\hat{c}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}\\ -\delta^{j}\dfrac{\partial\hat{r}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial b^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}-\dfrac{\partial\hat{b}^{il}_{jhm}(\eta^{il}_{jhm})}{\partial\eta^{il}_{jhm}}\}\}, & \text{otherwise}\end{cases}\tag{10.61}$$

where $\dot{\eta}^{il}_{jhm}$ denotes the rate of change of the relationship level η^{il}_{jhm}.

This shows that if the sum of the weighted relationship values for the source agent and the intermediary are higher than the total marginal costs plus the total weighted marginal risk, then the level of relationship between that financial source agent and intermediary pair will increase. If it is lower, the relationship value will decrease.

10.3.4 The Dynamics of the Relationship Levels Between the Source Agents and the Demand Markets

Here we describe the dynamics of the relationship levels between the source agents and the demand markets. The rate of change of the relationship level $\eta^{il}_{kh\hat{l}}$ in turn, responds to the difference between the weighted relationship value for source agent *il* and the sum of the marginal costs and weighted marginal risks. One also must guarantee that these relationship levels do not become negative (nor higher than one). Hence, one may write:

$$\dot{\eta}^{il}_{kh\hat{l}} = \begin{cases} \beta^{il} \dfrac{\partial v^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} - \dfrac{\partial c^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} - \dfrac{\partial b^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} \\ - \alpha^{il} \dfrac{\partial r^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}, & \text{if} \quad 0 < \eta^{il}_{kh\hat{l}} < 1, \\ \\ \min\{1, \max\{0, \beta^{il} \dfrac{\partial v^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} - \dfrac{\partial c^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} \\ - \dfrac{\partial b^{il}_{kh\hat{l}}(\eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}} - \alpha^{il} \dfrac{\partial r^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}, \eta^{il}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}\}\}, & \text{otherwise} \end{cases} \tag{10.62}$$

where $\dot{\eta}^{il}_{kh\hat{l}}$ denotes the rate of change of the relationship level $\eta^{il}_{kh\hat{l}}$. This shows that if the weighted relationship value for the source agent is higher than the total marginal costs plus the total weighted marginal risk, then the level of relationship between that financial source agent and demand market pair will increase. If it is lower, the relationship value will decrease. Of course, the bounds on the relationship levels must also hold.

10.3.5 The Dynamics of the Relationship Levels Between the Financial Intermediaries and the Demand Markets

The dynamics of the relationship levels between the financial intermediaries and demand markets are now described. The rate of change of the relationship level product $\dot{\eta}^{j}_{kh\hat{l}m}$ transacted via mode m is assumed to be equal to the difference between the weighted relationship value for intermediary j and the sum of the marginal costs and weighted marginal risks, where, of course, one also must guarantee that the relationship levels do not become negative nor exceed one. Hence, one may write:

$$\dot{\eta}^j_{kh\hat{l}m}=\begin{cases} \gamma^j \frac{\partial v^j_{kh\hat{l}m}(\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} - \delta^j \frac{\partial r^j_{kh\hat{l}m}(y^j_{kh\hat{l}m},\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} & \\ -\frac{\partial c^j_{kh\hat{l}m}(y^j_{kh\hat{l}m},\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} - \frac{\partial b^j_{khl}(\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} & \text{if } \; 0<\eta^j_{kh\hat{l}m}<1, \\ & \\ \min\{1,\max\{0,\gamma^j \frac{\partial v^j_{kh\hat{l}m}(\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} - \delta^j \frac{\partial r^j_{kh\hat{l}m}(y^j_{kh\hat{l}m},\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} & \\ -\frac{\partial c^j_{kh\hat{l}m}(y^j_{kh\hat{l}m},\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}} - \frac{\partial b^j_{khl}(\eta^j_{kh\hat{l}m})}{\partial \eta^j_{kh\hat{l}m}}\}\}, & \mathit{otherwise} \end{cases} \tag{10.63}$$

where $\dot{\eta}^j_{kh\hat{l}m}$ denotes the rate of change of the relationship level $\eta^j_{kh\hat{l}m}$. Expression (10.63) reveals that if the weighted relationship value for the intermediary with the demand market is higher than the total marginal costs plus the total weighted marginal risk, then the level of relationship between that intermediary and demand market pair will increase. If it is lower, the relationship value will decrease.

10.3.6 Demand Market Price Dynamics

We assume that the rate of change of the price $\rho_{3kh\hat{l}}$, denoted by $\dot{\rho}_{3kh\hat{l}}$, is equal to the difference between the demand for the financial product at the demand market in the currency and country and the amount of the product actually available at that particular market. Hence, if the demand for the product at the demand market at an instant in time exceeds the amount available from the various intermediaries and source agents, then the price will increase; if the amount available exceeds the demand at the price, then the price will decrease. Moreover, it is guaranteed that the prices do not become negative. Thus, the dynamics of the price $\rho_{3kh\hat{l}}$ for each $k,h,\hat{l}$ can be expressed as:

$$\dot{\rho}_{3kh\hat{l}}=\begin{cases} d_{kh\hat{l}}(\rho_3)-\sum_{j=1}^{J}\sum_{m=1}^{2} y^j_{kh\hat{l}m}-\sum_{i=1}^{I}\sum_{l=1}^{L} x^{il}_{kh\hat{l}}, & \text{if } \; \rho_{3kh\hat{l}}>0 \\ \max\{0, d_{kh\hat{l}}(\rho_3)-\sum_{j=1}^{J}\sum_{m=1}^{2} y^j_{kh\hat{l}m}-\sum_{i=1}^{I}\sum_{l=1}^{L} x^{il}_{kh\hat{l}}\}, & \text{if } \; \rho_{3kh\hat{l}}=0. \end{cases} \tag{10.64}$$

10.3.7 The Dynamics of the Prices at the Intermediaries

The prices at the intermediaries, whether they are physical or virtual, must reflect supply and demand conditions as well. In particular, we let $\dot{\lambda}_j$ denote the rate of change in the market clearing price associated with intermediary *j* and we propose the following dynamic adjustment for every intermediary *j*:

$$\dot{\lambda}_j = \begin{cases} \sum_{k=1}^{K} \sum_{h=1}^{H} \sum_{\hat{l}=1}^{L} \sum_{m=1}^{2} y^{j}_{kh\hat{l}m} - \sum_{i=1}^{I} \sum_{l=1}^{L} \sum_{h=1}^{H} \sum_{m=1}^{2} x^{il}_{jhm}, & \text{if } \lambda_j > 0 \\ \max\{0, \sum_{k=1}^{K} \sum_{h=1}^{H} \sum_{\hat{l}=1}^{L} \sum_{m=1}^{2} y^{j}_{kh\hat{l}m} - \sum_{i=1}^{I} \sum_{l=1}^{L} \sum_{h=1}^{H} \sum_{m=1}^{2} x^{il}_{jhm}\}, & \text{if } \lambda_j = 0. \end{cases} \tag{10.65}$$

Hence, if the financial flows from the source agents in the countries into an intermediary exceed the amount demanded at the demand markets from the intermediary, then the market-clearing price at that intermediary will decrease; if, on the other hand, the volume of financial flows into an intermediary is less than that demanded by the consumers at the demand markets (and handled by the intermediary), then the market-clearing price at that intermediary will increase.

10.3.7 The Projected Dynamical System

We now turn to stating the complete dynamic model. In the dynamic model the flows evolve according to the mechanisms described above; specifically, the financial flows from the source agents evolve according to (10.57) for all source agents *il*. The financial flows from the financial intermediaries to the demand markets evolve according to (10.60) for all financial intermediaries *j,* demand markets $kh\hat{l}$, and modes *m*. The relationship levels between source agents and financial intermediaries for all modes *m* evolve according to (10.61), the relationship levels between source agents *il* and demand markets $kh\hat{l}$ evolve according to (10.62), and the relationship levels between financial intermediaries j and demand markets $kh\hat{l}$ for all modes *m* evolve according to (10.63). Furthermore, the prices associated with the intermediaries evolve according to (10.64) for all intermediaries *j*, and the demand market prices evolve according to (10.65) for all *k*.

Let X and $F(X)$ be as defined following (10.56) and recall the feasible set $\mathbf{K}$. Then the dynamic model described by (10.57), (10.60)-(10.65) can be rewritten as a *projected dynamical system* (Nagurney and Zhang (1996a)) defined by the following initial value problem:

$$\dot{X} = \Pi_{\mathbf{K}}(X, -F(X)), \quad X(0) = X_0, \tag{10.66}$$

where $\Pi_{\mathbf{K}}$ is the projection operator of $-F(X)$ onto $\mathbf{K}$ at X (cf. (10.58)) and $X_0 = (x^{10}, x^{20}, y^0, \eta^{10}, \eta^{20}, \eta^{30}, \lambda^0, \rho_3^0)$ is the initial point corresponding to the initial financial flow and price pattern.

The trajectory of (10.66) describes the dynamic evolution of the relationship levels on the social network, the financial product transactions on the financial network, the demand market prices and the Lagrange multipliers or shadow prices associated with the intermediaries. The projection operation guarantees the constraints underlying the supernetwork system are not violated. Recall that the constraint set $\mathbf{K}$ consists not only of the conservation of flow constraints (cf. (10.1)) associated with the source agents but also the nonnegativity constraints associated with all the financial flows, the prices, as well as the relationships levels. Moreover, the relationship levels are assumed to not exceed the value of one.

Following Dupuis and Nagurney (1993) and Nagurney and Zhang (1996a), the following result is immediate.

Theorem 10.2: Set of Stationary Points Coincides with Set of Equilibrium Points

The set of stationary points of the projected dynamical system (10.66) coincides with the set of solutions of the variational inequality problem (10.54) and, thus, with the set of equilibrium points as defined in Definition 10.1.

With Theorem 10.2, we see that the dynamical system proposed in this Section provides the disequilibrium dynamics prior to the steady or equilibrium state of the international financial network. Hence, once, a stationary point of the projected dynamical system is reached, that is, when $\dot{X}=0$ in (10.66), that point (consisting of financial flows, relationship levels, and prices) also satisfies variational inequality (10.54); equivalently, (10.56), and is, therefore, an international financial network equilibrium according to Definition 10.1.

The above described dynamics are very reasonable from an economic perspective and also illuminate that there must be cooperation between tiers of decision-makers although there may be competition within a tier.

We now state the following:

Theorem 10.3: Existence and Uniqueness of a Solution to the Initial Value Problem

Assume that $F(X)$ *is Lipschitz continuous, that is, that*

$$\|F(X')-F(X'')\| \leq \mathbf{L}\|X'-X''\|, \quad \forall X',X'' \in \mathbf{K}, \text{ where } \mathbf{L}>0. \tag{10.67}$$

Then, for any $X_0 \in \mathbf{K}$, *there exists a unique solution* $X_0(t)$ *to the initial value problem (10.66).*

Proof: Lipschitz continuity of the function F is sufficient for the result following Theorem 2.5 in Nagurney and Zhang (1996a).

Under suitable conditions on the underlying functions (see also Nagurney and Dong (2002), Zhang and Nagurney (1995), and Nagurney, Wakolbinger, and Zhao (2004)), one can obtain stability results for the supernetwork. A similar result was obtained for a supply chain network model with electronic commerce and relationship levels in Wakolbinger and Nagurney (2004).

10.4 The Discrete Time Algorithm

In this Section, we propose the Euler method for the computation of solutions to variational inequality (10.54); equivalently, the stationary points of the projected dynamical system (10.66). The Euler method is a special case of the general iterative scheme introduced by Dupuis and Nagurney (1993) for the solution of projected dynamical systems. Besides providing a solution to variational inequality problem (10.54), this algorithm also yields a time discretization of the continuous-time adjustment process of the projected dynamical system (10.66). This discretization may also be interpreted as a discrete-time adjustment process. Conditions for convergence of this algorithm are given in Dupuis and Nagurney (1993) and in Nagurney and Zhang (1996a). In Section 5, we apply this algorithm to several numerical examples.

10.4.1 The Euler Method

The statement of the Euler method is the following: At iteration $\mathbf{T}$ compute

$$X_{\mathbf{T}} = P_{\mathbf{K}}(X_{\mathbf{T}-1} - a_{\mathbf{T}-1}F(X_{\mathbf{T}-1})), \tag{10.68}$$

where P_{K} denotes the projection operator in the Euclidean sense (cf. (10.57) and Nagurney (1999)) onto the closed convex set $\mathbf{K}$ and $F(X)$ is defined following (10.56). We discuss the sequence of positive terms a_{T} below. The complete statement of the method in the context of the dynamic supernetwork model is as follows:

Step 0: Initialization

Set $(x^{10}, x^{20}, y^{0}, \eta^{10}, \eta^{20}, \eta^{30}, \lambda^{0}, \rho_3^{0}) \in \mathbf{K}$. Let $\mathbf{T} = 1$ and set the sequence $\{a_{\mathbf{T}}\}$ so that $\sum_{\mathbf{T}=1}^{\infty} a_{\mathbf{T}} = \infty$, $a_{\mathbf{T}} > 0$, $a_{\mathbf{T}} \to 0$, as $\mathbf{T} \to \infty$ (such a sequence is required for convergence of the algorithm).

Step 1: Computation

Compute $(x^{1\mathbf{T}}, x^{2\mathbf{T}}, y^{\mathbf{T}}, \eta^{1\mathbf{T}}, \eta^{2\mathbf{T}}, \eta^{3\mathbf{T}}, \lambda^{\mathbf{T}}, \rho_3^{\mathbf{T}}) \in \mathbf{K}$ by solving the variational inequality subproblem:

$$\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}\Bigg[x_{jhm}^{il\mathbf{T}} + a_{\mathbf{T}}(\alpha^{il}\frac{\partial r_{jhm}^{il}(x_{jhm}^{il\mathbf{T}-1}, \eta_{jhm}^{il\mathbf{T}-1})}{\partial x_{jhm}^{il}} + \frac{\partial c_{jhm}^{il}(x_{jhm}^{il\mathbf{T}-1}, \eta_{jhm}^{il\mathbf{T}-1})}{\partial x_{jhm}^{il}}$$

$$+\delta^{j}\frac{\partial\hat{r}^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial x^{il}_{jhm}}+\frac{\partial c_{j}(x^{1\mathbf{T}-1})}{\partial x^{il}_{jhm}}+\frac{\partial\hat{c}^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial x^{il}_{jhm}}-\lambda^{\mathbf{T}-1}_{j})-x^{il\mathbf{T}-1}_{jhm}\Bigg]$$

$$\times\Big[x^{il}_{jhm}-x^{il\mathbf{T}}_{jhm}\Big]$$

$$+\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\Bigg[x^{il\mathbf{T}}_{kh\hat{l}}+a_{\mathbf{T}}(\alpha^{il}\frac{\partial r^{il}_{kh\hat{l}}(x^{il\mathbf{T}-1}_{kh\hat{l}},\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial x^{il}_{kh\hat{l}}}+\frac{\partial c^{il}_{kh\hat{l}}(x^{il\mathbf{T}-1}_{kh\hat{l}},\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial x^{il}_{kh\hat{l}}}$$

$$+\hat{c}^{il}_{kh\hat{l}}(x^{2\mathbf{T}-1},y^{\mathbf{T}-1},\eta^{2\mathbf{T}-1},\eta^{3\mathbf{T}-1})-\rho^{\mathbf{T}-1}_{3kh\hat{l}})-x^{il\mathbf{T}-1}_{kh\hat{l}}\Bigg]$$

$$\times\Big[x^{il}_{kh\hat{l}}-x^{il\mathbf{T}}_{kh\hat{l}}\Big]$$

$$+\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}\Bigg[y^{j\mathbf{T}}_{kh\hat{l}m}+a_{\mathbf{T}}(\delta^{j}\frac{\partial r^{j}_{kh\hat{l}m}(y^{j\mathbf{T}-1}_{kh\hat{l}m},\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j\mathbf{T}-1}_{kh\hat{l}m},\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial y^{j}_{kh\hat{l}m}}$$

$$+\hat{c}^{j}_{kh\hat{l}m}(x^{2\mathbf{T}-1},y^{\mathbf{T}-1},\eta^{2\mathbf{T}-1},\eta^{3\mathbf{T}-1})+\lambda^{\mathbf{T}-1}_{j}-\rho^{\mathbf{T}-1}_{3kh\hat{l}})-y^{j\mathbf{T}-1}_{kh\hat{l}m}\Bigg]$$

$$\times\Big[y^{j}_{kh\hat{l}m}-y^{j\mathbf{T}}_{kh\hat{l}m}\Big]$$

$$+\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{j=1}^{J}\sum_{h=1}^{H}\sum_{m=1}^{2}\Bigg[\eta^{il\mathbf{T}}_{jhm}+a_{\mathbf{T}}(\frac{\partial c^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}+\frac{\partial\hat{c}^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}$$

$$-\beta^{il}\frac{\partial v^{il}_{jhm}(\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}-\gamma^{j}\frac{\partial\hat{v}^{il}_{jhm}(\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}+\alpha^{il}\frac{\partial r^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}$$

$$+\delta^{j}\frac{\partial\hat{r}^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}+\frac{\partial b^{il}_{jhm}(\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}}+\frac{\partial\hat{b}^{il}_{jhm}(\eta^{il\mathbf{T}-1}_{jhm})}{\partial\eta^{il}_{jhm}})-\eta^{il\mathbf{T}-1}_{jhm}\Bigg]$$

$$\times\Big[\eta^{il}_{jhm}-\eta^{il\mathbf{T}}_{jhm}\Big]$$

$$+\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\Bigg[\eta^{il\mathbf{T}}_{kh\hat{l}}+a_{\mathbf{T}}(\frac{\partial c^{il}_{kh\hat{l}}(x^{il\mathbf{T}-1}_{kh\hat{l}},\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial\eta^{il}_{kh\hat{l}}}+\frac{\partial b^{il}_{kh\hat{l}}(\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial\eta^{il}_{kh\hat{l}}}$$

$$-\beta^{il}\frac{\partial v^{il}_{kh\hat{l}}(\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial\eta^{il}_{kh\hat{l}}}+\alpha^{il}\frac{\partial r^{il}_{kh\hat{l}}(x^{il\mathbf{T}-1}_{kh\hat{l}},\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial\eta^{il}_{kh\hat{l}}})-\eta^{il\mathbf{T}-1}_{kh\hat{l}}\Bigg]\times\Big[\eta^{il}_{kh\hat{l}}-\eta^{il\mathbf{T}}_{kh\hat{l}}\Big]$$

$$+\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\sum_{m=1}^{2}\Bigg[\eta^{j\mathbf{T}}_{kh\hat{l}m}+a_{\mathbf{T}}(\delta^{j}\frac{\partial r^{j}_{kh\hat{l}m}(y^{j\mathbf{T}-1}_{kh\hat{l}m},\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial\eta^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j\mathbf{T}-1}_{kh\hat{l}m},\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial\eta^{j}_{kh\hat{l}m}}$$

$$-\gamma^{j}\frac{\partial v^{j}_{kh\hat{l}m}(\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial\eta^{j}_{kh\hat{l}m}}+\frac{\partial b^{j}_{khl}(\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial\eta^{j}_{kh\hat{l}m}})-\eta^{j\mathbf{T}-1}_{kh\hat{l}m}\Bigg]\times\Big[\eta^{j}_{kh\hat{l}m}-\eta^{j\mathbf{T}}_{kh\hat{l}m}\Big]$$

$$+\sum_{j=1}^{J}\left[\lambda_j^{\mathbf{T}} + a_{\mathbf{T}}\left(\sum_{m=1}^{2}\left[\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H} x_{jhm}^{il\mathbf{T}-1} - \sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L} y_{kh\hat{l}m}^{j\mathbf{T}-1}\right]\right) - \lambda_j^{\mathbf{T}-1}\right]\times\left[\lambda_j - \lambda_j^{\mathbf{T}}\right]$$

$$+\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}\left[\rho_{3kh\hat{l}}^{\mathbf{T}} + a_{\mathbf{T}}\left(\sum_{j=1}^{J}\sum_{m=1}^{2} y_{kh\hat{l}m}^{j\mathbf{T}-1} + \sum_{i=1}^{I}\sum_{\hat{l}=1}^{L} x_{kh\hat{l}}^{il\mathbf{T}-1} - d_{kh\hat{l}}(\rho_3^{\mathbf{T}-1})\right) - \rho_{3kh\hat{l}}^{\mathbf{T}-1}\right]$$

$$\times\left[\rho_{3kh\hat{l}} - \rho_{3kh\hat{l}}^{\mathbf{T}}\right] \geq 0, \quad \forall (x^1, x^2, y, \eta^1, \eta^2, \eta^3, \lambda, \rho_3) \in \mathbf{K}. \qquad (10.69)$$

Step 2: Convergence Verification

If $|x_{jhm}^{il\mathbf{T}} - x_{jhm}^{il\mathbf{T}-1}| \leq \varepsilon$, $|x_{kh\hat{l}}^{il\mathbf{T}} - x_{kh\hat{l}}^{il\mathbf{T}-1}| \leq \varepsilon$, $|y_{kh\hat{l}m}^{j\mathbf{T}} - y_{kh\hat{l}m}^{j\mathbf{T}-1}| \leq \varepsilon$, $|\eta_{jhm}^{il\mathbf{T}} - \eta_{jhm}^{il\mathbf{T}-1}| \leq \varepsilon$, $|\eta_{kh\hat{l}}^{il\mathbf{T}} - \eta_{kh\hat{l}}^{il\mathbf{T}-1}| \leq \varepsilon$, $|\eta_{kh\hat{l}m}^{j\mathbf{T}} - \eta_{kh\hat{l}m}^{j\mathbf{T}-1}| \leq \varepsilon$, $|\lambda_j^{\mathbf{T}} - \lambda_j^{\mathbf{T}-1}| \leq \varepsilon$, $|\rho_{3kh\hat{l}}^{\mathbf{T}} - \rho_{3kh\hat{l}}^{\mathbf{T}-1}| \leq \varepsilon$, for all $i = 1, \cdots, I$; $l = 1, \ldots, L$; $\hat{l} = 1, \ldots, L$; $m = 1, 2$; $j = 1, \cdots, J$; $h = 1, \ldots, H$; $k = 1, \cdots, K$, with $\varepsilon > 0$, a pre-specified tolerance, then stop; otherwise, set $\mathbf{T} = \mathbf{T} + 1$, and go to Step 1.

Due to the simplicity of the feasible set $\mathbf{K}$ the solution of (10.69) is accomplished exactly and in closed form. In (10.69) the variational subproblem of the variables (x^1, x^2) can be solved using exact equilibration (cf. Dafermos and Sparrow (1969), Nagurney (1999)). The other variables can be obtained using the following explicit formulae:

Computation of the Financial Products from the Intermediaries

In particular, compute, at iteration $\mathbf{T}$, the $y_{kh\hat{l}m}^{j\mathbf{T}}$ s, according to:

$$y_{kh\hat{l}m}^{j\mathbf{T}} = \max\{0, y_{kh\hat{l}m}^{j\mathbf{T}-1} - a_{\mathbf{T}}(\delta^j \frac{\partial r_{kh\hat{l}m}^{j}(y_{kh\hat{l}m}^{j\mathbf{T}-1}, \eta_{kh\hat{l}m}^{j\mathbf{T}-1})}{\partial y_{kh\hat{l}m}^{j}} + \frac{\partial c_{kh\hat{l}m}^{j}(y_{kh\hat{l}m}^{j\mathbf{T}-1}, \eta_{kh\hat{l}m}^{j\mathbf{T}-1})}{\partial y_{kh\hat{l}m}^{j}}$$

$$+ \hat{c}_{kh\hat{l}m}^{j}(x^{2\mathbf{T}-1}, y^{\mathbf{T}-1}, \eta^{2\mathbf{T}-1}, \eta^{3\mathbf{T}-1}) + \lambda_j^{\mathbf{T}-1} - \rho_{3kh\hat{l}}^{\mathbf{T}-1})\}, \quad \forall j, k, h, \hat{l}, m. \qquad (10.70)$$

Computation of the Relationship Levels

At iteration $\mathbf{T}$ compute the $\eta_{jhm}^{il\mathbf{T}}s$ according to:

$$\eta_{jhm}^{il\mathbf{T}} = \min\{1, \max\{0, \eta_{jhm}^{il\mathbf{T}-1} - a_{\mathbf{T}}(\frac{\partial c_{jhm}^{il}(x_{jhm}^{il\mathbf{T}-1}, \eta_{jhm}^{il\mathbf{T}-1})}{\partial \eta_{jhm}^{il}} + \frac{\partial \hat{c}_{jhm}^{il}(x_{jhm}^{il\mathbf{T}-1}, \eta_{jhm}^{il\mathbf{T}-1})}{\partial \eta_{jhm}^{il}}$$

$$- \beta^{il} \frac{\partial v_{jhm}^{il}(\eta_{jhm}^{il\mathbf{T}-1})}{\partial \eta_{jhm}^{il}} - \gamma^{j} \frac{\partial \hat{v}_{jhm}^{il}(\eta_{jhm}^{il\mathbf{T}-1})}{\partial \eta_{jhm}^{il}} + \alpha^{il} \frac{\partial r_{jhm}^{il}(x_{jhm}^{il\mathbf{T}-1}, \eta_{jhm}^{il\mathbf{T}-1})}{\partial \eta_{jhm}^{il}}$$

$$+\delta^{j}\frac{\partial \hat{r}^{il}_{jhm}(x^{il\mathbf{T}-1}_{jhm},\eta^{il\mathbf{T}-1}_{jhm})}{\partial \eta^{il}_{jhm}}+\frac{\partial b^{il}_{jhm}(\eta^{il\mathbf{T}-1}_{jhm})}{\partial \eta^{il}_{jhm}}+\frac{\partial \hat{b}^{il}_{jhm}(\eta^{il\mathbf{T}-1}_{jhm})}{\partial \eta^{il}_{jhm}}))\}\},\quad \forall i,l,j,h,m. \quad (2.71)$$

Furthermore, at iteration $\mathbf{T}$ compute the $\eta^{il\mathbf{T}}_{kh\hat{l}}$ s according to:

$$\eta^{il\mathbf{T}}_{kh\hat{l}}=\min\{1,\max\{0,\eta^{il\mathbf{T}-1}_{kh\hat{l}}-a_{\mathbf{T}}(\frac{\partial c^{il}_{kh\hat{l}}(x^{il\mathbf{T}-1}_{kh\hat{l}},\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}+\frac{\partial b^{il}_{kh\hat{l}}(\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}-\beta^{il}\frac{\partial v^{il}_{kh\hat{l}}(\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}}$$

$$+\alpha^{il}\frac{\partial r^{il}_{kh\hat{l}}(x^{il\mathbf{T}-1}_{kh\hat{l}},\eta^{il\mathbf{T}-1}_{kh\hat{l}})}{\partial \eta^{il}_{kh\hat{l}}})\}\},\quad \forall i,l,k,h,\hat{l}. \quad (10.72)$$

At iteration $\mathbf{T}$ compute the $\eta^{j\mathbf{T}}_{kh\hat{l}m}$ s according to:

$$\eta^{j\mathbf{T}}_{kh\hat{l}m}=\min\{1,\max\{0,\eta^{j\mathbf{T}}_{kh\hat{l}m}-a_{\mathbf{T}}(\delta^{j}\frac{\partial r^{j}_{kh\hat{l}m}(y^{j\mathbf{T}-1}_{kh\hat{l}m},\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}+\frac{\partial c^{j}_{kh\hat{l}m}(y^{j\mathbf{T}-1}_{kh\hat{l}m},\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}$$

$$-\gamma^{j}\frac{\partial v^{j}_{kh\hat{l}m}(\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}}+\frac{\partial b^{j}_{khl}(\eta^{j\mathbf{T}-1}_{kh\hat{l}m})}{\partial \eta^{j}_{kh\hat{l}m}})\}\},\quad \forall j,k,h,\hat{l},m. \quad (10.73)$$

Computation of the Shadow Prices

At iteration $\mathbf{T}$ compute the $\gamma^{\mathbf{T}}_{j}$ s according to:

$$\gamma^{\mathbf{T}}_{j}=\max\{0,\gamma^{\mathbf{T}-1}_{j}-a_{\mathbf{T}}(\sum_{m=1}^{2}\left[\sum_{i=1}^{I}\sum_{l=1}^{L}\sum_{h=1}^{H}x^{il\mathbf{T}-1}_{jhm}-\sum_{k=1}^{K}\sum_{h=1}^{H}\sum_{\hat{l}=1}^{L}y^{j\mathbf{T}-1}_{kh\hat{l}m}\right])\},\quad \forall j. \quad (2.74)$$

Computation of the Demand Market Prices

Finally, at iteration $\mathbf{T}$ compute the demand market prices, the $\rho^{\mathbf{T}}_{3kh\hat{l}}$ s, according to:

$$\rho^{\mathbf{T}}_{3kh\hat{l}}=\max\{0,\rho^{\mathbf{T}-1}_{3kh\hat{l}}-a_{\mathbf{T}}(\sum_{j=1}^{J}\sum_{m=1}^{2}y^{j\mathbf{T}-1}_{kh\hat{l}m}+\sum_{i=1}^{I}\sum_{l=1}^{L}x^{il\mathbf{T}-1}_{kh\hat{l}}-d_{kh\hat{l}}(\rho^{\mathbf{T}-1}_{3}))\},\quad \forall k,h,\hat{l}. \quad (10.75)$$

As one can see in the discrete-time adjustment process(es) described above the algorithm is initialized with a vector of financial flows, relationship levels, and prices. For example, the relationship levels may be set to zero (and the same holds for the prices, initially). The financial flows, shadow prices, and the demand market prices are computed in the financial network of the supernetwork. The financial product transactions between intermediaries and demand markets are computed according to (10.70). The relationship levels are computed in the social network of the supernetwork according to (10.71), (10.72), and (10.73), respectively. Finally, the shadow prices are computed according to (10.74) and the demand market prices are computed according to (10.75).

The dynamic supernetwork system will then evolve according to the discrete-time adjustment process (10.70) through (10.75) until a stationary/equilibrium point of the projected dynamical system (10.66) (equivalently, and a solution to variational inequality (10.54)) is achieved. Once the convergence tolerance has been reached then the equilibrium conditions according to Definition 10.1 are satisfied as one can see from (10.70) through (10.75).

In the next Section, we apply the Euler method to solve several international financial network examples. Convergence results for this algorithm may be found in Dupuis and Nagurney (1993) and for a variety of applications in Nagurney and Zhang (1996a).

10.5 Numerical Examples

In this section, we applied the Euler method to several supernetwork examples as discussed in the preceding sections. The Euler method was implemented in FORTRAN and the computer system used was a Sun system located at the University of Massachusetts at Amherst. For the solution of the induced network subproblems in the (x^1,x^2) variables we utilized the exact equilibration algorithm (see Dafermos and Sparrow (1969), Nagurney (1999), and the references therein). The other variables were determined explicitly and in closed form as described in the preceding section.

The convergence criterion used was that the absolute value of the flows, prices, and relationship levels between two successive iterations differed by no more than 10^{-4}. The sequence $\{a_T\}$ used for all the examples was: $1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \dots$ in the algorithm.

We assumed in all the examples that the risk was represented through variance-covariance matrices for both the source agents in the countries and for the financial intermediaries (see also Nagurney and Cruz (2004)). We initialized the Euler method as follows: We set $x^{il}_{jh1} = \frac{S^{il}}{JH}$ for each source agent *i* and country *l* and for all *j* and *h*. All the other variables were initialized to zero.

10.5.1 Example 10.1

The first numerical example consisted of one country, two source agents, two currencies, two intermediaries, and two financial products. Hence, $L = 1$, $I = 2$, $H = 2$, $J = 2$, and $K = 2$. In this example, the electronic transactions were only between the source agents and the demand markets. In addition, for simplicity, we considered the possibility of the existence of relationship levels only between the source agents and the intermediaries, and the intermediaries and demand markets.

The data for the first example were constructed for easy interpretation purposes. The financial holdings of the two source agents were: $S^{11} = 20$ and $S^{21} = 20$. The variance-covariance matrices were equal to the identity matrices (appropriately dimensioned) for all source agents and all intermediaries. Note that since only physical transactions are allowed (except for as stated above), we have that $m = 1$.

The transaction cost functions faced by the source agents associated with transacting with the intermediaries (cf. (10.4)) were given by:

$$c^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm}) = .5(x^{il}_{jhm})^2 + 3.5x^{il}_{jhm} - \eta^{il}_{jhm};$$
$$i = 1,2; l = 1; j = 1,2; h = 1,2; m = 1.$$

The analogous transaction costs but associated with the electronic transactions between source agents and demand markets (cf. (10.4)) were given by:

$$c^{il}_{kh\hat{l}}(x^{il}_{kh\hat{l}}) = .5(x^{il}_{kh\hat{l}})^2 + x_{kh\hat{l}}; \quad \forall i,l,\hat{l},k,h.$$

Note that here we did not include relationship levels in the functional forms.

The handling costs of the intermediaries, in turn (see (10.27)), were given by:

$$c_j(x^1) = .5(\sum_{i=1}^{2}\sum_{h=1}^{2} x^{i1}_{jh1})^2; \quad j = 1,2.$$

The transaction costs of the intermediaries associated with transacting with the source agents were (cf. (10.25)) given by:

$$\hat{c}^{il}_{jhm}(x^{il}_{jhm},\eta^{il}_{jhm}) = 1.5x^{il\ 2}_{jhm} + 3x^{il}_{jhm}; \; i = 1,2; l = 1; j = 1,2; h = 1,2; m = 1.$$

The transaction costs, in turn, associated with the electronic transactions at the demand markets (from the perspective of the consumers (cf. (10.48)) were given by:

$$\hat{c}^{il}_{kh\hat{l}}(x^2, y, \eta^2, \eta^3) = .1x^{il}_{kh\hat{l}} + 1, \quad \forall i,l,\hat{l},k,h.$$

The demand functions at the demand markets (refer to (10.49)) were:

$$d_{111}(\rho_3) = -2\rho_{3111} - 1.5\rho_{3121} + 1000, \; d_{121}(\rho_3) = -2\rho_{3121} - 1.5\rho_{3111} + 1000,$$
$$d_{211}(\rho_3) = -2\rho_{3211} - 1.5\rho_{3221} + 1000, \; d_{221}(\rho_3) = -2\rho_{3221} - 1.5\rho_{3211} + 1000,$$

and the transaction costs between the intermediaries and the consumers at the demand markets (see (10.47)) were given by:

$$\hat{c}^j_{kh\hat{l}m}(x^2,y,\eta^2,\eta^3)=y^j_{kh\hat{l}m}+5-\eta^j_{kh\hat{l}m};\ k=1,2;h=1,2;\hat{l}=1;m=1.$$

The relationship value functions (10.14) and (10.38) were given by:

$$v^{il}_{jhm}(\eta^{il}_{jhm})=\eta^{il}_{jhm},\ \ \forall i,l,j,h,m;\ \ v^j_{kh\hat{l}m}(\eta^j_{kh\hat{l}m})=\eta^j_{kh\hat{l}m},\ \ \forall j,k,h,\hat{l},m,$$

with all other relationship value functions being set equal to zero.

The relationship cost functions (cf. (10.2) and (10.24)) were as follows:

$$b^{il}_{jhm}(\eta^{il}_{jhm})=2\eta^{il}_{jhm},\ \ \forall i,l,j,h,m;\ \ b^j_{kh\hat{l}m}(\eta^j_{kh\hat{l}m})=\eta^j_{kh\hat{l}m},\ \ \forall j,k,h,\hat{l},m.$$

Since all the weights associated with the criteria were set equal to one this means that the source agents as well as the intermediaries weight the criterion of risk minimization equally to that of net revenue maximization and relationship value maximization.

The Euler method converged in 2,998 iterations and yielded the following equilibrium financial flow pattern:

$$\begin{aligned}
x^{1*} &:= x^{11*}_{111}=x^{11*}_{121}=x^{11*}_{211}=x^{11*}_{221}=x^{211}_{11*}=x^{22*}_{211}=x^{21*}_{211}=x^{21*}_{221}=.0662;\\
x^{2*} &:= x^{11*}_{111}=x^{11*}_{121}=x^{11*}_{211}=x^{11*}_{221}=x^{21*}_{111}=x^{21*}_{121}=x^{21*}_{211}=x^{21*}_{221}=4.9938;\\
y^{*} &:= y^{1*}_{1111}=y^{1*}_{1211}=y^{1*}_{2111}=y^{1*}_{2211}=y^{2*}_{1111}=y^{2*}_{1211}=y^{2*}_{2111}=y^{2*}_{2211}=.0642.
\end{aligned}$$

Both source agents allocated the entirety of their funds to the instrument in the two currencies; thus, there was no non-investment.
The vector λ^* had components: $\lambda^*_1=\lambda^*_2=272.7246,$ and the computed demand prices at the demand markets were: $\rho^*_{3111}=\rho^*_{3121}=\rho^*_{3211}=\rho^*_{3221}=282.8586.$

All the relationship levels were identically equal to zero.

Note that due to the lower transaction costs associated with electronic transactions directly between the source agents and the demand markets a sizeable portion of the financial funds were transacted in this manner.

10.5.2 Example 10.2

Example 10.2 was constructed from Example 10.1 as follows. We kept the data the same except that we increased the weights associated with the relationship levels of the source agents from 1 to 20. Hence, in this example, the source agents weight relationship levels much higher than in Example 10.1. The Euler method again converged requiring the same number of iterations as in Example 10.1, but now the relationship levels of all the source agents increased to the maximum possible value of 1. All other computed equilibrium values remained as in Example 10.1.

10.5.3 Example 10.3

The third example consisted of two countries with two source agents in each country; two currencies, two intermediaries, and two financial products. Hence, $L = 2$, $I = 2$, $H = 2$, $J = 2$, and $K = 2$.

The data for Example 3 was constructed for easy interpretation purposes and to create a baseline from which additional simulations could be conducted. In fact, we essentially ``replicated" the data for the first country as it appeared in Example 10.1in order to construct the data for the second country.

Specifically, the financial holdings of the source agents were: $S^{11} = 20$, $S^{21} = 20$, $S^{12} = 20$, and $S^{22} = 20$. The variance-covariance matrices were equal to the identity matrices (appropriately dimensioned) for all source agents in each country and for all intermediaries, respectively.

The transaction cost functions faced by the source agents associated with transacting with the intermediaries were given by:

$$c^{il}_{jhm}(x^{il}_{jhm}, \eta^{il}_{jhm}) = .5(x^{il}_{jhm})^2 + 3.5x^{il}_{jhm} - \eta^{il}_{jhm};\ i = 1,2; l = 1,2; j = 1,2; h = 1,2; m = 1.$$

The handling costs of the intermediaries (since the number of intermediaries is still equal to two) remained as in Example 10.1, that is, they were given by:

$$c_j(x^1) = .5(\sum_{i=1}^{2}\sum_{h=1}^{2} x^{i1}_{jh1})^2;\ j = 1,2.$$

The transaction costs of the intermediaries associated with transacting with the source agents in the two countries were given by:

$$\hat{c}^{il}_{jhm}(x^{il}_{jhm}) = 1.5(x^{il}_{jhm})^2 + 3x^{il}_{jhm};\ i = 1,2; l = 1,2; j = 1,2; h = 1,2; m = 1.$$

The demand functions at the demand markets were:

$$\begin{aligned}
d_{111}(\rho_3) &= -2\rho_{3111} - 1.5\rho_{3121} + 1000, & d_{121}(\rho_3) &= -2\rho_{3121} - 1.5\rho_{3111} + 1000,\\
d_{211}(\rho_3) &= -2\rho_{3211} - 1.5\rho_{3221} + 1000, & d_{221}(\rho_3) &= -2\rho_{3221} - 1.5\rho_{3211} + 1000,\\
d_{112}(\rho_3) &= -2\rho_{3112} - 1.5\rho_{3122} + 1000, & d_{122}(\rho_3) &= -2\rho_{3122} - 1.5\rho_{3112} + 1000,\\
d_{212}(\rho_3) &= -2\rho_{3212} - 1.5\rho_{3222} + 1000, & d_{222}(\rho_3) &= -2\rho_{3222} - 1.5\rho_{3212} + 1000,
\end{aligned}$$

and the transaction costs between the intermediaries and the consumers at the demand markets were given by:

$$\hat{c}^j_{kh\hat{l}}(y^j_{kh\hat{l}m},\eta^j_{kh\hat{l}m}) = y^j_{kh\hat{l}} - \eta^j_{kh\hat{l}} + 5; \quad j=1,2; k=1,2; h=1,2; \hat{l}=1,2; m=1.$$

The data for the electronic links were as in Example 10.1 and were replicated for the other source agents.

The variance-covariance matrices were redimensioned and were equal to the identity matrices. The weights associated with the risk functions were set equal to 1 for all the source agents and intermediaries.

The Euler method converged in 1,826 iterations and yielded the following equilibrium international financial flow pattern: only the electronic links had positive flows with all other flows being identically equal to 0.000. In particular, the financial holdings of the source agents in the different countries were equally allocated via electronic transactions directly to the demand markets with $x^{il*}_{kh\hat{l}} = 2.5000$ for all $i,l,k,h,\hat{l}$.

The vector λ^* had components: $\lambda^*_1 = \lambda^*_2 = 279.6194$, and the computed demand prices at the demand markets were: $\rho^*_{3kh\hat{l}} = 282.8578$, $\forall k,h,\hat{l}$. In this example, all the financial transactions were conducted electronically.

10.5.3 Example 10.4

Example 4 was constructed from Example 10.3 in a similar manner to that in which Example 10.2 was constructed from Example 10.1. In other words, we now increased the weight associated with all the source agents in both countries from 1 to 20 regarding the criterion of relationship value maximization. The equilibrium pattern computed by the Euler method was identical to that obtained for Example 10.3 but now the relationship values associated with the source agents were all equal to their upper bound with a value of 1.

These examples, although stylized, have been presented to show both the model and the computational procedure. Obviously, different input data and dimensions of the problems solved will affect the equilibrium financial flow, price, and relationship level patterns. One now has a powerful tool with which to explore the effects of perturbations to the data as well as the effects of changes in the number of source agents, countries, currencies, and/or products, as well as the effects of the introduction of electronic transactions. Moreover, this supernetwork framework allows for the tracking of the co-evolution of the international financial flows, the product prices, as well as the relationship levels over space and time.

10.6 Summary and Conclusions

In this chapter, we developed a supernetwork model that integrated international financial networks with intermediation with social networks in which relationship levels were made explicit. Both networks had three tiers of decision-makers, consisting of: the sources of financial funds in the countries, the financial intermediaries, and the consumers associated with the demand markets for the financial products. We allowed for physical as well as electronic transactions between the decision-makers in the supernetwork. The relationship levels were allowed to affect not only the risk but also the transaction costs (by reducing them, in general) but did have associated costs. Moreover, we allowed for multicriteria decision-making behavior in which the source agents as well as the financial intermediaries were permitted to weight, in an individual fashion, their objective functions of net revenue maximization, total risk minimization, and total relationship value maximization.

We modeled the supernetwork in equilibrium, in which the international financial flows between the tiers as well as the relationship levels coincide and established the variational inequality formulation of the governing equilibrium conditions. We then proposed the underlying dynamics and the continuous-time adjustment process (es) and constructed its projected dynamical system representation. We established that the set of stationary points of the projected dynamical system coincides with the set of solutions of the variational inequality problem. We also provided conditions under which the dynamic trajectories of the financial flows, relationship levels, and prices are well-defined.

We proposed a discrete-time algorithm to approximate the continuous-time adjustment process and applied it to several simple numerical examples for completeness and illustrative purposes. The framework developed here further advances the work in financial equilibrium modeling and analysis, especially within a network context by explicitly considering the effect of relationship levels in financial networks and by establishing the optimal relationship levels as well as financial transaction quantities and prices. Finally, it also gives us insight into the optimal designs of the supernetworks.

This framework generalizes the recent work of Nagurney, Wakolbinger, and Zhao (2004) in the evolution and emergence of integrated social and financial networks with electronic transactions to the international dimension. In addition, it incorporates social networks into the international financial network framework of Nagurney and Cruz (2003).

Acknowledgments

This research was supported, in part, by NSF Grant No.: IIS-0002647 and, in part, by an AT&T Industrial Ecology Fellowship. This support is gratefully acknowledged.

References

Anthony DL (1997) Micro-lending institutions: Using social networks to create productive capabilities. International Journal of Sociology and Social Policy 17:156-178

Arrow KJ (1998) What has economics to say about racial discrimination? Journal of Economic Perspectives 12:91-100

Bazaraa MS, Sherali HD, Shetty CM (1993) Nonlinear programming: Theory and algorithms. John Wiley & Sons, New York

Claessens S (2003) Benefits and costs of integrated financial services provision. Forthcoming in Litan RE, Herring R (eds) Brookings-Wharton Papers on Financial Services 2002. The Brookings Institution, Washington, DC

Claessens S, Glaessner T, Klingebiel D (2000) Electronic finance: Reshaping the financial landscape around the world. Financial Sector Discussion Paper No. 4, The World Bank, Washington, DC

Claessens S, Glaessner T, Klingebiel D (2001) E-finance in emerging markets: Is leapfrogging possible?" Financial Sector Discussion Paper No. 7, The World Bank, Washington, DC.

Claessens S, Dobos G, Klingebiel D, Laeven L (2003) The growing importance of networks in finance and their effects on competition. In: Nagurney A (ed) Innovations in Financial and Economic Networks. Edward Elgar Publishers, Cheltenham, England, pp 110-135

Dafermos S, Sparrow FT (1969) The traffic assignment problem for a general network. Journal of Research of the National Bureau of Standards 73B:91-118

DiMaggio P, Louch H (1998) Socially embedded consumer transactions: For what kinds of purchases do people most often use networks? American Sociological Review 69:619-637

Dong J, Zhang D, Nagurney A (2002) Supply chain networks with multicriteria decision-makers. In: Taylor MAP (ed) Transportation and traffic theory in the 21st century. Pergamon, The Netherlands

Dupuis P, Nagurney A (1993) Dynamical Systems and Variational Inequalities. Annals of Operations Research 44:9-42

G-10 (2001) Report on consolidation in financial sector. Bank for International Settlements, Switzerland

Gabay D, Moulin H (1980) On the uniqueness and stability of Nash equilibria in noncooperative games. In: Bensoussan A, Kleindorfer P, Tapiero CS (eds) Applied stochastic control of econometrics and management science. North-Holland, Amsterdam, The Netherlands, pp 271-294

Ghatak M (2002) Exploiting social networks to alleviate credit market failures: On the endogenous selection of peer groups in microfinance programs," Prepared for the Conference on Credit, Trust and Calculation at the University of California.

Keeney RL, Raiffa H (1993) Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Cambridge University Press, Cambridge, England

Kim HM (1999) Globalization of international financial markets. Ashgate, Hants, England

Markowitz HM (1952) Portfolio selection. The Journal of Finance 7:77-91

Markowitz HM (1959) Portfolio selection: Efficient diversification of investments, John Wiley & Sons, Inc., New York

Nagurney A (1999) Network economics: A variational inequality approach, second and revised edition. Kluwer Academic Publishers, Dordrecht, The Netherlands

Nagurney A (ed) (2003) Innovations in financial and economic networks. Edward Elgar Publishing, Cheltenham, England

Nagurney A, Cruz J (2003) International financial networks with electronic transactions. In: Nagurney A (ed) Innovations in Financial and Economic Networks. Edward Elgar Publishing, Cheltenham, England, pp 136-168

Nagurney A, Cruz J (2004) Dynamics of international financial networks with risk management. Quantitative Finance 4:276-291

Nagurney A, Cruz J, Matsypura D (2003) Dynamics of global supply chain supernetworks. Mathematical and Computer Modelling 37: 963-983

Nagurney A, Dong J (2002) Supernetworks: Decision-making for the information age. Edward Elgar Publishing, Cheltenham, England

Nagurney A, Ke K (2001) Financial networks with intermediation. Quantitative Finance 1:441-451

Nagurney A, Ke K (2003) Financial networks with electronic transactions: Modeling, analysis, and computations," Quantitative Finance 3:71-87

Nagurney A, Siokos S (1997) Financial networks: Statics and dynamics. Springer-Verlag, Heidelberg, Germany

Nagurney A, Wakolbinger T, Zhao L (2004) The evolution and emergence of integrated social and financial networks with electronic transactions: A dynamic supernetwork theory for the modeling, analysis, and computation of financial flows and relationship levels. Isenberg School of Management, University of Massachusetts, Amherst; http://supernet.som.umass.edu

Nagurney A, Zhang D (1996a) Projected dynamical systems and variational inequalities with applications. Kluwer Academic Publishers, Dordrecht, The Netherlands

Nagurney A, Zhang D (1996b) Stability of spatial price equilibrium modeled as a projected dnamical system. Journal of Economic Dynamica and Control 20:43-63

Nagurney A, Zhang D (1997) Formulation, stability, and computation of traffic network equilibria as projected dynamical systems. Journal of Optimization Theory and its Applications 93:417-444

Nash JF (1950) Equilibrium points in n-person games. Proceedings of the National Academy of Sciences, USA 36:48-49

Nash JF (1951) Noncooperative games. Annals of Mathematics 54:286-298

Sharpe SA (1990) Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. The Journal of Finance 45:1069-1085

Uzzi B (1997) Social structure and competition in interfirm networks: The paradox of embeddedness. Administrative Science Quarterly 42:35-67

Uzzi B (1999) Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. American Sociological Review 64:481-505

Wakolbinger T, Nagurney A (2004) Dynamic supernetworks for the integration of social networks and supply chains with electronic commerce: modeling and analysis of buyer-seller relationships with computations. Isenberg School of Management, University of Massachusetts, Amherst; http://supernet.som.umass.edu

Zhang D, Nagurney A (1995) On the stability of projected dynamical systems. Journal of Optimization Theory and its Applications 85:97-124

Part C: Methodological Advances— General Equilibrium Models

11 Regional Adjustment to Globalization: A CGE Analytical Framework

James A. Giesecke and John R. Madden

Centre of Policy Studies, Monash University, Melbourne, Australia

11.1 Introduction

During the past decade or so regional modelers have been developing a new range of economic modeling tools with a capacity to analyze the causes and consequences of the current period of globalization. In this chapter, we examine some developments in multiregional computable general equilibrium (CGE) models that have enhanced their capacity to deal with economic globalization issues.[1]

Globalization, at least in terms of increases in the international trade of goods and services, and increases in the mobility of financial capital and labor, is not a new phenomenon.[2] However, the last few decades has seen a dramatic upsurge in the integration of the world economy.[3] This integration has seen an increase in international trade, particularly intra-industry trade and capital flows. At the back of this has been: (i) a global move towards deregulation and a reduction in external, and internal, barriers to trade; and (ii) technological developments resulting, in particular, in improved telecommunications and eCommerce. Increases in scale economies and reductions in transport costs have, *inter alia*, led to a fragmentation of production across large distances and reduced inventories.

The flexibility of the CGE framework has seen a multiplicity of studies employing models of this class to tackle a range of globalization questions. Our primary concern in this chapter is with (multi)regional CGE models where the regions are at the sub-national level. However, developments made with multi-

---

[1] CGE models are sometimes known as AGE (applied general equilibrium) models.

[2] For instance, World Bank (2000) notes that in the last wave of globalization around the end of the 19th century, trade and capital market flows between certain countries were, relative to GDP, around today's levels.

[3] Indeed, Salvatore (2004, p.421) claims "Globalization is a revolution, which in scope and significance is comparable to the industrial revolution, but while the industrial revolution took place over a century or so, the Globalization Revolution has taken place under our very eyes in one or two decades and is continuing unabated."

country CGE models and with national CGE models are of relevance, as innovations with these models are regularly transplanted to regional and multiregional models. Consequently, we will consider some treatments of various globalization issues by CGE models, regardless of their geographical level of specification, before putting such treatments into a multiregional context.

Of the various dimensions of economic globalization, it has been the international trade arrangements that have been the most analyzed by CGE models. A myriad of CGE studies have covered economic integration at various levels, including multilateral (global and regional) and bilateral trade arrangements, as well as more comprehensive political-economic unions. There are numerous studies of the North American Free Trade Agreement (see, for instance, the studies in Francois and Shiells 1994, and references in Kehoe 2005), the European Community/Union (see, for example, Burniaux and Waelbroeck 1992, Gasiorek et al. 2002, and Willenbockel 1994), and other regional arrangements, such as Asia Pacific Economic Cooperation (Adams 1998a, Walmsley 2002). Other studies examine the effects on a particular country of accessing the WTO (e.g. Diao et al. 2003), or of a variety of trade agreements (Gosh 2002, and Lee 2004).

Many of the CGE models employed in such studies are multi-country models, such as GTAP (Hertel 1997). Multi-country models have much in common with multiregional models of individual countries.[4] However, there are quite a few instances of multiregional models used. For instance, Diao et al. (2003) divide China into seven regions. On the other hand Gazel et al. (1995) employ a model with three US regions, plus Canada and RoW (the rest of the world). Haddad et al. (2002) use a top-down disaggregation of results from an economy-wide CGE model of Brazil, while Jean and Laborde (2004) use a two-stage process that feeds national CGE results into stand-alone CGE models of individual (sub-national) regions.

While trade modeling is considered one of the major strengths of CGE models, they have been employed in the examination of a wide-range of globalization issues. These include direct foreign investment (Verikios and Zhang, 2000), international technology spillovers (van Meijl and van Tongeren, 1999, and Lejour and Nahuis, 2000) and international financial crises (Adams, 1998b, and Adelman and Yeldan, 2000). Other CGE studies examine the effects of global climate change (Tsigas et al. 1997) and multilateral agreements to contain global pollution (Conrad and Schmidt 1998, and Adams 2004). There have been quite a number of CGE studies regarding the global movement of people, including the effects of foreign migration (Peter 1993, and Goto 1996), the effects of international tourism (Adams and Parmenter 1995), the effects of international foreign students (Dixon et al. 1998, and Giesecke 2004a) and the effects of international-terrorism reducing international travel (Adams et al. 2001). The new economy, which is seen as a major driver of globalization, has also been subject to CGE analysis, including

---

[4] Indeed, Hertel et al. (1997) refer to their multi-country model as a multi-region model. This is essentially because countries can be aggregated into supra-national regions. However, a multi-country model can be easily converted to a multiregional (single-country) model (Jones and Whalley 1989).

studies of the information technology industry in general (Tiwari et al. 2003), telecommunications (Giesecke 2003) and eCommerce (Dixon and Rimmer 2005).

Given the width of the above studies of globalization—and it is by no means an exhaustive list of globalization areas covered by CGE studies—this chapter confines itself to just a number of areas in which CGE models have been developed to handle globalization issues. Two key developments will be explored. The first, covered in Section 11.3, is historical modeling which allows the forces of globalization to be examined during the (normally recent) past within a general equilibrium framework. The second, covered in section 11.4, deals with modeling regional labor market adjustment to globalization forces. In the concluding section we consider, *inter alia*, some features of globalization that are not universally captured by CGE models. First, however, we briefly outline, in the next section, the key basic features of CGE models, and some of the more common additions to that framework in regional and multiregional CGE models.

11.2 CGE Framework

There is no one specification of a CGE model—quite the reverse. All CGE models do, however, share some common core features. At the most basic level, this involves constrained optimization behavior both by producers and consumers, market-clearing equations that determine prices, and one or more factors being in inelastic supply. A typical basic regional CGE model might contain the following features. Producers in each regional industry would be assumed to choose their inputs to minimize costs subject to a production technology that might allow substitution between primary factors (labor, capital and land) and between geographical sources of supply for intermediate inputs. Regional industry supply curves are usually determined by the assumption of zero pure profits. Normally there is at least one representative household (one, or more, in each region in a multiregional model) that purchases goods and services in order to obtain an optimal bundle in accordance with the household's preferences and its disposable income. Most CGE models carry some treatment of investment. For instance, it might be assumed that investors supply financial capital to regional industries so as to maximize their rate of return. Single country CGE models might treat foreigners' demands via export demand functions that capture the responsiveness of foreigners to changes in the home country's regional industries' supply prices.[5] Prices are determined in CGE models by a set of market-clearing equations.[6]

---

[5] There are various alternative treatments of exports, the simplest being a fixed world price for exports (the small-country assumption). One common treatment incorporates a transformation function between a domestically-produced good sold domestically and the same good sold as an export.

[6] The existence of market-clearing equations does not preclude a treatment of unemployment or changes to the unemployment rate. The latter could be modeled by a "sticky" wages assumption.

Outside these basic features, there are a wide variety of different specifications to be found in regional CGE models. Some models are designed to be well-equipped to deal with a specific question or class of questions. Such models may be quite stylized in areas that are not thought to play a major role in the matter under examination, but might contain a quite sophisticated treatment of economic mechanisms considered to play a key role in determining results in the particular instance. On the other hand, some CGE models are general-purpose models that can be expected to have an overall well-developed theoretical specification and considerable industry and commodity disaggregation.[7] However, extensions of such models' theory are often made when this is considered necessary to properly model a particular question. CGE models can also be adapted for particular applications by altering their closure (swapping variables between exogenous and endogenous categories).

There is a large number of ways that CGE models are open to extension, and many examples of particular extensions. For instance, particular attention might be given to the modeling of certain sectors of the economy, to modeling the behavior of investors, to the computation of regional household disposable income (perhaps distinguishing region of ownership from region where returns are earned), to the treatment of government revenue and expenditure[8] (perhaps for two or more tiers of government), or to the explicit treatment of transport and other margin services, and so on.

In general, CGE models assume constant returns to scale and perfectly competitive markets. Cross-hauling is generally allowed for by assuming that domestic and imported goods of the same type are imperfect substitutes, an approach pioneered by Armington (1969). It may be, however, that economies of scale and imperfect competition in one or more markets are more appropriate assumptions for a globalized world. A considerable proportion of CGE models used to examine economic integration employ these alternative assumptions. However, the degree to which these assumptions alter results is questionable. Harris (1984), who pioneered these assumptions, and some others since, together have found they significantly affect the results of trade liberalization simulations. However, this has not been a general finding by those who have employed these assumptions, and appears to be a product of the particular pricing rule used (Abayasiri-Silva and Horridge, 1998). These assumptions have been little employed to date in regional CGE models.[9] We will return to this matter towards the end of this chapter. For

---

[7] Australia has a strong tradition in large-scale general purpose CGE models, such as the national models ORANI (Dixon et al. 1982) and MONASH (Dixon and Rimmer 2002), and the multiregional models, FEDERAL-F (Giesecke and Madden 2003a) and MMRF-GREEN (Adams et al. 2000).

[8] Normally, CGE models treat the determination of government expenditure in a fairly ad hoc way, such as linking its movement to the movement in another variable such as GDP or tax revenue. Some models do not treat public consumption separately at all, aggregating it with private consumption. Only occasionally (for instance, Groenewold et al. 2003) are governments treated as optimizing agents.

[9] Example of the use of scale economies and imperfect competition in regional CGE modeling are Kilkenny (1993) and Bröcker (1998). Imperfect competition is introduced via the

the moment, we concentrate on the two developments outlined in the last paragraph of the introductory section.

The developments to be discussed in the next two sections are both dependent on dynamic modeling. In recent years there has been a significant increase in dynamic CGE models (see for instance the models outlined in Harrison *et al*., 2000). However, Partridge and Rickman (1998) note that there are very few examples of dynamic regional CGE models. One notable exception at that time was the single-region three-sector model of Scotland by McGregor et al. (1996). A quite limited number of dynamic regional CGE models have appeared since. McGregor *et al* (1999) expanded their model to two regions. Another example of a multiregional dynamic CGE model is the 14-region single-product model of Korea by Kim and Kim (2002). These models are quite small scale. However, two Australian CGE models are large-scale and contain sophisticated dynamics, FEDERAL-F (Giesecke and Madden 2003a) and MMRF-GREEN (Adams et al. 2000).

Both of these models employ dynamic properties that have been styled on the national CGE model MONASH (Dixon and Rimmer 2002). Hence, FEDERAL-F, for instance, is a recursive-dynamic multiregional CGE model, linking a sequence of single-period equilibria via stock-flow relationships. The equilibria thus computed change through time as the values for the model's stock variables change. Flows in previous periods (such as investment, inter-regional migration, and government borrowings) influence the values for the endogenous variables computed in each period through their contribution to the values of the model's stock variables (such as capital, population, and government debt) in each period. FEDERAL-F also possesses the sophisticated closure options of MONASH.

There are four major closure options that allow MONASH-style models to be run in four modes: historical, decomposition, forecasting and policy deviation. These alternative modes will be employed in the simulations discussed in the next two sections. Brief descriptions of the modes are:

> *Historical mode* – is used to uncover the movement in certain unobservable variables during a period of (recent) history. This requires a closure that is very different from that normally employed in CGE models. Various variables (such as: real private consumption by commodity; real investment; output, employment and capital usage by industry; commodity exports and imports) that are normally endogenous, are reassigned to the exogenous category and shocked by historical values recorded in statistical sources. Corresponding normally-exogenous variables (such as: shifts in household preferences; investment/capital ratios; various technical change variables; shifts in foreign demand curves and domestic/import preferences) are swapped to the endogenous category.
>
> *Decomposition mode* – is used to explain those factors driving the observed changes over the period under examination. Most of the swaps undertaken for the historical closure are reversed, with the variables switched back to

---

Dixit and Stiglitz (1977) assumption of product differentiation; an idea much used in economic geography by Krugman and others (see for example, Fujita et al. 1999).

the exogenous category being shocked with the values they attained under the historical simulation. The values of variables under both modes are thus identical. However, the decomposition mode allows the identification of the relative impact on the structure of the economy of such factors as technology and preference changes, policy decisions and foreign shocks.

Forecast mode – is used to obtain a baseline forecast for industries, occupations and regions that are consistent with expectations for macroeconomic variables and historical trends uncovered in historical simulations. The closure involves exogenizing macroeconomic variables which are shocked with the results from macroeconomic forecasts. Price and quantity forecasts for exports can be imposed by endogenizing shifters on export demand curves. Technology and taste variables are given shocks consistent with historical trends, while policy variables (such as tax rates) can be shocked in line with announcements and expectations.

Policy-deviation mode – The closure is changed to one suitable for modeling the particular policy scenario under consideration. Thus macroeconomic variables are reassigned to the endogenous category and the corresponding macroeconomic structural variables swapped into the exogenous category are shocked with their results from the forecast simulation. The results deviate from the baseline, however, by the imposition of the policy shock under examination. This deviation gives the effect of implementing the policy.

11.3 Examining Globalization

In this section we examine the power of historical/decomposition simulations to uncover underlying globalization forces. To date, there is only one example at the sub-national level of historical modeling aimed at uncovering underlying driving forces on economic variables. This is provided by Giesecke (2002) who employs the FEDERAL-F model to examine the reasons behind the growing disparity between the performance of the Tasmanian economy and that of the rest of Australia over the period 1992-93 and 1998-99. He found that there was no single cause of the growing disparity. However, he was able to rule out some commonly-postulated reasons, such as the effects of the federal and state governments' fiscal policies, real wage rate changes and demographic reasons, as the contribution of these variables to the increase in disparity in gross regional products was small. Some of the explanation for the increased disparity was associated with Tasmania's performance in a globalized world. Firstly, while Tasmania's productivity grew strongly, it was outpaced in this area by the rest of Australia. Secondly, and more significantly, while investors lowered their required rate of return from investment in both regions, this was less so for Tasmania than the rest of Australia. The relatively poorer Tasmanian investment prospects that are implicit in this result may reflect a perception by investors of greater risk when investing in Tasmania during this period. In a world of globalized capital, factors such as an uncertainty on state government policy on resource development can have significant effects on a region's economic performance.[10] Giesecke's results also show this relatively poorer performance in attracting investment as being compounded by a lower capital stock in subsequent years.

While Giesecke uncovered factors relevant to globalization, it is possible to use historical/decomposition modeling to directly examine globalization questions. An example of this is Dixon et al. (2000) who conduct simulations with the national MONASH model to explain the rapid growth in Australia's trade, particularly its intra-industry trade, during the period 1987 to 1994. The authors note that previous studies of trade growth generally regressed measures of trade against various variables that were thought to represent possible explanatory factors. Such variables were often aggregate in nature and frequently used proxies for potential economic factors. Dixon et al. (2000) argue that their historical modeling approach overcomes criticisms of the multiple regression approach by allowing the relationship between trade measures and primary variables, such as technological change, to be estimated directly. This is done within a general equilibrium framework, which the authors say allows "… clear-cut explanations of results, distinguishing between demand and supply effects" (Dixon et al. 2000).

Dixon *et al.* explain the changes in macroeconomic and trade variables in terms of ten groups of causal factors (see Table 11.2), each of which aggregate a group of exogenous variables in the decomposition simulation. Results are provided for the effects of each factor with the other factors unchanged. Adding across the contributing factors gives the total changes in the macroeconomic and trade variables over the period.[11] A feature of their paper is the careful way in which the authors proceed through each of the ten factors, carefully explaining the model mechanisms that generate the effects for that factor. Here we simply note their results for the trade variables.

Over the period modeled by Dixon et al. (2000), total trade volumes for Australia increased by almost 48 per cent compared with a real GDP increase of a little over 18 per cent. Australia's net trade increased by 41 per cent for the period, while intra-industry trade increased by just over 61 per cent. Dixon et al. define net trade as $\sum_i \left| X_i - M_i \right|$, where X_i and M_i are exports and imports of commodity i respectively, and intra-industry trade as $2\sum_i \min(X_i, M_i)$. What happens to these trade measures is dependent on the particular commodities that experience the greatest growth in exports and which experience the greatest growth in imports. This is demonstrated in table 11.1 which shows the effects on the trade measures of increasing either exports or imports for just the single commodity i by itself. The first row of the table is for a commodity which might be thought of as a traditional export (with exports greater than imports), while the third row might be

---

[10] This is consistent with explanations by Nixon (1997) and others that a period of minority and compromise governments during the period was seen as posing a significant risk to investors.

[11] The decomposition algorithm of Harrison et al. (2000) ensures that the sum of the contributions made by the 10 sets of exogenous variables is exactly equal to the total effect. This is explained further in the Appendix to this chapter.

thought of non-traditional exports (greater imports of *i* than there are exports). Looking at the export column, it can be seen that an increase in traditional exports does not affect the measure of intra-industry trade, but an increase in non-traditional exports will act to increase intra-industry trade.

Table 11.1. Effects on trade measures of an increase in exports/imports of a commodity.

Base-year relationship	Increase in X_i	Increase in M_i
$X_i > M_i$	NT ↑ , IIT unchanged	NT ↓ , IIT ↑
$X_i = M_i$	NT ↑ , IIT unchanged	NT ↑ , IIT unchanged
$X_i < M_i$	NT ↓ , IIT ↑	NT ↑ , IIT unchanged

NT indicates net trade; IIT indicates intra-industry trade

In Table 11.2 we show Dixon et al.'s results for just six variables. In their article the authors present results for 22 macro and trade variables in order that they can show the model mechanisms involved in generating the effects of each of the ten factors. Here we merely present results for GDP and five trade variables and simply make note of those factors that made the major contributions to the trade results. From Table 11.2, we see that, while the momentum of the economy and an increased required rate of return (in most industries) were trade contractionary, the remaining factors had a positive impact on total trade.

It can be seen that from Table 11.2 that the factors which gave the major boosts to the traditional exports (technical change and employment growth) were also the factors that yielded the bulk of the positive contributions to net trade. Those factors also stimulated non-traditional exports, and consequentially intra-industry trade. Foreign factors (export demands and import prices) also boosted non-traditional exports, but caused a substantial decrease in traditional exports as world prices for the primary products that Australia traditionally exports declined relative to the price of the nation's less traditional manufactured and service exports.

Giesecke (2004b) has undertaken historical simulations for the period 1996/97 to 2001/02. While that paper did not look at intra-industry trade, we compute that variable for this chapter. Contrasting to the earlier seven-year period studied by Dixon *et al.*, the rapid growth in trade and intra-industry trade as a proportion of GDP was not in evidence for this more recent five-year period.

Table 11.2. Macroeconomic and trade variables: Dixon *et al.* (2000) decomposition of changes from 1987 to 1994(a)

1	2	3	4	5	6	7	8	9	10	11	12
Percentage changes	Momentum(b)	Foreign demands and import prices	Protection	Technical change	Import-domestic preferences	Household tastes	Employment growth	Required rates of return	Export supply curves	Other factors	Total
Real GDP	0.5	1.7	0.3	6.4	0.1	-0.2	9.1	-0.8	0.0	1.2	18.3
Traditional exports	-13.3	-26.7	3.4	44.1	7.2	0.4	17.7	-7.2	1.1	4.5	31.2
Non-traditional exports	-23.3	17.7	6.9	22.0	12.2	1.3	27.2	-4.4	10.1	5.2	75.0
Total Trade	-7.3	2.9	3.4	27.6	7.6	0.5	14.0	-3.3	2.2	0.1	47.6
Net Trade	-4.2	-3.9	3.0	33.4	5.7	-0.2	11.5	-3.1	0.2	-1.6	40.7
Intra-industry trade	-13.3	16.2	4.2	16.1	11.2	2.0	19.0	-3.7	6.1	3.4	61.1

(a) This table reproduces only 6 of the 22 rows in Table II of Dixon, *et al.* (2000).

(b) Shows what would have happened to the Australian economy over period if there had been no shocks to the variables summarized by the other nine factors. It is an outcome of the base-year relationship between capital, savings and foreign liabilities. For further details, see Dixon et al. (2000) or Giesecke (2004b).

Table 11.3 Decomposition of results for GDP and trade variables from 1996/97 to 2001/02

	1	2	3	4	5	6	7	8	9	10	11
Percentage changes	Momentum(a)	Foreign demands and import prices	Protection	Technical change	Import-domestic preferences	Household tastes	Employment growth	Required rates of return	Export supply curves	Other factors	Total
Real GDP	0.1	0.8	0.0	18.6	0.0	-1.1	11.2	-4.8	-1.0	0.9	24.8
Import volumes	0.5	0.7	0.6	1.9	8.2	-0.6	9.1	-1.8	-1.4	5.2	22.4
Export volumes	-4.9	-4.0	0.8	15.5	10.5	-1.5	18.5	1.0	-1.8	-9.0	25.0
Traditional	-3.4	12.7	0.6	12.9	8.7	-2.3	15.1	-3.3	-14.1	-3.4	23.5
Non-traditional	-7.1	-20.4	1.2	18.1	13.4	-0.8	23.2	6.8	11.9	-16.2	30.1
Tourism	-5.4	-32.5	0.7	23.7	9.0	-2.0	17.8	-2.1	12.6	-13.5	8.3
Intra-industry trade	-3.6	-19.5	0.7	6.4	13.4	-0.6	15.9	3.7	4.6	-5.6	15.4

(a) See footnote (b) Table 11.2.

Table 11.3 shows that over the five years to 2001/02, export and import volumes (rows 3 and 2) grew at about the same rate as growth in real GDP (row 1), while growth in intra-industry trade (row 7) has been slower than growth in total trade. Two columns, describing the effects of changes in foreign trading conditions (column 2) and the effects of domestic changes in technical efficiency (column 4) account for much of the relatively slow growth in intra-industry trade.

While the growth in total trade and intra-industry trade is relatively closely matched for most columns, in columns 2 and 4 the growth in intra-industry trade is substantially less than the growth in total trade. The economic forces summarized by columns 2 and 4 are elaborated upon below.

As was the case for Table 11.2, column 2 of Table 11.3 summarizes the impacts of movements in foreign currency import prices and shifts in foreign demands for Australian exports. These shifts were slightly favorable to Australia, lifting the terms of trade and with it the capital / labor ratio and real GDP (row 1). While changes in foreign trading conditions had little impact on total exports (row 3), the composition of exports changed significantly (rows 4 – 6). Reversing the trend observed in Dixon, *et al.* (2000), shifts in foreign demands for traditional export commodities (particularly mining) over the study period were larger than those for non-traditional and tourism exports. Since the latter commodities experienced relatively weak shifts in their foreign demand schedules, the real exchange rate appreciation generated by the strong positive shifts in foreign demands for traditional exports caused a contraction in foreign demand for Australian tourism and non-traditional exports. As a result, non-traditional export volumes contracted by 20.4 per cent and tourism exports contracted by 32.5 per cent. Overall export volumes contracted by only 4.0 per cent because traditional export volumes increased by 12.7 per cent. Non-traditional and tourism exports consist of mainly manufacturing and service exports, commodities for which intra-industry trade tends to be relatively high. On the other hand, the set of traditional export commodities is heavily weighted towards agricultural and mining exports, commodities for which intra-industry trade tends to be relatively low.

Column 4 isolates the effects of changes in many commodity- and industry-specific variables describing the technical efficiency of the domestic economy.[12] For intra-industry trade, the most important of these technical change variables are those relating to the per unit requirements of each commodity for current production, capital formation, and consumption purposes. Overall, the changes in technical sectoral variables lifted economy-wide productivity and hence real GDP (row 1). However growth in import volumes (row 2) was remarkably slow: while gains in technical efficiency lifted real GDP by 18.6 per cent, they lifted real imports by only 1.9 per cent. This had a damping effect on growth in intra-industry trade. Changes in commodity-using technical efficiency tended to be against the usage of commodities that are relatively import intensive. In particular, the results of the historical simulation suggested that requirements per unit of output for import-intensive commodities such as *clothing*, *motor vehicles*, *electronic equipment*, *agricultural machinery*, *other machinery*, and *basic non-ferrous metal products* fell over the study period.

To date the above analysis of intra-industry trade has been carried out only at the national level. However, the same approach could be used with a model such

---

[12] In particular, column 4 isolates the impacts of movements in: primary factor productivity by industry; per-unit input requirements of each commodity for use in current production and capital formation; all-input-using technical change in current production by industry; and, all-input-using technical change in capital formation by industry.

as FEDERAL-F to analyze the factors underlying changes in intra-industry trade at the regional level. To date, however, the required historical modeling framework has been established only for a two-region model (Giesecke, 2002). That author is currently working on establishing the required bottom-up multiregional historical modeling framework to cover the eight state/territory regions of Australia. In this chapter, we present an intermediary modeling exercise, a top-down regional decomposition of the 1996/97 to 2001/02 results discussed above.

The top-down decomposition theory is the ORANI Regional Equation System (ORES) of Dixon et al. (1982). The central assumption of ORES is that industries can be divided into two categories: national and local. National industries are those producing commodities that are readily traded between regions, while local industries produce commodities that are primarily used only within the local region. In the ORES theory, the output of national industries at the regional level is determined independently of changes in regional activity, with the percentage change in the activity of national industry (*j*) located within region (*r*) being set equal to the percentage change in the Australia-wide activity of industry (*j*). With the changes in the regional activity of national industries thus calculated, the activity levels of local industries are determined by regional market clearing conditions for local commodities. This acts to capture local multiplier effects.

Table 11.4 presents the contributions of the ten sets of exogenous factors listed in this table to the growth in real gross state product (GSP) of Australia's eight states and territories. In explaining movements in real GSP, we must include an additional decomposition factor, which is summarized in column 11. This factor (regional industry growth shifts) takes account of the fact that the top-down decomposition theory on its own will not produce results for real GSP that match observed real GSP. Consider Victoria; Real GSP in this state grew by 27.2 per cent over the study period (column 12). The national shocks summarized in columns 1 – 10 explain, via the ORES theory, only 26.0 percentage points of growth. In the historical simulation, we require faster growth in Victorian national industries than that generated by the standard ORES theory. We allow this to occur via an endogenous shift variable on national industry growth rates within the region. Likewise, for the remaining regions, we endogenously determine region-specific shifts on national industry growth rates to accommodate the exogenous determination of real GSPs at their historically observed values. In the decomposition simulation, these shift variables are exogenous, and their contribution to each region's growth is summarized in column 12. The contributions of these industry growth shifts tend to be small relative to total growth in GSP (compare columns 11 and 12). Two exceptions are Queensland and Tasmania. Industries in Queensland tended to grow faster than the national average. This reflects shocks specific to Queensland (and thus not captured by the top-down method) such as the strong inward flow of retirees and the state's favorable investment prospects. On the other hand, industries in Tasmania grew more slowly than the national average. This is consistent with Giesecke (2002) who found strong adverse shifts in variables describing Tasmanian productivity and investment prospects over this period. Movements in region-specific factors such as these cannot be captured by the top-down method other than via exogenous shifts in regional industry growth rates, the effects of which are summarized in column 11.

Table 11.4. Decomposition results: real gross state product, 1996/97 to 2001/02.

	1	2	3	4	5	6	7	8	9	10	11	12
Percentage changes	Momentum(a)	Foreign demands and import prices	Protection	Technical change	Import-domestic preferences	Household tastes	Employment growth	Required rates of return	Export supply curves	Other factors	Regional industry growth shifts	Total
NSW	0.3	0.5	0.1	20.3	0.1	-1.1	10.8	-5.3	-0.4	1.6	-3.6	23.2
VIC	0.1	-0.2	-0.1	20.6	-0.5	-0.6	11.5	-5.1	-0.2	0.3	1.2	27.2
QLD	0.2	2.6	0.1	16.2	0.1	-1.2	11.5	-5.2	-2.1	0.9	6.4	29.5
SA	0.1	0.4	-0.1	18.5	-0.6	-1.2	11.0	-4.9	0.1	-0.1	-1.0	22.0
WA	-0.6	1.6	0.2	12.3	1.1	-1.5	12.0	-2.7	-3.7	0.0	2.1	20.8
TAS	0.2	0.3	0.1	12.0	0.5	-1.5	9.3	-2.0	-0.6	-0.2	-7.2	10.9
NT	-0.3	0.5	0.2	12.9	1.8	-1.2	12.3	-1.0	-4.0	0.5	2.8	24.5
ACT	1.2	1.8	0.0	19.7	0.3	-1.1	9.6	-5.3	-0.2	5.1	-5.5	25.6

(a) See footnote (b) Table 2.

Note: NSW (New South Wales); VIC (Victoria); QLD (Queensland); SA (South Australia); WA (Western Australia); TAS (Tasmania); NT (Northern Territory); ACT (Australian Capital Territory)

Growth in real GSP was largely driven by growth in labor supply (column 7) and improvements in productivity (column 4). Changes in foreign trading conditions (column 2) and domestic trade policy (column 3) had relatively little impact on growth in real GSP. Changes in foreign trading conditions tended to favor growth in the Australian Capital Territory, Queensland and Western Australia. For the latter two regions, this reflects the relatively high shares of GSP accounted for by industries producing traditional export commodities. The Australian Capital Territory, on the other hand, benefited from strong growth in foreign demands for education exports and technical services exports, activities with a relatively high representation in the region's GSP. Australia continued to reduce its already relatively low tariffs over the study period (column 3). This contributed to small (-0.1 per cent) reductions in the sizes of the two regions most heavily dependent on import-competing industries (Victoria and South Australia) and expansions (+0.2 per cent) in the sizes of the two regions most dependent on export industries (Western Australia and the Northern Territory).

11.4 Regional Labor Market Adjustment

11.4.1 The RELMAI Model

While the forces of globalization may be seen as generally bringing long-term benefits, two possible negative aspects are that inequality may be increased and there could be short-term disruptions in the labor-market as an economy adjusts to a globalization shock. CGE modeling has been used to examine these possible downsides on regions. For instance, Ferreira Filho and Horridge (2004) use a microsimulation attachment to a CGE model of Brazil to examine the effects of the formation of the proposed Free Trade Area of Americas on poverty and income distribution in 27 Brazilian regions. Giesecke and Madden (2003b) examine the regional labor market adjustment effects of a "globalization" policy on some small regions in Australia. In this section, we look at the second of these two applications, with a different focus than that provided in the original article. In particular, we examine the theoretical structure of the regional labor market adjustment model, before proceeding to use the model to obtain a measure of the extent of labor market disruption that follows a "globalization" shock.

The model was designed to be run in conjunction with the FEDERAL-F model, using the results from the latter model to compute a regional labor market adjustment index (RELMAI). The index is closely based on the LILI (labor input loss index) model of Dixon and Rimmer (2002). The latter model uses the sectoral, geographic, and occupational dimensions of the output on job gains and losses from the MONASH model to calculate an index of labor market adjustment costs for the nation.[13] The RELMAI model applies the LILI methodology to the regional level, providing a means of calculating indices of labor market adjustment costs for individual regions.[14]

RELMAI is an index that is computed on the basis of worker flows. Worker flows are defined by Borland (1996) as occurring when workers move between labor market states (i.e. employed, unemployed or not in the labor force) or between job categories (i.e. occupations and industries at our level of aggregation). For the RELMAI we include moves between regions as also constituting worker flows. A cost in terms of lost labor inputs to a particular region is associated with each worker flow.

The FEDERAL-F results for changes in employment (by occupation, industry and region) that form the input to the RELMAI, are in terms of changes in average levels of employment between year t and year *t*+1. For the purposes of computing RELMAI we assume these changes are from employment in the middle of year t to the middle of year *t*+1. In line with, Dixon and Rimmer (2002) we view this

---

[13] As noted previously the MONASH regional results are produced by a top-down decomposition of national results, while FEDERAL-F is a bottom-up model at the state level. However, the sub-state regional results from FEDERAL-F are generated via a top-down decomposition of the model's state results to sub-state regions.

[14] In the present version of the RELMAI model (we use this term for both the index and the model), the weighted average of the regional indices is equal to the LILI index.

mid-year level of employment simply as the employment in the relevant year. We therefore view worker flows also in terms of changes between labor market states (hereafter defined as also encompassing regions, and for the employed, job categories). Thus a person employed in a particular occupation in a particular region in year *t*, can in year *t*+1 have remained in the same job category (i.e. same occupation, industry and region), or moved to a different occupation/industry/region, become unemployed or left the labor force. Similarly, a person who was unemployed last year might remain unemployed this year, have found a job, or have left the labor force. A person not in the labor force may remain there or join the labor force in either an employed or unemployed capacity. Note that for the RELMAI model we include in the definition of flows persons who remain in the same labor market state.[15]

RELMAI is formally defined for region *r* for the year *t* to *t*+1 as a weighted sum of the worker flows associated with that region[16] expressed as a percentage of the labor force in region *r* in year *t*:

$$RELMAI^r(t,t+1) = \{100/LF^r(t)\}\textstyle\sum_v^{20} C_v^r WF_v^r(t,t+1) \tag{11.1}$$

where $LF^r(t)$ is the number of persons in region *r*'s labor force in year t, the C_v^r s are weights and $WF_v^r(t,t+1)$ are the worker flows between labor market states *v*. In order to compute equation (11.1) we first have to determine the WF_v^r s and assign a set of estimated values for the weights.

The remainder of the description of the RELMAI model explains how the number of workers in each of the twenty labor-market flow categories on the right-hand side of equation (11.1) are calculated. We assign an individual variable name to each of the flows. For instance, WF_1^r is named $E\_ESC\_SR_r$ to indicate employed (E) workers in year t who remain employed in the same occupation/industry category (ESC) in the same region (SR) in year *t*+1. Each of these twenty types of worker flows are described in table 11.5 which also contains the numerical value of the associated weight used, and the reason for assigning that value.

Table 11.5. Worker-flow variables in RELMAI (year *t* to *t*+1).

Variable		Weight	
Name	Description	Value	Explanation
$E\_ESC\_SR_r$	Number of employed workers remaining employed in same category in same region.	0	Such workers do not contribute to labor market adjustment costs.
$E\_ESC\_OR_r$	Number of employed workers remaining employed in the same category in some other	n.a.	These flows included in $E\_EOC\_OR_r$;

---

[15] This approach ensures that the cost of persons remaining unemployed over the year *t* to *t*+1 is taken into account in computing lost labor inputs.

[16] Note that when workers move between regions, it is assumed that the flow (and consequentially its adjustment cost) is associated with the region of origin.

	region.		$E\_ESC\_OR_r$ set exogenously at zero.
$E\_EOC\_SR_r$	Movement of employed workers to employment in another occupation/industry category within the same region.	0.25	These workers are assumed to require 3 months re-training involving loss of their productive capacity for that period.
$E\_EOC\_OR_r$	Movement of employed workers to employment in another occupation / industry category and/or another region.	0.25	As above.
$E\_U\_SR_r$	Workers moving from employment to unemployment while remaining within the region in which they were originally employed.	0.5	Assume flow from employment to unemployment occurs evenly over the year.
$E\_U\_OR_r$	Workers moving from employment to unemployment in some other region.	0.5	As above.
$E\_NLF\_SR_r$	Workers moving from employment to being outside of the labor force within their original region.	0	Retirement rate is assumed to be independent of the shocks to the model.
$E\_NLF\_OR_r$	Workers moving from employment to not in the labor force while also changing their region of residence.	0	As above.
$U\_NLF\_SR_r$	Workers moving from being unemployed in region *r* to being outside of the labor force in region *r*.	0.5	Assume flow occurs evenly over the year.
$U\_NLF\_OR_r$	Workers moving from unemployment in region *r* to being outside of labor force in some other region.	0.5	As above.
$U\_E\_SR_r$	Workers moving from unemployment to employment in the same region.	0.75	Reflects 2 assumptions: this flow occurs evenly over the year; and employment of these workers requires 3 months re-training.
$U\_E\_OR_r$	Workers moving from unemployment to employment in another region.	0.75	As above.
$NLF\_E\_SR_r$	Workers moving from being outside of the labor force in region *r* to being employed in region *r*.	0.25	Assume these workers require re-training, involving loss of potential labor input of 0.25 worker years.
$NLF\_E\_OR_r$	Workers moving from being outside of the labor force in region *r* to being employed in some region other than *r*.	0.25	As above.
$U\_U\_SR_r$	Workers unemployed in region *r* who remain unemployed in the same region.	1.0	Reflects loss of their potential labor input for a full year.
$U\_U\_OR_r$	Workers unemployed in region *r* in year t who remain unem-	1.0	As above.

	ployed in year t+1 but in some other region		
$NLF\_U\_SR_r$	Persons moving from outside of the labor force in region *r* to unemployment in region *r*.	0.5	Reflects assumption that flow from outside of the labor force to unemployment occurs evenly over the period.
$NLF\_U\_OR_r$	Persons moving from outside of labor force in region *r* to unemployment in some region other than *r*.	0.5	As above.
$NLF\_NLF\_SR_r$	Persons who are outside of the labor force in region *r*, and who remain so.	0	These persons are not part of the labor force at any time during the year.
$NLF\_NLF\_OR_r$	Persons outside of labor force in region *r*, who remain outside, while also moving region.	0	As above.

Equation (11.2) defines the number of workers that are employed in the same occupation / industry labor-market category (hereafter, “category’) o, j in the same region *r* in both the years *t* and *t*+1.

$$E\_ESC\_SR_r = \sum_{o \in OCC} \sum_{j \in IND} \text{MIN}[(1-\alpha^r)E\_T_{ojr} - VD_{ojr}\,; E\_T1_{ojr}] \tag{11.2}$$

where α^r is the fraction of the workforce that retires, E_T_{ojr} is employment in occupation *o* in industry *j* in region *r* in year *t*, VD_{ojr} is voluntary departures from category *ojr* to employment elsewhere, E_T1_{ojr} is employment in occupation *o* in industry *j* in region *r* in year *t*+1.

In equation (11.2), it is assumed that all the people who would like to continue in employment in category *o,j* in region *r* are able to do so, as long as there are sufficient jobs. Hence $E\_ESC\_SR_r$ is the lesser of two amounts: *(i)* The number of people in *r* who would *like* to continue in employment in category *o,j* between years *t* and *t*+1 - which is equal to employment in *o,j* in year *t* less retirements and voluntary departures; and *(ii)* actual employment in *o,j* in year *t*+1.[17]

Equation (11.3) defines the number of retrenchments in region *r*, E_TINVR_r. Retrenchments occur when there are insufficient positions in category *o,j* in region *r* in year *t*+1 to provide employment for those year *t* workers in category *o,j* in region *r* remaining after allowing for retirements and voluntary departures from *o,j*.

$$E\_TINVR_r = \sum_{o \in OCC} \sum_{j \in IND} \text{MAX}[0\,; (1-\alpha^r)E\_T_{ojr} - VD_{ojr} - E\_T1_{ojr}] \tag{11.3}$$

The total number of vacancies in *r*, $TVACR_r$, is defined by (11.4) as the difference between total employment in *r* in year *t*+1 and employment in *r* in year *t* adjusted for retirements and retrenchments. Equation (11.4) introduces total em-

---

[17] As a matter of definition, there are no region *r* workers employed in the same category *o,j* in another region in *t*+1, since a change in a worker’s region is treated as a change in the worker’s labor market category in (11.15).

ployment in r in year t, ER_T_r, and total employment in r in year t+1, ER_T1_r. These two variables are defined by (11.5) and (11.6).

$$TVACR_r = ER\_T1_r - [(1-\alpha^r)ER\_T_{r-} - E\_TINVR_r] \tag{11.4}$$

$$ER\_T_r = \sum_{o \in OCC} \sum_{j \in IND} E\_T_{ojr} \tag{11.5}$$

$$ER\_T1_r = \sum_{o \in OCC} \sum_{j \in IND} E\_T1_{ojr} \tag{11.6}$$

Equations (11.7) and (11.8) define the number of workers who are retrenched in r but who find re-employment in either region r, $TINV\_E\_SR_r$, or a region other than r, $TINV\_E\_OR_r$, respectively. In (11.9) it is assumed that no more than half of those retrenched workers who remain in region r find re-employment in r. Otherwise, it is assumed that retrenched workers have approximately double the probability of employment than that given by the simple share of their numbers in the total of all those retrenched workers, unemployed workers, and workers not in the labor force who are looking to take-up the available vacancies in r. Equation (11.7) assumes that potential workers in r are one quarter of those not in the labor force. Equation (11.7) also assumes that region r attracts a proportion of retrenched workers, the unemployed, and potential workers from regions other than r equal to its share in the national workforce.

$$TINV\_E\_SR_r = \mathrm{MIN}[\zeta_r E\_TINVR_r / 2\,;\, TVACR_r (2\zeta_r E\_TINVR_r) / LDEN_r] \tag{11.7}$$

where

$$LDEN_r = 2\zeta_r E\_TINVR_r + E\_Sh_r \sum_{i \neq r} (1-\zeta_i) E\_TINVR_i + \lambda_r UR\_T_r +$$

$$E\_Sh_r \sum_{i \neq r} (1-\lambda_i) UR\_T_i + 0.25[\rho_r NLFR_r + E\_SH_r \sum_{i \neq r} (1-\rho_i) NLFR_i]$$

and ζ_r is the proportion of retrenched workers who remain in r, λ_r is the proportion of region r's unemployed who stay in r, ρ_r is the proportion of those not in the labor force who remain in region r, E_Sh_r is region r's share of total unemployment, UR_T_r is the number of unemployed in region r, and $NLFR_r$ are people of working age who are not in the labor force.

$$TINV\_E\_OR_r =$$
$$\mathrm{MIN}[(1-\zeta)_r E\_TINVR_r / 2\,;\, \sum_{t \neq r} TVACR_t 2(1-\zeta_t) E\_TINVR_t / \sum_{t \neq r} LDEN_t] \tag{11.8}$$

Equation (11.8) defines $TINV\_E\_OR_r$. Consistent with (11.7), it is assumed that no more than half of those retrenched workers who leave region r find re-employment outside of the region. Otherwise, those retrenched workers who leave the region secure a proportion of the available vacancies outside of the region which is approximately equal to double their share of the total number of people looking to secure the available vacancies. Equation (11.9) defines the total number of retrenched workers from r who find re-employment, $TINV\_ER_r$, as the sum of those who find employment within *r* and those who find employment outside of *r*.

$$TINV\_ER_r = TINV\_E\_OR_r + TINV\_E\_SR_r \tag{11.9}$$

The number of unemployed in *r* that was introduced in (11.7) is calculated by (11.10) as the difference between the region's year *t* labor force, LFR_T_r ,and employment.

$$UR\_T_r = LFR\_T_r - ER\_T_r \tag{11.10}$$

Equation (11.7) also introduces region *r*'s share of total employment, and the number of people not in the labor force in region *r*. These are calculated by (11.11) and (11.12) respectively.

$$E\_Sh_r = ER\_T_r / \sum_k ER\_T_k \tag{11.11}$$

$$NLFR_r = WPOPR\_T_r - LFR\_T_r \tag{11.12}$$

where $WPOPR\_T_r$ is the region's working age population.

Equation (11.13) defines the total number of people moving from one employment category to another employment category within *r* as the total number of retrenched region *r* workers who find re-employment in *r*, plus the total number of region *r* workers making voluntary moves between employment categories in region *r*, less those voluntary region *r* movers who are moving between the same employment categories in region *r*.

$$E\_EOC\_SR_r = TINV\_E\_SR_r + \sum_{o \in OCC} \sum_{j \in IND} \sum_{u \in OCC} \sum_{k \in IND} V^{o,j,r}_{u,k,r} - \sum_{i \in OCC} \sum_{t \in IND} V^{i,t,r}_{i,t,,r} \tag{11.13}$$

where the $V^{o,j,r}_{u,k,t}$ s are voluntary moves from employment in category *u*,*k*,*t* to employment in category *u*,*k*,*t*.

Equation (11.14) defines the number of region *r* workers who move to another employment category in another region. This is equal to the number of retrenched region *r* workers who find employment outside of *r*, plus the sum of all region *r* workers making voluntary moves, less those region *r* workers making voluntary moves to employment categories within *r*.

$$E\_EOC\_SR_r = TINV\_E\_OR_r + \sum_o \sum_j \sum_u \sum_k \sum_t V^{o,j,r}_{u,k,t} - \sum_o \sum_j \sum_u \sum_k V^{o,j,r}_{u,k,r} \tag{11.14}$$

Equation (11.15) defines the number of region *r* workers moving to another employment category, E_EOCR_r , as the sum of those moving to another employment category within *r* and those moving to another employment category outside of *r*.

$$E\_EOCR_r = E\_EOC\_OR_r + E\_EOC\_SR_r \tag{11.15}$$

Equation (11.16) defines the number of people who move from employed in *r* to unemployed in *r* as the difference between those region *r* workers who are made redundant and stay in *r* and those region *r* workers who are made redundant and find re-employment in *r*.

$$E\_U\_SR_r = \zeta_r E\_TINVR_r + TINV\_E\_SR_r \tag{11.16}$$

Equation (11.17) defines the number of people who move between being employed in *r* in year t and being unemployed in a region other than *r* in *t*+1. This is equal to the difference between those region *r* workers who are made redundant and leave *r*, and those region *r* workers who are made redundant but find re-employment outside of *r*.

$$E\_U\_OR_r = (1-\zeta_r)E\_TINVR_r + TINV\_E\_OR_r \tag{11.17}$$

The total number of region *r* workers who move from employment to unemployment, E_UR_r, is defined by equation (11.18):

$$E\_UR_r = E\_U\_SR_r + E\_U\_OR_r \tag{11.18}$$

The number of people moving from being employed in *r* in *t*, to being outside of the labor force in *r* in *t*+1 is defined in (11.19) as equal to the number of retirements in *r*, multiplied by the proportion of newly retired persons who remain in *r*.

$$E\_NLF\_SR_r = \alpha_r \kappa_r ER\_T_r \tag{11.19}$$

where κ_r is the proportion of newly retirees who stay in their region.

The number of those employed persons who retire and go to live in another region is defined in equation (11.20) as equal to the number of retirements in *r*, multiplied by the proportion of newly retired persons who leave *r*.

$$E\_NLF\_OR_r = \alpha_r(1-\kappa_r)ER\_T_r \tag{11.20}$$

Equation (11.21) defines E_NLFR_r, the total number of people moving from being employed in *r* to being outside of the labor force either in *r* or some other region.

$$E\_NLFR_r = E\_NLF\_SR_r + E\_NLF\_OR_r \tag{11.21}$$

Equations (11.22) to (11.24) deal with the number of people who move from being unemployed to being outside the labor force. Equation (11.22) defines those in this category who stay in *r* as equal to the number of unemployed who retire multiplied by the proportion of new retirees who remain in *r*. Equation (11.23) deals analogously with those who leave *r*. Equation (11.24) defines the total number of unemployed people who retire from the labor force, U_NLFR_r, as the aggregate of the previous two variables.

$$U\_NLF\_SR_r = [\alpha \times UR\_T_r] \times \kappa_r \tag{11.22}$$

$$U\_NLF\_OR_r = [\alpha \times UR\_T_r] \times (1-\kappa_r) \tag{11.23}$$

$$U\_NLFR_r = U\_NLF\_SR_r + U\_NLF\_OR_r \tag{11.24}$$

Equation (11.25) defines the number of unemployed people in *r* who move to employment in region *r* as the number of jobs in *r* available to the unemployed and those not in the labor force multiplied by both the proportion of these jobs available to the unemployed, γ, and the proportion of the unemployed who remain in *r*. The number of jobs in *r* available to the unemployed and those not in the labor force is equal to the excess of employment in *t*+1 over the number of people employed in region *r* in year t who continue to be employed in *t*+1.

$$U\_E\_SR_r = \mathrm{MAX}[0\,;\gamma\lambda_r(ER\_T1_r - (1-\alpha^r)ER\_T_r + E\_UR_r) \tag{11.25}$$

The number of unemployed workers in *r* who find employment in a region outside of *r* is defined in (11.26). For any given region other than *r*, say *k*, region *r* workers are assumed to obtain a proportion of the jobs in k that are available to unemployed workers from outside *k*, $(1\text{-}\lambda_k)\gamma(ER\_T1_k - (1-\alpha^k)ER\_T_k + E\_UR_k)$, which is equal to region *r*'s share in the total number of unemployed workers in all regions other than k, $UR\_T_r \big/ \sum_{h\neq k} UR\_T_h$.

$$\begin{gathered} U\_E\_OR_r = \sum_{k\neq r}(UR\_T_r \,/ \sum_{h\neq k} UR\_T_h) \times \mathrm{MAX}[0\,; \\ ((1-\lambda_k)\gamma)(ER\_T1_k - (1-\alpha)ER\_T_k + E\_UR_k)] \end{gathered} \tag{11.26}$$

Equation (11.27) defines the number of unemployed workers from *r* who find employment as the sum of those unemployed region *r* workers who find employment within *r* and those who find employment outside of *r*.

$$U\_ER_r = U\_E\_SR_r + U\_E\_OR_r \tag{11.27}$$

Equation (11.28) defines the number of people outside the labor force in *r* who move to employment in *r*. This is equal to the number of jobs in *r* available to the unemployed and those not in the labor force multiplied by both the proportion of these jobs available to those outside the labor force and the proportion of those not in the labor force who remain in *r*. The number of jobs in *r* that are available to the unemployed and those not in the labor force is equal to the excess of employment in *t*+1 over the number of people employed in region *r* in *t* who continue to be employed in *t*+1.

$$NLF\_E\_SR_r = \mathrm{MAX}[0\,;\rho_r(1-\gamma)(ER\_T1_r - (1-\alpha^r)ER\_T_r + E\_UR_r)] \tag{11.28}$$

Equation (11.29) defines the number of those people who are outside the labor force in *r* who find employment in a region outside of *r*. For any given region other than *r*, say *k*, potential region *r* workers are assumed to obtain a proportion of the jobs in *k* that are available to those not in the labor force from outside k, $(1-\rho)(1-\gamma)(ER\_T1_k - (1-\alpha^k)ER\_T_k + E\_UR_{k)}$, which is equal to region *r*'s share in the total number of people not in the labor force in all regions other than k, $NLFR_r \Big/ \sum_{h \neq k} NLFR_h$.

$$NLF\_E\_OR_r \;=\; \sum_{k \neq r} \left(\frac{NLFR_r}{\sum_{h \neq k} NLFR_h} \right) \times \mathrm{MAX}\Big[0\,;\,\big((1\text{-}\rho_k)(1\text{-}\gamma\gamma\big)\big(ER\_T1_k - (1-\alpha^k) \times ER\_T_k + E\_UR_k\big)\Big] \tag{11.29}$$

Equation (11.30) defines the total number of people from outside the labor force *r* who find employment as the sum of those potential region *r* workers who find employment within *r*, and those potential region *r* workers who find employment outside of *r*.

$$NLF\_ER_r \;=\; NLF\_E\_SR_r \;+\; NLF\_E\_OR_r \tag{11.30}$$

Equation (11.31) defines the number of unemployed workers in *r* who remain unemployed in *r*. This is equal to the proportion of unemployed people who remain in *r* multiplied by the total of the number of unemployed people in *r* in year *t* less those that left the unemployment pool by either finding employment or leaving the labor force.

$$U\_U\_SR_r \;= \lambda_r \times \big(UR\_T_r \text{ - } U\_ER_r\text{-}U\_NLFR_r\big) \tag{11.31}$$

Equation (11.32) defines the number of unemployed workers in *r* who find employment outside of *r*. This is equal to the proportion of unemployed people who leave *r* multiplied by the total of the number of unemployed people in *r* in year *t* less those that left the unemployment pool by either finding employment or leaving the labor force.

$$U\_U\_OR_r \;= (1-\lambda_r) \times \big(UR\_T_r \text{ - } U\_ER_r\text{-}U\_NLFR_r\big) \tag{11.32}$$

Equation (11.33) defines the total number of unemployed people in *r* who remain unemployed as the total of those who remain unemployed within *r* and those who remain unemployed outside of *r*.

$$U\_UR_r \;=\; U\_U\_SR_r \;+\; U\_U\_OR_r \tag{11.33}$$

Equations (11.34) - (11.36) define the number of people in region *r* who move from being outside of the labor force to being unemployed. The derivation of these equations proceeds in a number of steps. First, we note the following relationship between the number of people not in the labor force in *t*+1 and the movement of people between various labor market categories:

$$NLFR\_T1_r = NLFR_r\text{- } NLF\_UR_r\text{- } NLF\_ER_r \;+\; E\_NLFR_r \;+\; U\_NLFR_r$$

Noting that:

$E\_NLFR_r + U\_NLFR_r = \alpha^r LFR\_T_r$, and $NLFR_r + LFR\_T_r = NLFR\_T1_r + LFR\_T1_r$, we can re-write the above expression as:

$NLF\_UR_r = LFR\_T1_r\text{-}(1\text{-}\alpha^r)LFR\_T_r \text{ - } NLF\_ER_r$,

which defines the number of people in *r* who move from outside of the workforce to unemployment. A proportion of these people, λ_r, are assumed to remain within region *r*, with the remainder departing to another region. This provides (11.34) and (11.35), while (11.36) is the usual aggregation, determining NLF_UR_r, the number of persons not in the labor force in *r* who become unemployed somewhere.

$$NLF\_U\_SR_r = \lambda_r \left[\left(LFR\_T1_r\text{-}(1\text{-}\alpha^r)LFR\_T_r\right) \text{- } NLF\_ER_r\right] \tag{11.34}$$

$$NLF\_U\_OR_r = (1-\lambda_r)\left[\left(LFR\_T1_r\text{-}(1\text{-}\alpha^r)LFR\_T_r\right) \text{- } NLF\_ER_r\right] \tag{11.35}$$

$$NLF\_UR_r = NLF\_U\_SR_r + NLF\_U\_OR_r \tag{11.36}$$

Equation (11.37) defines the number of people in *r* who were not in the labor force in period *t* and who remain not in the labor force in *r* in period *t*+1. This is equal to the number of people in *r* who were not in the labor force in period *t*, less those who moved into the labor force in year *t*+1, multiplied by the proportion of people not in the labor force who remain in region *r*.

$$NLF\_NLF\_SR_r = \rho_r \times [NLFR_r \text{ - } NLF\_ER_r \text{ - } NLF\_UR_r] \tag{11.37}$$

Equation (11.38) defines the number of people in *r* who were not in the labor force in *t* and who remain not in the labor force in *t*+1 but in a region other than *r*. This is equal to the number of people in *r* who were not in the labor force in *t*, less those who moved into the labor force in *t*+1, multiplied by the proportion of people not in the labor force who move to a region other than *r*.

$$NLF\_NLF\_OR_r = (1\text{-}\rho_r) \times [NLFR_r \text{ - } NLF\_ER_r \text{ - } NLF\_UR_r] \tag{11.38}$$

Equation (11.39) defines voluntary moves by workers from occupation *o* in industry *j* in region *r* to employment in occupation *u* in industry *k* in region *t*. The last term on the RHS of (11.39) defines the total number of vacancies in employment category *u*,*k*,*t* that are available to those not initially employed in the category. The second term on the RHS of (11.40) determines how these vacancies are distributed among workers departing from category *o*,*j*,*r*. The parameter $\Delta^{o,j,r}_{u,k,t}$ is a measure of the closeness of employment categories *o*,*j*,*r* and *u*,*k*,*r*. If these employment categories are very similar (e.g. very similar occupations in similar industries in the same region) then $\Delta^{o,j,r}_{u,k,t}$ will be close to or equal to 1. If, however, the employment categories are very dissimilar (e.g. occupations requiring widely different skills, located in different industries and in a different region) then $\Delta^{o,j,r}_{u,k,t}$

will have a value close to or equal to 0. The coefficient $S_{o,j,r}$ is the share of employment category *o,j,r* in total employment. This coefficient weights the share of the vacancies in category *u,k,t* that are filled by workers from category *o,j,r* towards *o,j,r*'s share of total employment. The parameter φ is set so that five per cent of employees make a voluntary move each year. Equation (11.40) defines total voluntary departures *o,j* from category, *r*.

$$V_{u,k,t}^{o,j,r} = \varphi\left(S_{o,j,r} \times \mathit{\Delta}_{u,k,t}^{o,j,r}\right) \times \mathrm{MAX}\left[0\,; E\_T1_{u,k,t} - (1-\alpha)E\_T_{u,k,t} + VD_{u,k,t}\right] \tag{11.39}$$

$$VD_{o,j,r} = \sum_{u \neq o} \sum_{k \neq j} \sum_{t \neq r} V_{u,k,t}^{o,j,r} \tag{11.40}$$

11.4.2 Regional Adjustment to Competition Policy

During the 1980s, Australia faced an apparent major external debt problem. This problem and a global trend in reforms saw Australia introduce financial market reforms and trade liberalization, before in the 1990s turning to the reforms of internal markets (Madden, 2006). In April 1995, the federal government and the eight state and territory governments agreed to implement a national competition policy (NCP). The pro-competitive reforms that formed the NCP were then implemented by the nine governments over the following years. A particular focus of NCP was reform of the Utilities industry. Using a two-region version of FEDERAL-F, Giesecke and Madden (2003b) undertook historical simulations and found a large jump in the rate of the Tasmania Utilities sector's productivity improvement following the 1995 NCP agreement, but only a modest increase in the Rest of Australia Utilities' productivity improvement.[18] This fitted in with the fact that Tasmania made a number of major reforms to the Utilities sector in 1995, while the Rest of Australia had undertaken substantial reforms to the sector prior to the NCP agreement.[19]

The RELMAI model was employed to examine labor market adjustment to the increase in the rate of productivity improvement post 1995. We shall refer to this

---

[18] Dividing the six-year historical period in two, we noticed a substantial increase in Tasmanian Utilities' average annual primary factor saving technical change (9 per cent) in the second half compared to an average productivity improvement of 6 per cent for the first three years. The corresponding figure for the Rest of Australia was 7.7 per cent for the latter period, compared with 7.0 per cent for the earlier period.

[19] The Utilities sector covers electricity, gas and water. Prior to competition reforms, these industries were dominated by government ownership and public provision. In 1995 Tasmania passed legislation commercializing the government business enterprise that was the sole producer and distributor of electricity in the state, allowed third party access to the electricity grid and established a Prices Oversight Commission for the industry. Reforms were also taken to put the State's water authorities on a more commercial footing.

as adjustment to NCP. We consider this NCP application a good example of how the regional adjustment impacts of a globalization shock might be modeled. As outlined in section 11.1, many aspects of the manifestations of globalization – e-commerce, reductions in trade barriers, international technology spillovers, and improvements in transport and communications services – can in part be implemented in CGE models as a set of productivity shocks. Part of the popular backlash against the forces of globalization no doubt has its origins in a fear of the adjustment costs that such productivity shocks might bring. However, such fears often ignore important general equilibrium effects. Any accounting of potential job losses in sectors directly affected by a globalization shock must also take stock of potential job gains elsewhere in the economy. Even in the directly affected sectors, the implications of job losses for total adjustment costs will depend on whether underlying prospects for the sectors are favorable or unfavorable. Government fiscal reactions add another dimension to the adjustment cost calculation.

The NCP simulations we report cover the period 1996/97 to 2003/04. Consistent with the discussion in section 11.2, we first undertook a base case simulation of FEDERAL-F which reflected history except that we excluded from the Utilities sector's productivity gains a part attributable to NCP. Results from this simulation were input to the RELMAI model to calculate base case adjustment costs. We then undertook two policy simulations of FEDERAL-F in which we included the NCP productivity gains. In the first policy simulation, we assume that regional governments use the proceeds from NCP to improve their budget positions.[20] In the second policy simulation, we assume that regional governments disburse the proceeds from NCP via payroll tax cuts. Otherwise, leaving aside these regional fiscal reactions, the shocks to the model in the policy simulations were identical to those in the base case simulation in all respects other than the rate of productivity improvement occurring in the Tasmanian and Mainland Utilities sectors. For these sectors, we increased the annual rates of primary factor productivity improvement by 3.0 percentage points and 0.7 percentage points respectively, for each of the years 1996/97 to 1998/99 inclusive. Results from the policy simulations were input to the RELMAI model to calculate policy-case adjustment costs. We report below the deviation in the regional adjustment cost indices in the policy cases away from their base case values in each year.

Figure 11.1 shows the deviation in the RELMAI in the case where regional governments use the additional Utilities dividends to decrease their short-run net borrowing requirements relative to base case. The RELMAI increases for all regions, reflecting the effects of job losses in the Utilities sectors and negative re-

---

[20] Our model takes explicit account of regional government capital ownership and price setting in the Utility sectors. Over the period covered by this simulation, regional governments owned the bulk of the capital in the directly affected sectors. In our modeling, we assume that the regional governments impose CPI-X pricing rules on the Utility sectors. This also reflects the institutional features of the study period. Under this pricing rule, the improvements in productivity in the Utilities sectors are manifested in the short-run by lower Utilities wage bills (secured via job losses) and corresponding higher dividend payments to the regional governments. How the regional governments then react to these higher dividend payments is the distinguishing feature of our two policy simulations.

gional multiplier effects from lower household consumption (due to job losses) and lower investment (since less capital is now required in the Utilities sector). The dispersion of the regional RELMAI impacts is driven by differences in regional industrial composition. Hobart is worst affected because it has the highest share of activity accounted for by Utilities. Northern and Southern are the least affected. These regions benefit not only by having relatively little Utilities activity but also by having relatively large amounts of Agricultural activity. Agriculture, and other trade-exposed sectors such as Mining, expand in this simulation, because the decline in Tasmanian consumption and investment spending (see above) requires that the region's balance of trade moves towards surplus. Regional adjustment costs peak in 1999/00. The fall in adjustment costs thereafter reflects our regional labor market assumptions. Regional real wages are sticky, adjusting gradually to return regional unemployment in the policy simulation to its base case level. The initial job losses cause regional unemployment to rise above the base case. Part of the adjustment is borne by inter-regional immigration, but much of it is borne by falling Tasmanian wages. This gradually returns Tasmanian unemployment, and with it employment, to its baseline path.

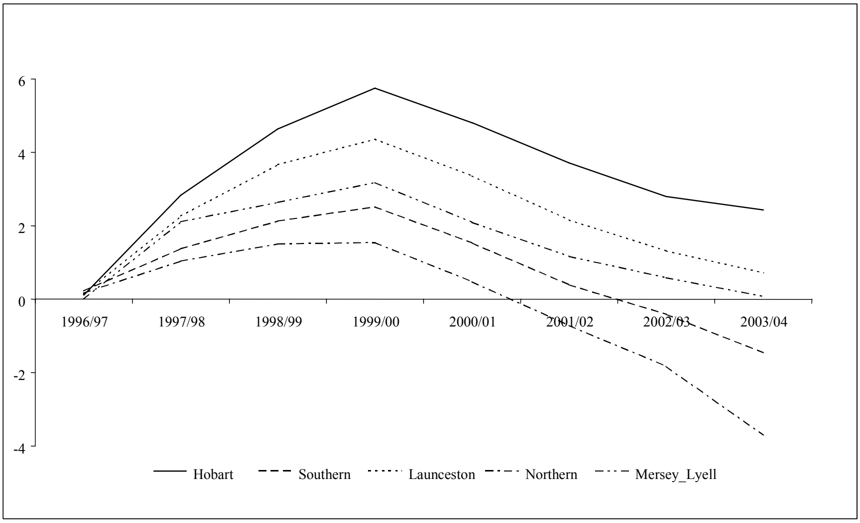


Fig. 11.1 Regional RELMAI indices (PSBRs endogenous): Percentage point deviation from the base.

The results from our first simulation demonstrate how a productivity gain can add to regional adjustment costs. However the regional government's fiscal reaction is central to this result: with the gains from the productivity improvement not passed on to the wider regional economy via either Utility price cuts or lower taxes, the short-run effects must be contractionary. This might reflect the case of a regional government seeking to immediately improve its long-term budgetary position by using the current profitability gains to avoid future tax increases. However we can bring this idea of an exogenous path for the regional government's budgetary position more clearly within the time frame of our simulation period. We do this in the second policy simulation (Figure 11.2). Here, we assume regional governments act to neutralize the impact of rising profits from Utilities on their net financing requirements by reducing regional payroll taxes. Essentially, we assume that government borrowing requirements are determined independently of the NCP shock, and that either the NCP gains thus allow taxes to be lower or that in the absence of the NCP gains taxes would have had to be higher. In this simulation, payroll tax rates now adjust in each year of the policy simulation to ensure that regional government borrowing requirements are the same as their base case values. With regional real consumer wages sticky, the reductions in payroll tax rates reduce the employer cost of labor, leading to short-run employment gains relative to the base case. This causes regional adjustment costs to fall below base case levels, leading to the short-run negative RELMAI values reported in Figure 11.2. The RELMAI deviations reach their lowest point in 1999/00, and thereafter return towards the base case as the regional real wage rises to return regional unemployment to its baseline path.

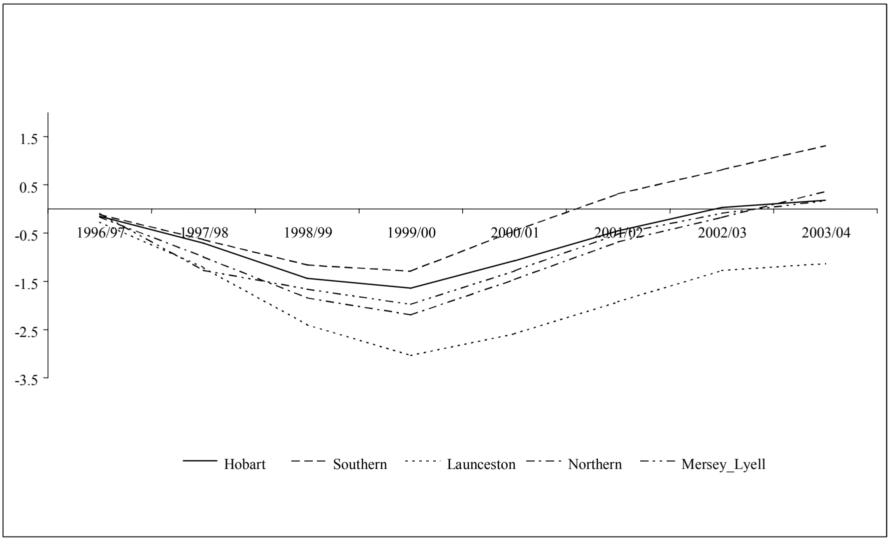


Fig. 11.2. Regional RELMAI indices (PSBRs exogenous): Percentage point deviation from base.

11.5 Concluding Comments

This chapter has examined some developments in CGE modeling that have enhanced regional economists' tool kit for analyzing globalization. These developments relate primarily to dynamic CGE models. We outlined how a change in model closure can allow historical analysis of globalization forces and presented a sub-model of a labor-input-loss index to analyze regional labor market adjustment costs.

The development of modern dynamic multiregional CGE models is the outcome of a continuous development process in this type of modeling over the past quarter of a century. The flexibility and detail of such models allow them to be applied to a large array of economic problems with little or no adjustment to their basic theoretical structure. While such CGE models have greatly increased in scale over the years, advancement in software for solving general equilibrium models has meant that ease of operating the models has, if anything, improved. The strong theoretical foundations of CGE models also mean that model results can be interpreted in terms of model mechanisms and data, at least by economists with the requisite skills. It is of course important that regional economic modelers carefully explain their results in order that policy lessons can be drawn, and as an important computational check.

While demonstrating the power of the present CGE tools, it is important to draw out the areas where limitations remain and the most promising areas for future model development. Globalization and other economic forces are continually reshaping our economy just as model builders seek to improve their treatment of the current behavior of economic agents. Most regional CGE models continue to employ assumptions of market-clearing (except for the labor market), perfect competition and constant returns to scale. However, many national CGE models now employ imperfect competition and scale economies, and as mentioned in section 11.1, some regional models are following suit. There are, however, a considerable number of alternatives for introducing these features, and it is important to ask how best to undertake this task, what approaches are best for particular economic sectors, and whether the introduction of these features in a particular way significantly affects results, is theoretically and empirically justified, and is not hindered in its implementation by lack of data.

A start on shedding light on these questions can be found in Abayasiri-Silva and Horridge (1998). They investigated the short-run and long-run effects of changes in domestic protection using a CGE model that allowed for the implementation of different combinations of three different technology possibilities (constant returns to scale, increasing returns to scale internal to the firm, or increasing returns to scale external to the firm), three different pricing rules (marginal cost pricing, Lerner pricing, or Harris' (1984) pricing rule) and two entry/exit possibilities (fixed number of firms, or free entry). Abayasiri-Silva and Horridge (1998) found that their results under Lerner pricing and internal economies of scale were very similar to those obtained under constant returns to scale. Only under the Harris pricing rule or external economies of scale were the results notably different from those obtained under constant returns to scale. This confirmed earlier work, such as that of Nguyen and Wigle (1992), suggesting that Harris' dramatic trade liberalization results were due in part to his mixed Lerner/Eastman-Stykolt (see Eastman and Stykolt, 1966) pricing rule and not simply the imperfect competition and economies of scale features of his model. Abayasiri-Silva and Horridge (1998) conjecture that neither Harris pricing nor external economies have compelling theoretical bases, leaving the striking conclusion that CGE models adopting the conventional assumptions of constant returns to scale and perfect competition might quite accurately reproduce the results of CGE models adopting plausible assumptions relating to increasing returns to scale and imperfect competition.

The unconvincing behavioral assumptions that underpin the Eastman-Stykolt proposition and a dearth of empirical support for pervasive significant economies of scale (Dixon and Rimmer, 2002) provides a solid justification for CGE modelers retaining assumptions of perfect competition and constant returns to scale for most applications. At present, the best approach may be to treat the phenomena of imperfect competition and economies of scale when there is evidence that they may be important to the particular application. Dixon and Rimmer (2002), for instance, report examples where the sensitivity to a lowering of protection on Australian-produced motor vehicles may induce a lowering of unit costs, or impact on Australia's trading partners policies on access to their markets.

Nevertheless, the matters of imperfect competition and scale economies need to be pursued by regional CGE researchers, as they may well prove important when modeling in a spatial context. Haddad and Hewings (2005) simulations with a multiregional model of Brazil suggest that scale economies, while not necessarily greatly effecting welfare results, may give rise to asymmetric effects from transport investment. The Dixit-Stiglitz specification, employed in Abayasiri-Silva and Horridge's (and others) modeling of external economies, has provided a fertile research path for regional economists investigating intra-industry trade and agglomeration. Dixon (1998) demonstrates under Lerner pricing that firm numbers in an industry move close to proportionally to output, and consequently so to do (fixed and variable) costs. This explains why CGE models can give similar results under constant-returns and internal-economies specifications. However, this need not be the case with other pricing rules (and indeed it is not the case with the Harris specification). We would agree with Dixon (1998) that more empirical work is necessary to determine actual pricing behavior. We would also urge that such work be at a reasonably disaggregated level as market conduct is likely to be influenced by the characteristics of a particular sector.

There is also a clear research agenda in the area of historical modeling. Results from modeling of this type have suggested the plausibility of the approach. Historical simulations invariably show trends in technical change and taste changes that are consistent with known technological and taste changes.[21] However, it should be borne in mind that the numerical results for these variables are a product of both theory and data and will thus contain some degree of error (similar to error terms in econometric modeling). One source of error may be in the model's parameter specification. An interesting avenue for future research might be to examine alternative values for certain parameters such as those governing inter-regional substitution possibilities across commodities. The aim would be to

---

[21] For instance, the historical simulation we report in section 11.3 shows increased intermediate usage of, for instance, finance and business services, and reduced wholesale margins – consistent with lower inventories, and a change in household tastes towards entertainment, pharmaceuticals and communication services.

adjust (within reasonable bounds) the values for these parameters such that changes in relative prices and activity levels explained a higher share of the observed regional output responses, thereby reducing the need to resort to endogenously calculated inter-regional sourcing twists and source-specific technical changes.

References

Abayasiri-Silva K, Horridge M (1998) Economies of scale and imperfect competition in an applied general equilibrium model of the Australian economy. In: Arrow KJ, Ng YK, Yang X (eds) Increasing returns and economic analysis. Macmillan, London, pp 307–334

Adams PD (1998a) Long-run effects of APEC trade liberalisation: An applied general equilibrium analysis. World Economy 21:931-952

Adams PD (1998b) Prospects for the Australian economy and the impact of the Asian crisis. Australian Bulletin of Labour 24:247-278

Adams PD (2004) Medium term prospects for the Australian economy and the impact of Kyoto-compliance. Australian Bulletin of Labour 30:1-15

Adams PD, Parmenter BR (1995) An applied general equilibrium analysis of the economic effects of tourism in a quite small open economy. Applied Economics 27:985-994

Adams PD, Horridge JM, Parmenter BR (2000) Forecasts for Australian regions using the MMRF-Green model. Australasian Journal of Regional Studies 6:293-323

Adams PD, Dixon PB, Rimmer MT (2001) The September 11 shock to tourism and the Australian economy from 2001-02 to 2003-04. Australian Bulletin of Labour 27:241-257

Adelman I, Yeldan E (2000) The minimal conditions for a financial crisis: a multiregional intertemporal CGE model of the Asian crisis. World Development 28:1087-1100

Armington PS (1969) The geographic pattern of trade and the effects of price changes. IMF Staff Papers 16:176-199

Conrad K, Schmidt TFN (1998) Economic effects of an uncoordinated carbon dioxide policy in the European Union: an applied general equilibrium analysis. Economic Systems Research 10:161-182

Borland J (1996) Labour market flows data for Australia. Australian Economic Review, 2nd Quarter: 225-235

Bröcker J (1998) How would an EU-membership of the Visegrád-countries affect Europe's economic geography? Annals of Regional Science 32:91-114

Burniaux JM, Waelbroeck J (1992) Preliminary results of two experimental models of general equilibrium and imperfect competition. Journal of Policy Modeling 14:65-92

Diao X, Somwaru A, Tuan F (2003) Regional and national perspectives of China's integration into the WTO: A CGE inquiry with emphasis on the agriculture sector. Review of Urban and Regional Development Studies 15:84-105

Dixit AK, Stiglitz JE (1977) Monopolistic competition and optimum product diversity. American Economic Review 67:297-308

Dixon PB (1998) Comment on "Economies of scale and imperfect competition in an applied general equilibrium model of the Australian economy." In: Arrow KJ, Ng YK, Yang X (eds) Increasing Returns and Economic Analysis. Macmillan, London, pp 335-337

Dixon PB, Parmenter BR, Sutton J, and Vincent DP (1982) ORANI: A multisectoral model of the Australian economy. North-Holland, Amsterdam

Dixon, PB, Parmenter BR, Rimmer MT (1998) The effects of foreign students on the Australian economy and its regions. Paper presented at the 27th Conference of Economists, University of Sydney, Sept 28 - Oct 1.

Dixon PB, Menon J, Rimmer MT (2000) Changes in technology and preferences: a general equilibrium explanation of rapid growth in trade. Australian Economic Papers 39:33-55

Dixon PB, Rimmer MT (2002) Dynamic, General Equilibrium Modelling for Forecasting and Policy: A Practical Guide and Documentation of MONASH. North Holland, Amsterdam

Dixon PB, Rimmer MT (2005) Explaining a dynamic CGE simulation with a trade-focused back-of-the-envelope analysis: the effects of eCommerce on Australia. In: S. Jayasuriya (ed) Trade Theory and Analytical Models: Papers in Honour of Professor Peter Lloyd, Vol. 1, Edward Elgar, Cheltenham, pp 223-256

Eastman HC, Stykolt S (1966) The tariff and competition in Canada. University of Toronto Press, Toronto

Ferreira Filho JBS, Horridge JM (2004) Economic integration, poverty and regional inequality in Brazil. Paper presented to 44th Congress of the European Science Association, August 25-29, Porto, Portugal

Fujita M, Krugman P, Venables AJ (1999) The spatial economy: Cities, regions and international trade. MIT Press, Cambridge MA

Francois JF, Shiells CR (1994) AGE models of North American free trade. In: Francois JF, Shiells CR (eds.) Modeling Trade Policy: Applied General Equilibrium Assessments of North American Free Trade. Cambridge University Press, Cambridge

Gasiorek M, Smith A, Venables AJ (2002) The accession of the UK to the EC: A welfare analysis. Journal of Common Market Studies 40:425-447

Gazel R, Hewings GJD, Sonis M (1995) Trade, sensitivity and feedbacks: Interregional impacts of the US-Canada free trade agreement. In: van den Bergh JCJM, Nijkamp P, and Rietveld P (eds) Recent advances in spatial equilibrium modeling: Methodology and applications. Springer-Verlag, Berlin

Giesecke JA (2002) Explaining regional economic performance: an historical application of a dynamic CGE model. Papers in Regional Science 81:247-278

Giesecke JA (2003) The regional impacts of the growth of the “new economy” industries – the case of Australian telecommunications. Paper presented to 43rd Congress of the European Regional Science Association, August 27-30, Jyvaskyla, Finland

Giesecke JA (2004a) Foreign students and regional economies: a multiregional general equilibrium analysis. Australasian Journal of Regional Studies 10(1), forthcoming

Giesecke JA (2004b) The extent and consequences of recent structural changes in the Australian economy, 1997 – 2002: Results from historical / decomposition with MONASH. General working paper No. 151, Centre of Policy Studies, Monash University, Melbourne, www.monash.edu.au/policy

Giesecke JA, Madden JR (2003a) A large-scale dynamic multi-regional CGE mModel with an illustrative application. Review of Urban and Regional Development Studies 15:2-25

Giesecke JA, Madden JR (2003b) Regional labour market adjustment to competition policy reforms: A dynamic CGE framework for assessment. Australian Journal of Labour Economics 6:409-433

Gosh M (2002) The revival of regional trade arrangements: a GE evaluation of the impact on small countries. Journal of Policy Modeling 24:83-101

Goto J (1996) The impact of migrant workers on the Japanese economy: trickle vs. flood. Yale Economic Growth Center discussion paper 749, Yale University, New Haven

Groenewold N, Hagger AJ, Madden JR (2003) The Effects of Interregional Transfers: A Political-Economy CGE Approach. Papers in Regional Science 82:535-554

Haddad EA, Domingues EP, Perobelli FS (2002) Regional effects of economic integration: the case of Brazil. Journal of Policy Modeling 24:453-482

Haddad EA, Hewings GJD (2005) Market imperfections in a spatial economy: some experimental results. Quarterly Review of Economics and Finance 45:476-496

Harris R (1984) Applied general equilibrium analysis of small open economies with scale economies and imperfect competition. American Economic Review 74:1016-1032

Harrison WJ, Horridge JM, Pearson KR (2000) Decomposing simulation results with respect to exogenous shocks. Computational Economics 15: 227-249

Harrison GW, Hougaard Jensen SE, Haagen Pedersen L, Rutherford T (eds) (2000) Using dynamic general equilibrium models for policy analysis. North Holland, Amsterdam

Hertel TW (ed) (1997) Global trade analysis: Modeling and applications. Cambridge, Cambridge University Press.

Hertel TW, Ianchovichina E, McDonald BJ (1997) Multi-region general equilibrium modeling. In: Francois JF, Reinert KA (eds) Applied methods for trade policy analysis. Cambridge University Press, Cambridge

Jean S, Laborde D (2004) The impact of multilateral liberalization on European Unions: A CGE assessment. CEPII working paper No. 2004-20, Paris www.cepii.fr

Jones R, Whalley J (1989) A Canadian general equilibrium model and some applications. Journal of Urban Economics 25:368-404

Kehoe TJ (2005) An evaluation of the performance of applied general equilibrium models of the impact of NAFTA. In: Kehoe TJ, Srinivasan TN, Whalley J (eds) Frontiers in Applied General Equilibrium Modeling: Essays in Honor of Herbert Scarf. Cambridge University Press, Cambridge

Kilkenny M (1993) Agricultural liberalization in segmented or integrated markets, with scale economies. Journal of Economic Integration 8:201-218

Kim E, Kim K (2002) Impacts of regional development strategies on growth and equity of Korea: a multiregional CGE model. Annals of Regional Science 36:165-180

Lee JE (2004) The sequential issue in free trade areas: policy implications for Korea. International Journal of Finance and Economics 9:165-174

Lejour A, Nahuis R (2000) Openess, growth and R&D spillovers: uncovering the missing link? Paper presented to the third annual conference on Global Economic Analysis, Melbourne, 27-30 June.

McGregor PG, Swales JK, Yin YP (1996) A long-run interpretation of regional-input output analysis. Journal of Regional Science 36: 479-501

McGregor PG, Swales JK, Yin YP (1999) Spillover and feedback effects in general equilibrium interregional models of the national economy: a requiem for interregional input-output. In: Hewings GJD, Sonis M, Madden M, Kimura Y (eds) Understanding and Interpreting Economic Structure. Springer, Berlin, pp 167-190

Madden JR (2006) Australia: Central fiscal dominance, collaborative federalism and economic reform. In Srinivasan TN, Wallack JS (eds) Federalism and economic reform: International perspectives. Cambridge University Press, New York

Nixon P (1997) Commonwealth-state inquiry into the Tasmanian economy. Report to the Prime Minister of Australia and the Premier of Tasmania, Hobart.

Nguyen T, Wigle R (1992) Trade liberalization with imperfect competition: the large and small of it. European Economic Review 36:17-35

Partridge MD, Rickman DS (1998) Regional computable general equilibrium modeling: A survey and critical appraisal. International Regional Science Review 21:205-248

Peter MW (1993) The use of the ORANI model in the immigration debate. General paper No. G-103, Centre of Policy Studies and the Impact Project, Monash University, Melbourne www.monash.edu.au/policy/

Salvatore D (2004) Globalization, growth and poverty: editor's introduction. Journal of Policy Modeling 26: 421-424

Tiwari P, Doi M, Itoh M, Itoh H (2003) A CGE analysis of the potential impact of information technology on the Japanese economy. Policy Reform 6:17-33

Tsigas ME, Frisvold GB, Kuhn B (1997) Global climate change and agriculture. In: Hertel TW (ed) Global trade analysis: Modeling and applications. Cambridge University Press, Cambridge

Van Meijl H, van Tongeren F (1999) Endogenous international technology spillovers and biased technical change in agriculture. Economic Systems Research 11:31-48

Verikos G, Zhang X (2000) Sectoral impacts of liberalizing trade in services. Paper presented to the Third Annual Conference on Global Economic Analysis, Melbourne, June 27-30

Walmsley T (2002) Incorporating international ownership of endowments into a global applied general equilibrium model. Economic Modelling 19:679-707

Willenbockel D (1994) Applied General Equilibrium Model: Imperfect Competition and European Integration. Wiley, Chichester

World Bank (2000) Assessing globalization. World Bank briefing papers. PREM Economic Policy Group and Development Economics Group. http://www1.worldbank.org/economicpolicy/globalization/

Appendix

Brief Note on Decomposition Method

For the case of one endogenous variable Z, Harrison *et al.* (1999: 4-5) summarize their decomposition as follows. Assume Z can be expressed as a function of n exogenous variables $X_1, X_2, ..., X_n$

$$Z = f\left(X_1, X_2, ..., X_n\right)$$

Next, assume that the vector of exogenous variables $X = \left(X_1, X_2, ..., X_n\right)$ moves along some path which begins at $X_{INITIAL}$ and ends at X_{FINAL}:

$$X_{INITIAL} = \left(X_{10}, X_{20}, ..., X_{n0}\right)$$

$$X_{FINAL} = \left(X_{11}, X_{21}, ..., X_{n1}\right) = \left(X_{10} + \Delta X_1, X_{20} + \Delta X_2 +, ..., + X_{n0} + \Delta X_n\right)$$

Assume that the shocks are divided in h equal installments. Provided h is sufficiently large, the effect of applying the first $(1/h)^{\text{th}}$ installment of the total shock can be accurately approximated by:

$$dZ = f_1 dX_1 + f_2 dX_2 + ... + f_n dX_n \quad \text{where } f_i = \partial f / \partial X_i \text{ and } dX_i = \Delta X_i / h$$

If h is sufficiently large (ie the dX_i are sufficiently small), then the approximation will be exact and the right hand side terms provide a decomposition of the total change dZ for this first installment of the total shock. We would then go on to apply the remaining h-1 installments of the shocks. The f_i will depend on the value of Z and X at each step, and so will change with each step. This provides no additional computational burden for GEMPACK, since values for the f_i are required anyway for the standard GEMPACK solution algorithm. Finally, we can calculate the contribution made by each shock ΔX_i to the total change in $Z\left(\Delta Z\right)$ by adding up the results for $f_i dX_i$ over each of the h steps. Note that some path for the exogenous variables must be chosen in order to implement such a decomposition. In the above example, and in the algorithm implemented in GEMPACK by Harrison *et al.*, (1999) a straight line path for the movements of the exogenous variables from their pre- to their post-shock values is chosen. Harrison *et al.* (1999) argue that, among the possible choices for the path of the exogenous variables, a straight-line path will typically be among the more natural of the possible paths.

12 Modeling Small Area Economic Change in Conjunction with a Multiregional CGE Model

Ian Sue Wing[1] and William P. Anderson[2]

[1] Department of Geography and Center for Energy and Environmental Studies, Boston University
[2] Department of Geography and Center for Transportation Studies, Boston University

12.1 Introduction

Regional economic projections are often made at the level of spatial aggregation that is most convenient from political, data and computational perspectives. For example, in the US projections are often made at the level of 50 states. From a policy perspective this makes some sense since disbursements of federal funds for infrastructure are generally made to state governments and sub-state governments have little political representation at the federal level. Economic data is much richer at the state level than at any sub-state level. Also, for many models a larger number of regions than the 50 implies excessive computational burden.

There are, however, a number of economic issues that cannot be adequately addressed at the state level. A few examples include:

Economically depressed regions which seldom coincide with state borders. Even a state as small as New Hampshire has significant differences in economic well being from north to south and large states like Texas may simultaneously contain some of the most affluent and some of the most depressed regions in the country.

Land use change such as the transformation of agricultural land and the general phenomenon of urban sprawl cannot be realistically addressed at the state level.

Environmental pollution studies benefit greatly from spatial disaggregation. It may be possible to estimate aggregate emissions at the state level but not to estimate exposed populations or to provide input to atmospheric models.

In *transportation studies*, origins and destinations of trips are points rather than areas, so greater spatial detail generates more useful results. Furthermore, the spatial impacts of transportation infrastructure do not necessarily respect state borders.

Our purpose in this paper is to specify a comprehensive and rigorous framework for sub-state economic analysis. In order to do this we must first define a set of sub-state spatial units. There have been attempts in the past to develop models from economically functional regions rather than political jurisdictions. These include the BEA regions and National Transportation Analysis Regions. From the perspective of data availability it is easier to work with regional definitions that have long histories and that are used consistently by different agencies that collect and publish data. In the US this means either counties or Census Metropolitan Areas (CMAs). CMAs are preferable for many types of analysis because they constitute functional regions with high levels of internal spatial interaction. They are not, however, exhaustive (they exclude all rural areas) and only useful for analyzing urban and interurban phenomena. For the model described in this paper we have chosen the 3000 plus US counties as our spatial units. This is an exhaustive regionalization that provides sufficient spatial detail to address some, if not all, of the issues mentioned above. Also, counties can be aggregated up to CMAs and BEA regions.

While a great deal of regional economic analysis has been based on Leontief type Input-Output (IO) models, we have chosen instead to develop a spatially disaggregate Computable General Equilibrium (CGE) model. IO models have the advantages of relatively modest data requirements, a high degree of industrial specificity and a good ability to capture multiplier effects. They are limited, however, to modeling exogenous shocks that can be represented as changes in final demand and they fail to capture supply side effects. As an example of the latter, in the presence of labor supply limitations, large infrastructure investments may result in a combination of increased employment and increased wages. IO models can only capture, and may exaggerate, the employment effect.

How does one develop a CGE model that can make projections at the county level? Is it just a question of applying a state level specification to county level data? The answer is no, for three reasons. First, many types of economic data are not available at the county level. Data for population and employment are available, but data for economic categories like value added and investment are not. Second, models that may be computationally manageable for 50 states may become infeasible for 3000 counties. Finally, there are number issues that become more important for smaller geographical units and must therefore be addressed in a county level model. For example, the issue of available space, which can be neglected at the national or state level, becomes an important constraint on the spatial evolution of activities. Population growth in a county depends not only on its ability to yield utility to new residents but also to the extent to which it is "built out."

The approach we have adopted is to develop a county level model in conjunction with a state level CGE model. Endogenous variables in the state model serve as boundary conditions for the county level. The county level model is not simply an allocation mechanism. It determines a spatial equilibrium distribution of economic activities and population at the county level within each state.

Our emphasis is to develop a model that has a rigorous specification of the supply side – an emphasis that is often lacking in regional models. At this point the demand side of the model is rather aggregate and simple. In later versions of the model, we plan to develop a more comprehensive and explicitly spatial demand side.

The remainder of the chapter is organized as follows. The next two sections describe the state-level and county-level models respectively. While the model is not yet fully operational, we have solved a proof-of-concept model at the state level and conducted empirical estimation of population distribution mechanisms at the state and county levels. These are described in a subsequent section. Finally in the conclusion we speculate about policy analyses that conducted with the model.

12.2 A State-Level Computable General Equilibrium Economic Model

Our first step is to construct a state-level economic simulation model to project the trajectories of output, employment, prices and wages by industry in each of the 50 states over the time horizon 2000-2050 on a series of five-year time-steps.[1] Our approach is deliberately simple, and, in keeping with the long-run nature of our projections, focuses on the supply side of the economy. We treat the individual industries within each state as representative firms and simulate the dynamics of each according to a Solow-Swan growth model that captures the decision to invest in accumulating a stock of capital in a recursive-dynamic fashion (i.e., based myopically on the values of exogenous and endogenous variables in the current period, rather than looking forward over the entire simulation horizon).[2] These decisions determine how industry capital stocks in each state evolve from one time-period to the next. Each industry produces output according to a production function that specifies how capital, intermediate goods, and labor—which is assumed to be mobile among states – are combined. In each period, the decisions of the population to migrate among and within states determines the intra-state supply of labor. Industries' competition for workers then simultaneously determines their

---

[1] The sectoral disaggregation in Table 1 captures the major sources of air pollutant emissions without being overly detailed. Although simulation modeling is a well established technique for projecting future emissions, there are comparatively few examples of its use to generate long-run forecasts for regions of the U.S. The best example is the REMI economic-demographic model and its antecedents (Stevens and Treyz 1986; Treyz et al 1992; Treyz 1993), which are demand-driven simulations of the spatial economy.

[2] This may be contrasted with an intertemporal optimization framework used in multi-regional climate policy simulations (e.g., Manne and Richels 1992; Nordhaus and Boyer 1999) and dynamic rural-urban computable general equilibrium (CGE) models (e.g., Kelly and Williamson 1984; Becker et al 1992), in which representative agents choose trajectories of investment to maximize their present value of discounted utility from consumption, looking forward over the entire simulation horizons. This approach, although theoretically appealing, is too expensive computationally to work here.

wages, the prices and quantities of their output, and also the resources available for capital investment in the next period.

12.2.1 The Within-Period Spatial General Equilibrium Model

Our first task is to build and calibrate the within-period component of the model, which is a static spatial price equilibrium simulation of the U.S. economy.[3] In each of the s states in the U.S., assume that there are j industries, each of which produces a single homogeneous commodity indexed by i. For each industry in a given state, gross output (y) is produced using capital (k), labor (l) and a composite of intermediate inputs (x), according to a constant-returns-to-scale Cobb-Douglas production function:[4]

$$y(j,s) = \xi(j,s) l(j,s)^{\alpha(j,s)} k(j,s)^{\beta(j,s)} x(j,s)^{\gamma(j,s)} . \qquad (12.1)$$

Here, ξ is a Hicks-neutral shift parameter, and the parameters α, β and γ are the value shares of l, k and m in output. α and β are derived from the components of value added by state recorded in BEA's Gross State Product (GSP) accounts. γ is derived using the coefficients on intermediate input from the latest U.S. social accounting matrix (SAM), constructed from the make and use tables in BEA's Input-Output (I-O) accounts.

We do not observe gross output at the state level—industries' gross state product is synonymous with their value added (v). However, y can be easily inferred from v if we assume that the economy has the same structure of inter-industry demand at both aggregate and the state levels. In the SAM, the coefficient on the use of intermediate input i by industry j ($\bar{X}$) is the average value share of i in j's gross output. It follows that over all i intermediate inputs:

$$p_X(j) x(j,s) = \sum_i p(i) \bar{X}(i,j) p(j) y(j,s) , \qquad (12.2)$$

where p and p_X are the price indices of output and the composite of intermediate materials. Constant returns to scale in production implies that:

---

[3] Throughout, we stick to the following notational convention: lower-case letters for industry-level variables (e.g., prices), upper-case letters for state- and aggregate-level variables (e.g., value added), and tildes (~) for county-level variables (e.g., housing units). We suppress time subscripts when describing the static equilibrium simulation but include them in our explication of the model's dynamics.

[4] Scale externalities in production are central to economic models of urbanization and agglomeration (Duranton and Puga 2004). However, including them directly in the production function would result in a non-convex optimization problem which is very difficult to solve (see, e.g, Fan et al 2000). We therefore opt to introduce the effect of external economies parametrically, imposing a neutral decline in the unit cost of production by scaling the value of ψ downward as a concave function of j's output in state s (i.e., learning-by-doing) or input-output based indices of within-state supply or demand linkages (e.g., Bartelsman et al 1994; Paul and Siegel 1999).

$$p(j)y(j,s) = v(j,s) + p_X(j)x(j,s)\,, \tag{12.3}$$

which enables the value of gross output to be easily imputed as:

$$p(j)y(j,s) = v(j,s) \Big/ \left(1 - \sum_i p(i)\bar{X}(i,j)\right). \tag{12.4}$$

To keep things simple, we propose to model the supply of capital to each industry as being fixed within a period. Competitive equilibrium in output markets results in a zero-profit condition that equates the price of output and the short-run unit cost function that is dual to (12.1):

$$p(j) = \left(1+\tau_Y(j,s)\right)\left(\psi(j,s)w(j)^{\frac{\alpha}{\alpha+\gamma}}\,p_X(j)^{\frac{\gamma}{\alpha+\gamma}}\,k(j,s)^{\frac{-\beta}{\alpha+\gamma}}\,y(j,s)^{\frac{1-\alpha-\gamma}{\alpha+\gamma}} + \bar{R}\frac{k(j,s)}{y(j,s)}\right) \tag{12.5}$$

Here, p is the aggregate price of output in industry j, $\bar{R}$ is the economy-wide average capital rental rate, which we treat as exogenous for simplicity, w is the wage in that industry, τ_Y is the state's tax rate on j's output in the base year,[5] and ψ is a Hicks-neutral shift parameter whose initial value is calibrated to set the prices of all commodities to unity in the base year.[6] Simplifying (12.2) to (12.4) enables output in the base year to be calibrated, and permits the quantity and price of intermediate input to be modeled according to a Leontief specification, as follows:

$$x(j,s) = \sum_i \bar{X}(i,j)y(j,s)\,, \tag{12.6}$$

and

$$p_X(j) = \sum_i p(i)\bar{X}(i,j)p(j)\,. \tag{12.7}$$

The initial condition for w is derived from statistics on employment and compensation by industry tabulated in BEA's Annual State Personal Income (ASPI) accounts. The capital rental rate is based on the average value of corporate bond yields. τ is easily derived from output and tax revenue tabulated in the GSP accounts. It is less straightforward to determine the initial stocks of capital. Following Garofalo and Yamarik (2002), we plan to use the historical data series on aggregate-level investment and depreciation by industry from the BEA's Fixed Assets (FA) accounts, allocate these values among the states in proportion to their shares in value added, and then cumulate the resulting state-level investment and depreciation schedules into stocks of capital using the perpetual inventory method.

We assume that industries face an aggregate demand for their output (YD) that is the sum of own- and other-industry intermediate uses, and final uses by con-

---

[5] For simplicity, we simplify the actual structure of industry taxation, treating taxes as if they were levied on output.

[6] This calibration trick is widely employed in CGE models (Ginsburgh and Keyzer 2000); Dawkins et al 2001).

sumers. Intermediate use is determined by the economy's input-output structure and industrial composition. Final use is modeled according to a downward-sloping demand curve that shifts outward with rising aggregate income (GDP):[7]

$$Y_D(i) = \sum_j \bar{X}(i,j) \sum_s y(j,s) + \varpi(i) GDP / p(i) \, , \tag{12.8}$$

where the parameter ϖ denotes good i's share of total expenditure on final uses, derived from the SAM. Income balance is achieved by specifying two components to GDP, the sum across states of state-level value added (V) and aggregate tax revenue ($TAXREV$):

$$GDP = \sum_s V(s) + TAXREV \tag{12.9}$$

where

$$V(s) = \sum_j v(j,s) = \sum_j \left(w(j,s) l(j,s) + \bar{R} k(j,s) \right) , \tag{12.10}$$

and[8]

$$TAXREV = \sum_j \sum_s \tau_Y(j,s) P(j) y(j,s) \, . \tag{12.11}$$

Finally, market clearance for the i^{th} good implies that aggregate demand equal aggregate supply:

$$Y_D(j) = \sum_s y(j,s) \, . \tag{12.12}$$

The within-period equilibrium is closed by defining the labor market at the state level. The demand for labor by industry and state is:

$$l(j,s) = \tfrac{\alpha}{\alpha+\gamma} p(j) y(j,s) / w(j) \, . \tag{12.13}$$

Employment in each state is the sum of the labor demands by that state' industries:

$$L(s) = \sum_j l(j,s) \, . \tag{12.14}$$

This expression allows us to specify the average wage at the state level:

$$W(s) = \sum_j w(j) l(j,s) / L(s) \, , \tag{12.15}$$

---

[7] This formulation models final consumers as an aggregate representative agent with Cobb-Douglas preferences.

[8] This assumes that tax revenue is recycled as income in a lump sum, a standard practice in CGE modeling.

whose value is determined by the distribution of activity levels among industries therein. We propose to employ the simple model of state labor supply developed by Gallin (in press), which models the employment rate (H) in each state as a function of the average wage using a simple labor supply curve:

$$H(s) = \chi(s)W(s)^{\eta}, \qquad (12.16)$$

where the parameters η and χ denote the average elasticity of labor supply and the base-year employment rate, respectively. The former is usually taken to be around 0.3 in macroeconomic studies, while latter is tabulated in the BLS Local Area Unemployment Statistics. All that is left is to pin down state-level labor supplies, which we model as the product of the employment rate and the population (N), which we assume to be fixed in each period:

$$L(s) = H(s)N(s). \qquad (12.17)$$

Equations (12.5) to (12.17) specify the core within-period sub-model. They may be collapsed into a square system of nonlinear simultaneous equations in six unknowns (primal variables l, x and y, and corresponding dual variables w, PX and P) which can be solved for the allocations of labor, intermediate input and output, and the supporting vectors of wages, composite intermediate input prices, and commodity prices that constitute a spatial price equilibrium.[9]

12.2.2 The Dynamic Process of the Economy

We now describe the model's dynamic process, which consists of the equations of motion of state-level industry capital stocks and population from one time-step to the next. By specifying the boundary conditions that close the within-period equilibrium problem outlined above, the dynamic problem determines the temporal evolution of the spatial pattern of production.

Our first challenge is to determine the geographic distributions of growth and decline in industries' capital stocks.[10] We employ a recursive-dynamic approach, which is simple, easily implemented, and allows us to bring to bear on the model's regional forecasts the vast empirical literature on industrial location (e.g., Coughlin et al. 1991; Friedman et al. 1992; Woodward 1992).

---

[9] It is a straightforward task to specify and solve the model as a mixed complementarity problem (MCP—see Ferris and Pang 1997) in GAMS (Brooke et al 1998) with the MPSGE sub-system (Rutherford 1999). The only potential difficulty is the computational cost of finding a solution to a problem of this size (50 states × 20 industries × 6 equations). The authors will provide an operational proof-of-concept model upon request.

[10] These are driven by the industry-state pattern of investment in each period, which is the equilibrium outcome of industries that seek to maximize their returns to capital. Due to the difficulty of computing a solution to the true inter-temporal spatial investment allocation problem, which is highly dimensional, we adopt a simpler approach.

Letting the index t denote time periods, we propose to model capital accumulation in each industry within a state using the standard perpetual inventory equation:

$$k(j,s,t+1) = inv(j,s,t) + \kappa(j)k(j,s,t)\,, \tag{12.18}$$

where inv is the quantity of investment and the parameter κ is the average per-period capital survival share in each industry, derived from BEA's Fixed Assets (FA) accounts. As a first cut, we plan to model investment simply as a fixed share (ι) of output of the corresponding industry in each state:

$$inv(j,s,t) = \iota(j)y(j,s,t)\,. \tag{12.19}$$

With (12.19) as a start, our second major task is to develop, numerically calibrate, and test the performance of a more realistic, empirically-based specification for the investment equation. Our preferred approach is the econometrically-calibrated investment accelerator model of Treyz et al. (1992) and Rickman et al. (1993). We propose to adapt this specification to work at the level of individual industries, focusing on two sets of influences on investment:

> Those affecting its *level*—from the macroeconomic literature on capital age structure and turnover (e.g., Caballero et al. 1995; Doms and Dunne 1998; Cooper et al. 1999), and
>
> Those affecting its *spatial distribution*—from the regional science literature on the effects of state taxes (e.g., Bartik 1985; Holmes 1998), location externalities (e.g., Head et al. 1995).[11]

The second challenge in characterizing the spatio-temporal evolution of economic activity is to specify the determinants of state-level population and labor supply, which depends upon both the growth and migration behavior of the population. Accordingly, we propose to model the evolution of the population in each state as a function of the net growth rate of state populations (G), and economic immigration (MI) and emigration (ME):

$$N(s,t+1) = N(s,t)\big(1+G(s,t)\big) + M^I(s,t) + M^E(s,t)\,. \tag{12.20}$$

Our third major research task is to estimate the terms on the right-hand side of (12.20). To keep things simple, we avoid explicit characterization of the demographic structure of either state populations or migrants. Instead, we propose to model state population growth using crude birth and death rates by state (b and d, respectively), which are assumed to follow exogenous trends:

$$G(s,t) = B(s,t) + D(s,t) = b_0(s)e^{b_1(s)t} - d_0(s)e^{d_1(s)t}\,. \tag{12.21}$$

---

[11] Note that if the technology parameter ψ in the unit cost function (12.5) declines with increasing output, the simple formulation of investment demand in eq. (12.19) will lead to progressive spatial agglomeration.

Here, the parameters b_0, b_1, d_0 and d_1 are estimated on time series data from the National Center for Health Statistics. We also plan to estimate migration flows empirically, in the spirit of Greenwood *et al.* (1991) and Treyz *et al.* (1992). We specify gross immigration and emigration at the state level, as well as population redistribution within-state migration (M^R), as functions of states' populations, average population densities (ρ_N), and their wage and employment rates:

$$M^{\omega}(s,t) = \mu_0^{\omega}(s)\rho_N(s,t)^{\mu_1^{\omega}} N(s,t)^{\mu_2^{\omega}} H(s,t)^{\mu_3^{\omega}} W(s,t)^{\mu_4^{\omega}}, \quad \omega = \{I, E, R\}, \qquad (12.22)$$

in which $\rho_N(s,t) = N(s,t)/A(s)$, where A is the land area of state s, and μ_0-μ_4 are parameters estimated on data from the IRS migration database and BEA's Regional Economic Information System (REIS).

In developing equations for both investment and migration it is natural for tensions to arise between the theoretical correctness of candidate specifications and the consistency of the simulated spatial economy's behavior with our priors when these specifications are included in the model. A key test of the workability of specifications as we implement them is therefore whether the simulated evolution of the geography of production and population is consistent with historical trends (e.g., Kim 1995; Black and Henderson 1999; Holmes and Stevens 2004) or official forecasts (e.g., the Census Bureau's state population projections—Campbell 1996).

12.2.3 Disaggregating Economic Activity and Population to the County Level

Our next step is to project the spatial pattern of industry output, population and the demand for transportation at this county scale. This is a process of down-scaling the spatio-temporal evolution of the economic and demographic variables solved for above from the state level to individual counties. We propose to do this via a simple two-step procedure. First, estimating empirical relationships that govern the geographic distribution of output, growth of population and housing stocks, and land use change, and then simulating the resulting set of behavioral equations as a system to solve for the equilibrium spatial distribution of economic activity, population and land use in future periods.

This procedure is the lynchpin of our framework, as it nests fifty county-level spatial allocation problems within the state-level spatial allocation problem outlined above using the solution for the latter problem as the boundary condition for each of the former ones.

12.2.4 Spatially Disaggregating Output

Our first challenge is to distribute output by industry down to the level of individual counties. Employment and compensation are the only variables for which there is comprehensive industry-by-industry coverage at this scale. Accordingly, we begin with labor input. For industry j in state s, labor demand ($\tilde{l}$) and total employment ($\tilde{L}$) in each of that state's constituent counties $c\{s\}$ are given by the analogues of equations (12.13) and (12.14):

$$\tilde{l}(j,c\{s\}) = \tfrac{\alpha}{\alpha+\gamma}\, p(j)\tilde{y}(j,c\{s\})/w(j) \tag{12.23}$$

and

$$\tilde{L}(c\{s\}) = \sum_j \tilde{l}(j,c\{s\}), \tag{12.24}$$

where $\tilde{y}$ denotes j's level of activity in county c. Similarly, the average wage at the county level ($\tilde{W}$) is given by the analogue of equation (12.15):

$$\tilde{W}(c\{s\}) = \sum_j w(j)\tilde{l}(j,c\{s\})/\tilde{L}(c\{s\}). \tag{12.25}$$

In the base year, $\tilde{l}$ is easily calculated from the Census Bureau's County Business Patterns (CBP) and the ASPI accounts, enabling the initial conditions for $\tilde{y}$ and $\tilde{W}$ to be computed directly from equations (12.23) to (12.25). At subsequent time-steps, $\tilde{y}$ is determined by distributing the equilibrium level of production of industry j in state s found by the state model among counties, enabling (12.23) to (12.25) to be used to calculate $\tilde{l}$ and $\tilde{W}$. Following the Figueiredo *et al.* (2002) empirical model of manufacturing plant births at the county level, we propose an apportionment procedure that utilizes a logistic sharing rule:

$$\tilde{y}(j,c\{s\},t) = y(j,s,t)e^{\tilde{\sigma}(j,c\{s\},t)} \Big/ \sum_{c\{s\}} e^{\tilde{\sigma}(j,c\{s\},t)}. \tag{12.26}$$

which represents industry j's propensity to locate its production in a given county, depends on several county-level variables: average population density ($\tilde{\rho}_N$), total population ($\tilde{N}$), the average wage, and lagged output (a proxy for local agglomeration externalities):

$$\tilde{\sigma}(j,c\{s\},t) = \sigma\left\langle \tilde{\rho}_N(c\{s\},t), \tilde{N}(c\{s\},t), \tilde{W}(c\{s\},t), \tilde{y}(j,c\{s\},t-1)\right\rangle. \tag{12.27}$$

Our fourth major research task is to specify a functional form for σ and econometrically estimate (12.26) and (12.27) using data from the REIS.[12]

---

[12] Recent studies (Henderson 1996; Becker and Henderson 2000; Greenstone 2002) have identified non-attainment of the NAAQS as a large significant influence on industrial location at the county level. While we do not propose to include this feedback in our simulation, we note that it can easily be introduced once a baseline projection of the spatial pattern of emissions is run though an air quality model to give county-level air pollutant concentrations.

12.2.5 Spatially Distributing Population

The second challenge is to distribute migration flows calculated in the state model. This enables industrial location in future periods to be simulated by pinning down the level and density of population that affect the right-hand side of (12.27). Census Bureau data give an initial condition for population at the county level ($\tilde{N}$) in the base year. We propose to model the evolution of $\tilde{N}$ in a manner analogous to equation (12.21), where at subsequent time-steps it depends on both population growth (determined by the parent state's birth and death rates in (12.21), and gross immigration ($\tilde{M}^I$) and emigration ($\tilde{M}^E$) for that county in the preceding period:

$$\tilde{N}(c\{s\},t) = \tilde{N}(c\{s\},t-1)\big(1+G(s,t-1)\big) + \tilde{M}^I(c\{s\},t-1) - \tilde{M}^E(c\{s\},t-1)\,. \tag{12.28}$$

County-level migration must sum to the state-level total calculated in (12.22), which is ensured by modeling migration flows according to the logistic sharing rule:

$$\tilde{M}^\omega(c\{s\},t) = \big(M^\omega(s,t) + M^R(s,t)\big)e^{\tilde{\mu}^\omega(c\{s\},t)} \Big/ \sum_{c\{s\}} e^{\tilde{\mu}^\omega(c\{s\},t)}\,, \quad \omega = \{I, E\}, \tag{12.29}$$

where the propensity to immigrate or emigrate ($\tilde{\mu}$) at the county level mimics equation (12.22):

$$\tilde{\mu}^\omega(c\{s\},t) = \tilde{\mu}_0^\omega(c\{s\})\tilde{\rho}_N(c\{s\},t)^{\tilde{\mu}_1^\omega}\, \tilde{N}(c\{s\},t)^{\tilde{\mu}_2^\omega}\, \tilde{H}(c\{s\},t)^{\tilde{\mu}_3^\omega}\, \tilde{W}(c\{s\},t)^{\tilde{\mu}_4^\omega}\, \tilde{O}(c\{s\},t)^{\tilde{\mu}_5^\omega}\,. \tag{12.30}$$

In this expression, the county employment rate is given by the analogue of (12.17) :

$$\tilde{L}(c\{s\}) = \tilde{H}(c\{s\})\tilde{N}(c\{s\}), \tag{12.31}$$

$\tilde{\rho}_N$ is the population density for a county with land area $\tilde{A} : \tilde{\rho}_N(c\{s\},t) = \tilde{N}(c\{s\},t) / \tilde{A}(c\{s\})$,

and $\tilde{O}$ is the occupancy rate of housing units ($\tilde{U}$) in c: $\tilde{O}(c\{s\},t) = \tilde{N}(c\{s\},t) / \tilde{U}(c\{s\})$. Our fifth major research task is to estimate (12.29) and (12.30) using data from the IRS migration database and the REIS to recover values for the parameters $\tilde{\mu}_0$ - $\tilde{\mu}_5$.

12.2.6 The Spatial Pattern of Housing and Land Use

Our third challenge is to model the process by which growth of population and income generates demand for both new housing and the conversion of land from agricultural to residential, commercial and industrial uses. Specifying this process enables us to simulate future population growth by determining the effect of $\tilde{U}$ on occupancy in (12.30). $\tilde{U}$ is given initially by CBP data for the base year. Over time it responds to the demand-side forces of population, occupancy and income (proxied for by the average wage), and to the supply-side forces of spatial constraints on new builds (proxied for by the average unit density, $\tilde{\rho}_U$: $\tilde{\rho}_U(c\{s\},t) = \tilde{U}(c\{s\},t) / \tilde{A}(c\{s\})$), and the availability of land that is "potentially convertible" to residential use ($\tilde{A}_{PC}$, proxied for by the acreage under agriculture), according to a function υ:

$$\begin{aligned}\tilde{U}(c\{s\},t) &= \tilde{U}(c\{s\},t-1)\\ &+ \upsilon\left\langle \tilde{\rho}_U(c\{s\},t-1), \tilde{A}_{PC}(c\{s\},t-1), \tilde{N}(c\{s\},t-1), \tilde{W}(c\{s\},t-1), \tilde{O}(c\{s\},t-1)\right\rangle.\end{aligned} \tag{12.32}$$

Our sixth major research task is to specify and estimate a reduced-form empirical model for υ.

To simulate future values of $\tilde{U}$ we need to determine $\tilde{A}_{PC}$ in (12.32), which requires us to model land use change. We do this simply by treating each county's total land area as comprising areas that are under residential, industrial and commercial uses ($\tilde{A}_{ICR}$), areas that are potentially convertible ($\tilde{A}_{PC}$) and areas that are "non-convertible" ($\tilde{A}_{NC}$, e.g. unusable, wilderness or otherwise protected areas, which we assume to be constant):

$$\tilde{A}(c\{s\}) = \tilde{A}_{ICR}(c\{s\},t) + \tilde{A}_{PC}(c\{s\},t) + \tilde{A}_{NC}(c\{s\}). \tag{12.33}$$

Urban sprawl is the progressive conversion of agricultural land to industrial, commercial and residential land at the county level (i.e., growth of $\tilde{A}_{ICR}$ at the expense of $\tilde{A}_{PC}$). The initial conditions for $\tilde{A}_{ICR}$ and $\tilde{A}_{PC}$ are calculated from CBP and Census of Agriculture data.[13] As in (12.32), $\tilde{A}_{ICR}$ responds to the demand-side forces of population, income and housing unit density, and to availability of potentially-convertible land on the supply side, according to a function ζ:

$$\begin{aligned}\tilde{A}_{ICR}(c\{s\},t) &= \tilde{A}_{ICR}(c\{s\},t-1)\\ &+\zeta\left\langle \tilde{\rho}_U(c\{s\},t-1), \tilde{N}(c\{s\},t-1), \tilde{W}(c\{s\},t-1), \tilde{O}(c\{s\},t-1), \tilde{A}_{PC}(c\{s\},t-1)\right\rangle.\end{aligned} \tag{12.34}$$

Our seventh major research task is to specify and estimate a reduced-form empirical model for ζ. This will enable urban growth management policies to be

---

[13] The Census of Agriculture is the only data source we could find on land use by area at the county level.

simulated via mandated reductions in $\tilde{A}_{PC}$ that attenuate the growth of $\tilde{A}_{ICR}$. The variables $\tilde{U}$ and $\tilde{O}$ then act as channels through which such policies exert feedback effects on migration and, ultimately, industrial location.

12.2.7 Spatial Equilibrium at the County Level

Estimation of equations (12.27), (12.30), (12.32) and (12.34) yields numerical expressions which can be used to simulate the county-level distribution of economic activity, population and land use. The variables that determine the levels of these indicators within each county are for the most part endogenous to the sub-state spatial pattern of growth.[14] We must therefore solve the system of equations (12.23) through (12.34) for the equilibrium allocation of industries and population.[15]

---

[14] In particular, at each future time-step $\tilde{y}$ depends on $\tilde{W}$ by (12.26), but $\tilde{W}$ depends on $\tilde{l}$, and, ultimately, $\tilde{y}$ by (12.23)-(12.25), implying that $\tilde{y}$ is a function of its own level. Furthermore, $\tilde{y}$ also depends on $\tilde{N}$, which by (12.28)-(12.30) is in turn a function of both its own level in addition to $\tilde{W}$ —and therefore $\tilde{y}$, as well.

[15] It is a straightforward task to specify and solve such an equilibrium problem in GAMS as a nonlinear program (NLP) with a dummy objective. The only potential difficulty may be solution problems caused by the global properties of the logistic functions in eqs. (12.26) and (12.29) (Perroni and Rutherford 1998), but this seems unlikely.

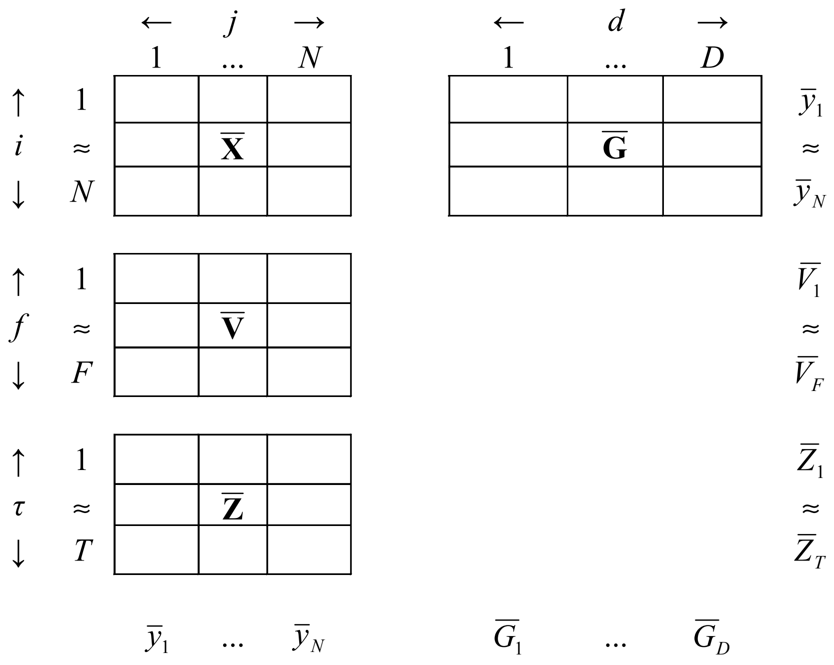


Fig. 12.1. Schematic of the Social Accounting Matrix

12.3 Preliminary Calibration Efforts

We calibrate the model on a set of social accounting matrices (SAM) for the U.S. states. The SAM that we use, shown schematically in figure 12.1, comprises an $N \times N$ matrix $\bar{\mathbf{X}}$ of inter-industry transactions, an $F \times N$ matrix $\bar{\mathbf{V}}$ of value-added activities, an $N \times D$ matrix $\bar{\mathbf{G}}$ of final demands, and a $T \times N$ matrix $\bar{\mathbf{Z}}$ of revenue flows due to tax and subsidy distortions. The indices $i, j = \{1, ..., N\}$ denote the set of industry sectors, $f = \{1, ..., F\}$ denotes the set of primary factors (labor and capital), $d = \{1, ..., D\}$ denotes the set of final demands (e.g. consumption, investment, government, and net exports), and $\tau = \{1, ..., T\}$ denotes the set of distortions.

Official data on state-level SAMs are not published. Traditionally, a SAM must be created for each individual state or region using regional multiplier techniques. However, since our goal is to develop a regional economic model that both explicitly represents state-level detail on the supply side and is consistent with macroeconomic linkages at the aggregate level, we employ a different approach to data development.

The procedure that we use starts by creating a national SAM for the U.S. in the format of figure 12.1, and then disaggregating it into s state-level SAMs. Each of these replicates the structure of the national table (the sub-matrices $\bar{\mathbf{X}}$, $\bar{\mathbf{V}}$, $\bar{\mathbf{G}}$ and $\bar{\mathbf{Z}}$) through corresponding state-level sub-matrices $\bar{\mathbf{X}}^S$, $\bar{\mathbf{V}}^S$, $\bar{\mathbf{G}}^S$ and $\bar{\mathbf{Z}}^S$. We use two datasets to perform the disaggregation. The first is 1999 input-output data published by the Bureau of Economic Analysis (BEA), which is used to develop a year 2000 national social accounting matrix (SAM) for the U.S.[16] The second is year 2000 data on gross state product by industry (GSP) and its constituent components, and annual state personal income (SPI), also from BEA. These data are used to derive each state's share of national value-added and final use according to its fractions of the total across all states of GSP and SPI, respectively.

The components of GSP that are tabulated in the data are labor, property-type income (i.e., a proxy for capital input) and indirect business taxes. Thus, letting the index comp denote these components and GSPC denote their individual contributions to GSP, we have for state s:

$$GSP(s) = \sum_{comp} GSPC(comp, s)\,, \tag{12.35}$$

where $comp = \{f, \tau\}$. This notation proves useful in formally describing our disaggregation procedure, to which we now turn.

In keeping with our assumption of a Leontief structure of inter-industry demands, we assume that the relationship between the intermediate inputs to a given industry and its value-added is the one given in the national SAM, and does not depend on state location. Therefore, the values of the column elements of the input-output matrix $\bar{\mathbf{X}}^S$ at the state level are determined by the GSP of the corresponding industries, which implies that the national input-output table can be disaggregated according the shares of each state in each industry column:

$$\bar{x}^S(i, j, s) = \bar{x}(i, j)\frac{GSP(j, s)}{\sum_s GSP(j, s)}\,. \tag{12.36}$$

Although the structure of intermediate demand is fixed, the substitutability of labor for capital implies that the relative intensities of use of these inputs in a given industry may differ across states, a fact which is borne out by actual data on the components of GSP. The values of elements of $\bar{\mathbf{V}}^S$ are thus imputed by separately apportioning among states each individual component of value added in

---

[16] The SAM is constructed from the BEA's 92-sector "Make of Commodities by Industries" and "Use of Commodities by Industries" tables for 1999 using the industry technology assumption. Its components of value added are disaggregated using data on industries' shares of labor, capital, taxes and subsidies in GDP published by BEA. The resulting benchmark flow table is aggregated up to three sectoral groupings (primary industries, manufacturing and services), scaled to approximate the U.S. economy in the year 2000 using the growth rate of real GDP from 1999-2000 (3.75 percent), and deflated to year 2000 dollars using the GDP deflator from the NIPAs.

each industry, according to the fraction of each state's contribution to that component of GSP in that industry:

$$\overline{v}^S(f,j,s) = \overline{v}(f,j)\frac{GSPC(f,j,s)}{\sum_s GSPC(f,j,s)}. \tag{12.37}$$

Since taxes and subsidies also differ by state, we use the same procedure to impute values for the elements of $\overline{\mathbf{Z}}^S$:

$$\overline{z}^S(\tau,j,s) = \overline{z}(\tau,j)\frac{GSPC(\tau,j,s)}{\sum_s GSPC(\tau,j,s)}. \tag{12.38}$$

Finally, we make the simplifying assumption that the structure of state-level final demands reflects the pattern in the national SAM, and that the elements of $\overline{\mathbf{G}}^S$ depend not on location but on states' incomes. We therefore used the simple procedure of disaggregating the aggregate final use matrix based on states' shares of total income:

$$\overline{g}^S(i,d,s) = \overline{g}(i,d)\frac{SPI(s)}{\sum_s SPI(s)}. \tag{12.39}$$

The results of this procedure are shown in figure 12.2, which illustrates the disaggregation of a three-sector U.S. national SAM into four SAMs that correspond to the Census regions. It is interesting to note that using the assumptions of (12.39), the row and column totals for each industry do not balance at the state level, but do at the national level. Given that the key assumption of our spatial equilibrium framework is that the law of one price holds for each commodity across all states, the difference between the row and column totals in a state's SAM indicates the magnitude of its net commodity trade flows. This information, along with interstate distances and data from the Department of Transportation's commodity flow survey, may be used to develop detailed state-to-state trade matrices, which can be used to elaborate the simple final demand system in (12.13). (See Appendix, equations A3 and A4.)

With these disaggregate data in hand, it is a simple matter to calibrate the spatial equilibrium model. We use the standard CGE calibration technique of setting all prices to unity and solving for the values of the technical coefficients that replicate the benchmark dataset.[17] The computational model is formulated and solved using the MPSGE subsystem (Rutherford, 1995, 1999) for GAMS numerical

---

[17] Sue Wing (in preparation) provides the details. Also see the illustrations of CGE calibration in Mansur and Whalley (1983) and Kehoe (1988). Dawkins et al (2001) provides an comprehensive comparison of different approaches.

simulation language (Brooke et al., 1999), reproducing the benchmark equilibrium with a residual of 10^{-6} (i.e., 10,000 dollars).[18]

Northeast

	P	M	S	U	
P	1.09	7.46	4.61	19.14	*32.30*
M	6.67	29.00	15.17	43.00	*93.83*
S	5.28	15.79	64.65	154.80	*240.53*
L	6.52	21.22	103.19		*130.93*
K	4.01	11.36	58.95		*74.32*
T	0.09	1.90	6.79		*8.79*
	23.66	*86.73*	*253.37*	*216.93*	*580.69*

South

	P	M	S	U	
P	5.69	14.70	6.56	28.24	*55.19*
M	14.95	47.35	21.92	63.44	*147.67*
S	14.59	26.00	93.03	228.42	*362.05*
L	13.59	30.42	148.15		*192.15*
K	10.79	22.26	80.16		*113.21*
T	0.40	4.35	10.48		*15.22*
	60.02	*145.07*	*360.29*	*320.10*	*885.49*

Midwest

	P	M	S	U	
P	2.09	8.86	4.11	19.39	*34.44*
M	8.98	47.00	13.65	43.57	*113.20*
S	7.28	24.04	58.21	156.85	*246.39*
L	8.87	33.10	93.65		*135.62*
K	5.88	16.90	49.38		*72.16*
T	-0.20	2.12	6.56		*8.48*
	32.90	*132.03*	*225.56*	*219.81*	*610.30*

West

	P	M	S	U	
P	3.67	7.93	4.82	19.90	*36.33*
M	10.73	34.40	16.04	44.71	*105.88*
S	9.48	17.31	68.20	160.96	*255.94*
L	10.41	22.11	104.98		*137.50*
K	7.07	13.47	64.61		*85.16*
T	0.31	1.52	7.04		*8.87*
	41.67	*96.74*	*265.69*	*225.56*	*629.67*

U.S. National SAM

	P	M	S	U	
P	12.53	38.96	20.09	86.68	*158.26*
M	41.33	157.74	66.78	194.71	*460.57*
S	36.63	83.15	284.10	701.03	*1104.90*
L	39.39	106.85	449.97		*596.21*
K	27.77	63.99	253.10		*344.85*
T	0.60	9.89	30.87		*41.36*
	158.26	*460.57*	*1104.90*	*982.42*	*2706.15*

Key:

P	Primary industries	L	Labor input
M	Manufacturing industries	K	Capital input
S	Service industries	T	Net tax revenue
U	Final uses		

Row and column totals in italics

All table entries in $10 billion 2000 U.S.

Fig. 12.2. Year 2000 aggregate social accounting matrices for U.S. census regions.

---

[18] The software automatically calibrates the technical coefficients of the excess demand functions, formulates the general equilibrium problem as square system of nonlinear equations, and solves this system as a mixed complementarity problem using the PATH solver (Dirkse and Ferris, 1995).

12.4 Population Dynamics

Equations of population movement (12.30), which are used to update the state population from one time step to the next, must be estimated econometrically using Census data on migration and data on socioeconomic characteristics that are endogenous to or can be calculated based on variables that are endogenous to, the within period equilibrium model. Equations (40a-c) were estimated for population movements over the 1995-2000 interval. (All variables as defined in table 12.1). Since decisions to move into, out of, or from one location to another within a particular state are all governed by the same set of forces, we employ seemingly unrelated regression specification for the state-level migration equations, which are estimated using three-stage least squares:

$$\begin{aligned} inm &= \alpha_1 + \beta_1 pop + \gamma_1 unemp + \delta_1 wage + \phi_1 popden + \theta_1 occu + \lambda_1 unitden && \text{(a)} \\ outm &= \alpha_2 + \beta_2 pop + \gamma_2 unemp + \delta_2 wage + \phi_2 popden + \theta_2 occu + \lambda_2 unitden && \text{(b)} \\ intm &= \alpha_3 + \beta_3 pop + \gamma_3 unemp + \delta_3 wage + \phi_3 popden + \theta_3 occu + \lambda_3 unitden && \text{(c)} \end{aligned} \qquad (12.40)$$

Table 12.1 List of Variables and Sources

Variable	Description	UNITs	Source
inm	State in-migration, 1995-2000	1000 persons	Census Bureau data files
outm	State out-migration, 1995-2000	1000 persons	Census Bureau data files
intm	State internal migration, 1995-2000	1000 persons	Census Bureau data files
pop	State population, 1995	1000 persons	BEA State Economic Profiles
unemp	State average unemployment rate, 1995	Percent	BLS Local Area Unemployment Statistics
wage	Av. wage per job, 1995	current dollars	BEA State Economic Profiles
popden	State population density, 1995	1000 persons per sq. mi.	BEA REIS and Census Gazetteer data files
unitden	State av. density of housing units, 1995	Units per sq. mi.	Census Gazetteer data files
occu	State av. occupancy rates of housing units, 1995	Persons per unit	BEA REIS and Census Gazetteer data files

The results are shown in table 12.2. Since *popden* = *occu* × *unitden*, two variants of equation (12.28) are estimated, one with *popden* (specification I) and the other with *occu* and *unitden* (specification II). Specifications (III) and (IV) in the table control for the influence of population size in the dependent variables, which express the numbers of in- out- and internal migration for each state as fractions of the respective state populations. We attempt to capture the effects of spatial autocorrelation by including spatial lags of the covariates. For each state, the spatial lag of a variable is computed as the average of the values of that variable over all contiguous states. In this calculation AK and WA, and CA and HI are treated as contiguous. These results are shown in table 12.3.

The fit of the regressions is generally good, and is improved by the addition of the spatial lags of the explanatory variables. The size of a state's population is the strongest predictor of all three types of migration, with positive effects on absolute levels of migration and the rate of internal migration, and negative effects on the rates of in- and out-migration.

The effect of average population density—or, equivalently, the combination of unit density and occupancy rates—is negative and significant throughout. While the effects on in-migration and internal migration are of the expected sign (reflecting the congestion costs incurred by migrants in obtaining new lodging), the impact on out-migration defies simple explanation. One might be tempted to conclude that high population density may be picking up the influence of access to urban amenities, whose attractiveness attenuates individuals' propensity to move out-of-state, but specification II shows that occupancy rates exert a much stronger negative influence.[19] This result, which means that the rate of out-migration is declining in the average number of persons per unit, is suggestive of a "life-cycle" effect, whereby small households, comprising singles or couples without children, have a higher propensity to make out-of-state moves. Controlling for spatial autocorrelation, *popden*'s effect on out-migration becomes insignificant, but the negative direct effects of occupancy and unit density remain. The weaker negative effect of lagged population density (and, in specification II, unit density) is consistent with the attenuating influence of congestion costs in neighboring states' housing markets to emigration there.

We find that unemployment has a negative and significant influence on internal migration, but an insignificant effect on in- or out-migration. It is well known that unemployment exerts two countervailing influences on migration—on one hand it reduces households' labor income, and with it the resources necessary to undertake the pecuniary costs of relocation, while on the other hand it acts as a psychic "push" factor, simultaneously inhibiting in-migration and inducing residents to emigrate in search of employment. The results indicate that the former pecuniary effect seems to be the dominant factor. The spatial lag of unemployment has positive and significant effects on both immigration and, to a lesser degree, emigration. The former reflects the influence of relative economic conditions in neighboring states on the propensity of residents of other states to undertake cross-border moves, while the second indicates a regional phenomenon, namely economically-induced migration away from groups of contiguous states which are economically depressed. The estimates of the effect of the average wage, while generally not significant, tend to corroborate this story: they have a positive and

---

[19] However, the effect of occupancy rates on out-migration is weaker than on both in-migration and internal migration, and the effect of unit density on out-migration understates that on internal migration, as one might expect based on intuition.

significant impact on the propensity to move out-of-state, which presumably involves larger expenditures than internal migration.[20]

12.5 Concluding Remarks

The signal benefit of this research is the creation of a simple, transparent, theoretically and methodologically rigorous simulation model that is suitable for a potentially broad range of applications. Our approach has a number of advantages in this regard, including the following. First, the Hicks-neutral shift parameter ψ may be specified to decline from its calibrated value in the base period one period to the next according to projections of productivity growth by industry, facilitating inquiry into the effect of technological progress within industries on the spatial patterns of economic growth. Secondly, because the distribution of industry tax rates in different states affects the spatial pattern of production, the modeler, by changing τ_Y as a policy variable, can explore the impact of future state tax policy scenarios on output and employment. Thirdly, since the consumption of land in the creation of new housing is explicitly modeled, it will be possible to simulate the effect of policies such as limitations on the conversion of non-urban land on county level population growth and economic growth.

The modeling framework is particularly beneficial for conducting environmental analysis for a number of reasons. Because of the relatively detailed distributions of economic activities it will be possible to apply emissions factors to generate spatial patterns of criteria pollutant emissions. This will be useful in identifying areas of high exposures and as input to atmospheric models for ozone and acid deposition. Further, the model may be extended to produce estimates of transportation activity levels and emissions from mobile sources, as described in the Appendix. In addition, the model, by explicitly representing industries' use of intermediate inputs, sheds light on the potential for macro-level climate change policy to affect both regional growth and the spatial distribution of secondary air pollution benefits from reduced combustion activity. A carbon tax (τ_C) can be simply represented as the additional term $\tau_C\varepsilon_C(j_C)$ in the unit cost function of the model's fossil fuel sectors (j_C) differentiated according to the average carbon emission coefficients (ε_C) on these sectors' outputs. Moreover, the income effects of this tax are easily represented by including the revenue that it generates ($\sum_{j_C} \tau_C \varepsilon_C(j_C) Y(j_C)$) as an additional term in equation (12.11). This feedback facilitates investigation of the spatial impacts of double dividend policies use the revenue from τ_C to lower τ_Y.

---

[20] Absent controls for spatial autocorrelation, wages have a counterintuitively strong negative impact on internal migration, but this effect disappears with the inclusion of the spatial lag.

Table 12.2 Seemingly Unrelated Regressions for State-Level In-, Out- and Internal Migration

	(I)			(II)			(III)			(IV)		
	inm	*outm*	*intm*	*inm*	*outm*	*intm*	*inmfrac*	*outmfrac*	*intmfrac*	*inmfrac*	*outmfrac*	*intmfrac*
pop	0.834	0.806	1.330	0.826	0.803	1.331	-0.166	-0.194	0.330	-0.174	-0.197	0.331
	(0.055)***	(0.027)***	(0.048)***	(0.053)***	(0.027)***	(0.048)***	(0.055)***	(0.027)***	(0.048)***	(0.053)***	(0.027)***	(0.048)***
unemp	-0.086	0.007	-0.404	-0.128	-0.010	-0.398	-0.086	0.007	-0.404	-0.128	-0.010	-0.398
	(0.198)	(0.099)	(0.174)**	(0.194)	(0.098)	(0.176)**	(0.198)	(0.099)	(0.174)**	(0.194)	(0.098)	(0.176)**
wage	0.327	1.022	-0.851	0.257	0.994	-0.840	0.327	1.022	-0.851	0.257	0.994	-0.840
	(0.402)	(0.201)***	(0.354)**	(0.393)	(0.198)***	(0.355)**	(0.402)	(0.201)***	(0.354)**	(0.393)	(0.198)***	(0.355)**
popden	-0.089	-0.096	-0.135				-0.089	-0.096	-0.135			
	(0.041)**	(0.020)***	(0.036)***				(0.041)**	(0.020)***	(0.036)***			
occu				-0.166	-0.127	-0.123				-0.166	-0.127	-0.123
				(0.060)***	(0.030)***	(0.054)**				(0.060)***	(0.030)***	(0.054)**
unitden				-0.073	-0.090	-0.137				-0.073	-0.090	-0.137
				(0.041)*	(0.020)***	(0.037)***				(0.041)*	(0.020)***	(0.037)***
Constant	-1.937	-4.146	-1.761	-1.491	-3.968	-1.827	-1.937	-4.146	-1.761	-1.491	-3.968	-1.827
	(1.297)	(0.647)***	(1.140)	(1.287)	(0.649)***	(1.163)	(1.297)	(0.647)***	(1.140)	(1.287)	(0.649)***	(1.163)
Obs.	50	50	50	50	50	50	50	50	50	50	50	50
R-sq.	0.87	0.97	0.95	0.88	0.97	0.95	0.40	0.73	0.53	0.43	0.74	0.54

All variables in logarithms. Standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12.3 SUR Results for State-Level In-, Out- and Internal Migration: Effect of Spatial Lags

	(I)			(II)			(III)			(IV)		
	Inm	*outm*	*intm*	*inm*	*outm*	*intm*	*inmfrac*	*outmfrac*	*intmfrac*	*inmfrac*	*outmfrac*	*intmfrac*
pop	0.779	0.773	1.309	0.779	0.780	1.288	-0.221	-0.227	0.309	-0.221	-0.220	0.288
	(0.068)***	(0.031)***	(0.061)***	(0.067)***	(0.032)***	(0.060)***	(0.068)***	(0.031)***	(0.061)***	(0.067)***	(0.032)***	(0.060)***
unemp	-0.296	-0.104	-0.300	-0.321	-0.111	-0.293	-0.296	-0.104	-0.300	-0.321	-0.111	-0.293
	(0.211)	(0.097)	(0.187)	(0.209)	(0.100)	(0.185)	(0.211)	(0.097)	(0.187)	(0.209)	(0.100)	(0.185)
wage	0.501	1.100	-0.644	0.455	1.064	-0.585	0.501	1.100	-0.644	0.455	1.064	-0.585
	(0.441)	(0.204)***	(0.393)	(0.437)	(0.209)***	(0.387)	(0.441)	(0.204)***	(0.393)	(0.437)	(0.209)***	(0.387)
popden	-0.012	-0.039	-0.136				-0.012	-0.039	-0.136			
	(0.059)	(0.027)	(0.053)***				(0.059)	(0.027)	(0.053)***			
occu				-0.115	-0.089	-0.111				-0.115	-0.089	-0.111
				(0.064)*	(0.031)***	(0.057)*				(0.064)*	(0.031)***	(0.057)*
unitden				-0.029	-0.058	-0.105				-0.029	-0.058	-0.105
				(0.054)	(0.026)**	(0.048)**				(0.054)	(0.026)**	(0.048)**
lag(*pop*)	-0.017	0.015	-0.166	-0.028	-0.003	-0.104	-0.017	0.015	-0.166	-0.028	-0.003	-0.104
	(0.093)	(0.043)	(0.083)**	(0.102)	(0.049)	(0.090)	(0.093)	(0.043)	(0.083)**	(0.102)	(0.049)	(0.090)
lag(*unemp*)	0.982	0.491	-0.085	0.943	0.541	-0.271	0.982	0.491	-0.085	0.943	0.541	-0.271
	(0.390)**	(0.180)***	(0.347)	(0.395)**	(0.188)***	(0.349)	(0.390)**	(0.180)***	(0.347)	(0.395)**	(0.188)***	(0.349)
lag(*wage*)	-0.454	0.047	-0.916	-0.879	-0.369	-0.386	-0.454	0.047	-0.916	-0.879	-0.369	-0.386
	(0.790)	(0.365)	(0.703)	(0.738)	(0.352)	(0.652)	(0.790)	(0.365)	(0.703)	(0.738)	(0.352)	(0.652)
lag(*popden*)	-0.125	-0.118	0.126				-0.125	-0.118	0.126			
	(0.085)	(0.039)***	(0.075)*				(0.085)	(0.039)***	(0.075)*			
lag(*occu*)				-0.532	-0.120	-0.599				-0.532	-0.120	-0.599
				(0.676)	(0.322)	(0.598)				(0.676)	(0.322)	(0.598)
lag(*unitden*)				-0.036	-0.048	0.038				-0.036	-0.048	0.038
				(0.051)	(0.024)**	(0.045)				(0.051)	(0.024)**	(0.045)
Constant	-1.819	-5.169	2.342	0.408	-3.496	0.614	-1.819	-5.169	2.342	0.408	-3.496	0.614
	(2.495)	(1.153)***	(2.222)	(2.191)	(1.045)***	(1.937)	(2.495)	(1.153)***	(2.222)	(2.191)	(1.045)***	(1.937)
Obs.	50	50	50	50	50	50	50	50	50	50	50	50
R-sq.	0.89	0.98	0.96	0.89	0.98	0.96	0.51	0.81	0.61	0.51	0.80	0.62

All variables in logarithms. Standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%.

References

Agyemang-Duah K, Anderson WP, Hall FL (1995) Trip generation for shopping travel. Transportation Research Record 1493:12-20

Babiker MH et al (2001) The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Revisions, Sensitivities, and Comparison of Results. MIT joint program on the science & policy of global change report No. 71, Cambridge MA

Bartelsman E, Caballero RJ, Lyons R (1994) Customer and supplier driven externalities. American Economic Review 84:1075-1084

Bartik T (1985) Business location decisions in the United States: Estimates of the effects of unionization, taxes, and other characteristics of states. Journal of Business and Economic Statistics 3:14–22

Black D, Henderson JV (1999) Spatial evolution of population and industry in the United States. American Economic Review 89:321-327

Brooke A, Kendrick D, Meeraus A, Raman R (1998) GAMS: A user’s guide. GAMS Development Corp., Washington DC

Becker R, Henderson JV (2000) Effects of air quality regulations on polluting industries. Journal of Political Economy 108:379-421

Becker G, Mills E, Williamson JG (1992) Indian urbanization and economic growth since 1960. Johns Hopkins Press, Baltimore

Caballero RJ, Engel EMRA, Haltiwanger JC (1995) Plant-level adjustment and aggregate investment dynamics. Brookings Papers on Economic Activity 2:1-39

Campbell PR (1996) Population Projections for States by Age, Sex, Race, and Hispanic Origin: 1995 to 2025. U.S. Bureau of the Census, Population Division, PPL-47.

Cooper R, Haltiwanger JC, Power L (1999) Machine replacement and the business cycle: Lumps and Bumps. American Economic Review 89:921-46

Coughlin C, Terza J, Arromdee V (1991) State characteristics and the location of foreign direct investment within the United States. Review of Economics and Statistics 73:675–683

Dawkins C, Srinivasan TN, Whalley J (2001) Calibration. In: Heckman J, Leamer E (eds) Handbook of Econometrics Vol. 5. Elsevier Science, Amsterdam, pp 3653-3703

Doms M, Dunne T (1998) Capital adjustment patterns in manufacturing plants. Review of Economic Dynamics 1: 409-29.

Duranton G, Puga D (2004) Microfoundations of urban agglomeration economies. In: Henderson JV, Thisse JF (eds) Handbook of urban and regional Economics Vol. 4. North Holland, Amsterdam.

EPA (2003) Compilation of air pollutant emission factors, AP-42, 5th ed., Vol. 1.

Fan W, Treyz F, Treyz GI (2000) An evolutionary new economic geography Model. Journal of Regional Science 40:671-695

Ferris MS, Pang JS (1997) Engineering and economic applications of complementarity problems. SIAM Review 39:669-713

Figueiredo O, Guimaraes P, Woodward DP (2002) Modeling industrial location decisions in U.S. counties. NIMA working paper 18, Universidade do Minho.

Friedman J, Gerlowski D, Silberman J (1992) What attracts foreign multinational corporations? Evidence from branch plant location in the United States. Journal of Regional Science 32, 403–418.

Gallin JH (in press). Net migration and state labor market dynamics. Journal of Labor Economics

Garofalo G, Yamarik S (2002) Regional Convergence: Evidence from a New State-by-State Capital Stock Series. Review of Economics and Statistics 84:316-323

Ginsburgh V, Keyzer M (2002) The structure of applied general equilibrium mModels. MIT Press, Cambridge, MA

Greenstone M (2002) The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1997 clean air act amendments and the census of manufactures. Journal of Political Economy 110:1175-1219

Greenwood MJ, Hunt GL, RickmanDS, Treyz GI (1991) Migration, regional equilibrium, and the estimation of compensating differentials. American Economic Review 81:1382-1390.

Head K, Ries J, Swenson D (1995) Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. Journal of International Economics 38:223–247

Henderson JV (1996) Effects of air quality regulation. American Economic Review 86:789-813

Holmes TJ (1998) The effects of state policies on the location of industry: Evidence from state borders. Review of Economics and Statistics 81:314-325

Holmes TJ, Stevens JJ (2004) Spatial distribution of economic activities in North America. In: Henderson JV, Thisse J-F (eds) Handbook of urban and regional economics Vol. 4. North Holland, Amsterdam

Kelly AC, Williamson JG (1984) What drives third world city growth? A dynamic general equilibrium approach. Princeton University Press, Princeton, NJ

Kehoe TJ (1998) Social accounting matrices and applied general equilibrium models. In: Begg I, Henry SGB (eds) Applied economics and public policy. Cambridge University Press, Cambridge, pp 59-87

Kim S (1995) Expansion of markets and the geographic distribution of economic activity. The trends in U.S. regional manufacturing structure, 1860-1987, Quarterly Journal of Economics 110:881-908

Mansur A, Whalley J (1983) Numerical specification of applied general equilibrium models: Estimation, calibration and data. In Scarf H, Shoven J (eds) Applied general equilibrium analysis, Cambridge University Press, Cambridge, pp 69-127

Paul CJ, Siegel D (1999) Scale economies and industry agglomeration externalities: A dynamic cost function approach. American Economic Review 89:272-290

Perroni C, Rutherford TF (1998) A comparison of the performance of flexible functional forms for use in applied general equilibrium analysis. Computational Economics 11:245-263

Rickman DS, Shao G, Treyz GI (1993) Multiregional Stock Adjustment Equations of Residential and Nonresidential Investment in Structures. Journal of Regional Science 33:207-219

Rutherford TF (1995) Extensions of GAMS for complementarity problems arising in applied economic analysis. Journal of Economic Dynamics and Control 19:1299-1324

Rutherford TF (1999) Applied general equilibrium modeling with MPSGE as a GAMS subsystem: An overview of the modeling framework and syntax. Computational Economics 14:1-46

Stevens BH, Treyz GI (1986) A multiregional model forecast for the United States through 1995. American Economic Review 76:304-307

Sue Wing I (in preparation) Computable general equilibrium models and their use in economy-wide policy analysis: Everything you ever wanted to know (but were afraid to ask). Discussion paper.

Treyz GI, Friedlaender AF, Stevens BH (1980) The employment sector of a regional policy simulation model. Review of Economics and Statistics 62:63-73

Treyz GI, Rickman DS, Hunt GL, Greenwood MJ (1993) The dynamics of U.S. internal migration. Review of Economics and Statistics 75:209-214

Treyz GI, Rickman DS, Shao G (1992) The REMI economic-demographic forecasting and simulation Model. International Regional Science Review 14:221-253

Treyz GI (1993) Regional Economic Modeling A Systematic Approach to Forecasting and Policy Analysis. Kluwer, Boston

Woodward D (1992) Locational determinants of Japanese manufacturing start-ups in the United States. Southern Economic Journal 58:690–708

Appendix: Estimating Transportation Activity Levels and Mobile Source Emissions

Personal Transportation Emissions

We focus on road transportation, and propose a two-track empirical approach that relates vehicle emissions to VMT and driving behavior, both of which are influenced by congestion. Using cross-section data from the National Household Transportation Survey (NHTS), we first estimate a trip generation model of average annual VMT ($\tilde{T}$) as a function of income and size of urban area, and then employ the ordered logit model of Aygemang-Duah et al (1995) to apportion $\tilde{T}$ among a number of average speed categories ϑ according to geographic attributes. In county c, the probability of being in speed class ϑ thus depends on a function π of income, population density and size of urban area (both of which proxy for congestion):

$$\tilde{\pi}(\vartheta, c\{s\}) = 1\Big/\Big(1+\exp\Big[\pi\Big\langle \tilde{\rho}_N(c\{s\}), \tilde{N}(c\{s\}), \tilde{W}(c\{s\})\Big\rangle - \lambda_{\vartheta}\Big]\Big), \tag{A1}$$

where λ_{ϑ} is an estimated vector of probability cutoffs. The result is a vector of county-level speed distributions projected at each time-step as a function of the variables solved for at the county level, which serve as a proxy for driving cycle characteristics that affect emissions per mile. This allows us to use the distribution of emissions per VMT in each velocity category from the emission factors in the MOBILE6 model (ε_M) to estimate mobile emissions ($\tilde{e}_M$) by pollutant, speed class and county:

$$\tilde{e}_M(z, \vartheta, c\{s\}, t) = \tilde{N}(c\{s\}, t)\tilde{T}(c\{s\}, t)\tilde{\pi}(\vartheta, c\{s\}, t)\varepsilon_M(z, \vartheta). \tag{A2}$$

Interstate Freight Transportation Emissions

An important advantage of the proposed simulation framework is that the model's demand structure facilitates investigation of the impact of regional growth on interstate freight transportation and associated air pollutant emissions. The state-level counterpart of equation (12.8) gives the demand in each state (y_D) for subset of commodities that are transported, i_F:[21]

---

[21] Here we assume a simple lump-sum recycling rule that divides aggregate tax revenue equally among members of the population. More realistic assumptions can be made, although at the cost of increasing the model's complexity.

$$y_D(i_F,s,t) = \sum_j \bar{X}(i_F,j)y(j,s,t) + \varpi(i)\left(V(s,t) + TAXREV / N(s,t)\right)\big/P(i_F,t)\,. \tag{A3}$$

This expression allows us to use a production-constrained gravity model to approximate inter-state flows of goods *f* from production in state *s* to uses in state *r* along transportation mode *q*:

$$f(i_F,s,r,q,t) = y(i_F,s,t)y_D(i_F,r,t)\Delta(s,r,q)^{-\phi(i.q)}\Big/\sum_r\left(y_D(i_F,r,t)\Delta(s,r,q)^{-\phi(i_F.q)}\right). \tag{A4}$$

Here, Δ is the distance from *s* to *r*, calculated from the BTS North American Transportation Atlas Database, and ϕ is a measure of the friction of distance, which our tenth research task is to estimate using cross-sectional data from the 1997 Commodity Flow Survey. We propose to assign the estimated proportions of freight flows by mode to the appropriate transportation network based on shortest path routes, and to focus once again on road transport emissions, which can be easily estimated by multiplying the number of truck miles assigned to each highway link by average highway emissions factors from MOBILE6.[22]

---

[22] Additional off-interstate emissions are likely to be significant due to local congestion at origin and designation counties. To estimate them it is necessary to account for demand by county as a proportion of each interstate flow.

13 Impact Assessment of Clean Development Mechanisms in a General Spatial Equilibrium Context

Shunli Wang and Peter Nijkamp

Department of Regional Economics, Free University, Amsterdam, The Netherlands

13.1 Introduction

The climate on our earth has always been in a state of flux. It has, for instance, exhibited dramatic cycles of glacier movements in the past and it will continue to change in the future. Human beings have – with more or less success – adjusted themselves to such changes, and this has prompted creative survival strategies with far reaching social-economic and geographical implications. In the past century, the intensification of human activity on our earth has also led to man-made climatological changes, e.g., as a consequence of atmospheric pollution, in particular greenhouse gases (such as carbon dioxide, methane and nitrous oxide). A higher concentration of such gases tends to make the earth's temperature higher. If this greenhouse effect continues (as a result of further industrial growth, mobility, household consumption or agriculture), then the earth' climate may show significant changes, such as changes in the biosphere, variations in wind and weather conditions, melting of glaciers, sealevel rise, wetland loss and erosion. Clearly, a precise prediction of such variations is fraught with many uncertainties, but it is possible that the changes will be rather drastic in the long run.

The threat of climate change has caused much political concern world-wide. The Kyoto Protocol is an institutional arrangement that seeks to develop a shared global policy consensus on the necessary reduction of greenhouse gases. This protocol was adopted by many countries, viz. industrialised countries, countries with economies in transition, and developing countries. The Kyoto Protocol comprises a wide array of various measures for various groups of stakeholders. They all aim to reduce the emission of greenhouse gases, at the lowest costs, with an average of 5.2% for the developed world (with a possibility for individual deviations). The three specific policy measures foreseen are Joint Implementation (JI), Clean Development Mechanisms (CDMs) and Emission Trading (ET). A flexible combination of these three instruments might likely create the highest level of international support.

Despite intensive negotiation rounds, a consensus is still far to be reached in the international political debate on global climate change issues. Nevertheless, global warming is an on-going autonomous process in all regions of our world. In coping with this complicated process, many kinds of policy instruments and choice directions were designed. In recent years, the debate has concentrated on the question whether a mitigation policy or adaptation policy should be adopted. In this debate, one of the instruments designed in the course of the Kyoto Protocol, i.e. CDMs, may not have received the attention it deserved. CDM is a relatively modern climate change policy that is able to create a win-win situation for both developed and developing countries in coping with emission reduction targets. This policy aimed in particular at a cost-effective reduction target by improving the level of technology in non-Annex I (i.e., the developing) countries.

Currently, the main body of literature on CDM policies is largely “policy-oriented.” Much attention is given to review of climate change negotiations (see, e.g., Toman and Hourcade 2000; Michaelowa 2003; IISD 2002). Other authors, like Hardner et al. (1999), or Woerdman (2000), offered an account of usefulness of CDM policies. The majority of the literature however, provides and discusses proposals for the functioning of critical aspects of CDM policies (Parkinson et al. 1999; Noble and Scholes 2001; Geres and Michaelowa 2002; Forner and Jotzo 2002; and Schwarze and Niles 2000). In addition, several empirically-oriented studies can be found which address CDM policies on a project level (see e.g., De Leeuw and van Ierland 2003), by analyzing whether or not concrete CDM proposals are attractive from a policy point of view. In addition, there are other empirically-oriented studies that try to estimate the CDM potential (see, e.g., Bosi and Ellis 1999; Jackson 1995; Fichtner et al. 2001; Duic et al. 2003; and Halsnæs 2002). From an applied modeling perspective, the CDM market potential has been estimated by deploying inter alia specific models such as the marginal abatement cost approach (see, e.g., Zhang 2001; Jotzo and Michaelowa 2002; and Chen 2003). In the extensive literature, there are also models which capture the technology aspects of CDM (e.g., RICE; see Nordhaus and Yang 1996) or a substitution between energy commodities due to CDM policies (e.g., WorldScan; see Bollen et al. 1999a). Consequently, the research field on CDM policy is in full motion.

13.1.2 New Directions for Analyzing CDM

This chapter aims to model and evaluate CDM in a general spatial equilibrium framework, by deploying in particular an adjusted version of the GTAP-E (Global Trade Analysis Project including energy) model which incorporates energy and technology (see later). For this reason, this chapter offers an economic foundation of this instrument by discussing the underlying economic processes that would be affected by the use of this instrument. This is done by (i) discussing the modeling of necessary extensions of the CDM approach, (ii) presenting a general sketch of economic impacts of CDM strategies, (iii) providing a method for the calibration of CDM extensions in a general spatial equilibrium framework, in particular the GTAP-E model, and (iv) discussing the numerical results on the economic impacts of this instrument in different global regions.

Modeling CDM in a general spatial equilibrium framework needs a formulation of several assumptions on important aspects of this instrument. In order to model CDM policies in a spatial CGE (Computable General Equilibrium) model like GTAP-E, this chapter uses a simple working definition of CDM, which is: "CDM is an instrument designed to facilitate the cooperation between source countries and receiving countries, which enables (the producers in) the source countries to meet a part of their greenhouse gas (GHG) emission reduction targets by obtaining a certain amount of CERs (Certified Emission Reductions) as a result of their effort to increase the technological level (of the producers) in the receiving countries." Consequently, CDM policies are operationalized as follows:

- producers in source countries raise their CDM investments by levying 'self-imposed' taxes. These investments are collected by a CDM coordinator;
- producers in receiving countries receive these investments from the CDM coordinator, and will be encouraged to induce 'experience' technological change in their production processes;
- technological change leads, ceteris paribus, to a reduced use of inputs and thus to emission reductions in the receiving countries, i.e., it is biased towards energy-efficient technologies;
- emission reductions in the receiving countries due to technological change in these countries will be assigned to producers in source countries and allow these producers to acquire a certain amount of CERs.

The main features of modeling CDM in our study are:

- it proposes the use of a 'conditional technological change function' for capturing the technological change aspect of CDM (i.e., this function presents the conditional relationship between technological change and CDM investments).
- it models only CERs in terms of carbon emission reductions (although the reduction of other greenhouse gasses (GHGs) may be dealt with analogously).
- it neglects a meso-project approach to CDM, as the level of aggregation in an CGE framework is higher (i.e., a regional level).
- it introduces, for modeling purposes, an aggregate CDM distribution mechanism. This is done in the form of a 'CDM coordinator', who is assumed to (i) distribute the CDM investments, collected from self-imposed taxes from representative producers in Annex I countries, to representative producers in non-Annex I countries; and to (ii) attribute CERs to producers in Annex I countries. In terms of institutions, this approach is similar to what is called the `multilateral approach'. The reason for this multilateral modeling ap-

proach is simplicity of modeling, as modeling based on decisions in bilateral structures requires complicated decision rules and information that is not available.

It should be noted that in modeling decisions in source countries, the so-called development additionality condition is not explicitly taken into consideration. Implicitly, all CDM activities, through their beneficial impact on the state of technology in receiving countries, satisfy this requirement. For this reason, the producers in Annex I countries would be compensated through receiving CERs. The value of CERs depends on the price of emissions, which in turn is determined by the emission restriction targets and other rules of the climate change regime. Also, we neglect the net impact of afforestation and deforestation, which are, for example, capped at 1% of a source country's base-year emission levels.

This chapter is set up as follows. Section 13.2 gives a brief introduction in the CGE model context, the GTAP-E model. Section 13.3 discusses CDM extensions in the form of changes in behavioral rules of relevant economic actors, i.e. the producers and the CDM coordinator. Section 13.4 presents a general sketch of economic impacts of CDM. Section 13.5 shows a method for calibration of parameters for CDM extensions in the GTAP-E model, so that in Section 13. 6, numerical results for various CDM regimes could be presented. This chapter concludes in Section 13.7.

13.2 A Brief Introduction in CGE Model Context: GTAP-E

13.2.1 Introduction

GTAP-E is an applied general equilibrium model with a database for the world economy as a whole. As were shown by Kremers, Wang, and Nijkamp (2002), GTAP-E could be characterized as a spatial-extended CGE model for analyzing the interaction between the ecological and the economic systems. The modeling of the economic system is according to CGE analysis. Partly owing to the current computational capacities, the CGE models are widely applied tools for analyzing the complex relationships between various markets, and connections between different countries or regions. CGE analysis involves, as Shoven and Whalley (1984) pointed out, a numerically-specified general-equilibrium model for policy evaluation. Generally, it is accepted that the earlier works of Johansen (1960), Harberger (1962), and Scarf (1967) form the starting point for numerically-based general equilibrium analysis (see Shoven and Whalley 1984; Shoven and Whalley 1992; Ginsburgh and Keyzer 1997).

13.2.2 Model Context: The Economic Structure in GTAP-E

As discussed by Wang, Nijkamp and Verhoef (2003), an archetype economic system can be described by defining a set of inputs, a set of outputs, a set of production processes, a set of utility functions and a set of feedback relationships. The relationship between the set of inputs and the set of outputs is described by the set of production processes. Furthermore, the amount of the outputs is evaluated by the economic subjects through the utility function. From an economic point of view, it is interesting to extract the optimal solution, i.e. the (general) equilibrium that maximizes the overall utility.

In CGE models, this is achieved by decomposing the economic decision process into two sub-problems. The first sub-problem is solved by the producers, who maximize their profits subject to the technological feasibilities of the production process. The second sub-problem is solved by the consumers, who maximize their utility subject to their budget constraint. Typically, on both sides of the market, price-taking behavior is assumed.

In a CGE model, a consumer (↻) maximizes the utility under the budgetstriction, i.e.:

∓↦¬ ∪↻∼¬⌒

such that:

$$\sum_{x} P_{x}^{*} x_{i}^{*} \leq \sum_{k} P_{k}^{*} k_{i}^{*} + \sum_{j} \theta_{i,j} \Pi_{j}^{*} \qquad (13.1)$$

where ∪↻∼¬⌒ stands for the utility function[1] of good ¬ by consumer ↻, ‖¬ for price of good ¬, and ¬↻, due to notational simplicity, for quantity of good ¬ consumed by consumer ↻, ‖ for price of endowment , and ↻ for the for quantity of endowment of consumer ↻, θ↻ ↺ for the share of consumer ↻ in firm ↺, Π↺ for the profit from firm ↺, and Σ is the symbol for summation. In addition, * denotes that this relationship holds in the equilibrium situation.

In GTAP-E, there is a representative consumer for each region in the world. These consumers spend their income on the consumption of goods, for which three aggregate goods, consumer goods, government expenditure, and savings, are formulated (see Hertel and Tsigas, 1997 for more detail). Like other general equilibrium models, a representative consumer earns their income from their resources (e.g. natural resources, capital and labour), which are inputs for production.

On the production side of the market, CGE models such as GTAP-E often deduce producer's behavior from the assumption of profit maximization calculations of the producers, i.e.:

---

[1] For notational ease, we use ¬ for the quantity as well as index of good ¬. This holds also for the use of inputs in other parts of this paper.

$$\max \ \Pi_j^* = P_j^* Q_j^* - \sum\nolimits_k P_k^* k_j^*$$

subject to: (13.2)

$$\because_{\circlearrowleft} \ll \succ_{\circlearrowleft\smile} \frown$$

Where $\Pi_{\circlearrowleft}$ stands for the profit of firm $\circlearrowleft$, $\|_{\circlearrowleft}$ for the price of outputs from firm $\circlearrowleft$, $\because_{\circlearrowleft}$ for the quantity of output by firm $\circlearrowleft$, index the endowment , and $_{\circlearrowleft}$ for the quantity of endowments used by firm $\circlearrowleft$. In addition, the producer j is restraint by the production possibility function $\succ_{\circlearrowleft\smile} \frown$.

Of course, one important restriction is the production possibilities within the economy. In GTAP-E, the production possibilities are described by a nested Constant Elasticity of Substitution function (see Hertel and Tsigas 1997 for more detail on the formulation of the demand side of the economy (consumers' behavior), Hertel 1999 for the supply side of the economy (producers' behavior), and see Kremers, Wang and Nijkamp 2002 for a discussion of GTAP-E in comparison to other CGE models).

13.2.3 Link to the Ecological System in GTAP-E

As discussed by Kremers, Wang, and Nijkamp (2002), GTAP-E model is a spatially-extended CGE model for analyzing the interaction between the ecological and the economic systems. For analyzing the effectiveness of climate policy, energy-capital substitution is an important issue (Truong 1999). In GTAP-E, the capital energy composite has a CES form with capital and energy as inputs. In addition, the energy nest is further produced by a multilevel structure of electric energy and a nest of non-electric energy. The non-electric nest is composed of coal and a non-coal nest. The non-coal nest is composed of the inputs gas, oil and petroleum products. Each of the energy inputs, viz. coal, gas, oil, petroleum and electricity, has a domestic variant and a imported variant (see also Kremers *et al.*, 2002).

Relevant for analyzing climate change issues is also that GTAP-E includes a carbon emission module and the associated energy taxes. Truong (1999) has not fully documented this module. The carbon emission module of GTAP-E assumes that there is a fixed relationship between the emission of, for example, one of the GHGs, viz. CO_2, and the use of the underlying energies. This relationship may be formalized by the following mathematical notation:

$$\llcorner_{\nearrow} \ll \chi_{\nearrow} \nearrow \qquad (13.3)$$

which states that the unintended economic output ($\_{\nearrow}$, i.e. CO_2–emissions) from the use of emission-generating input ($\nearrow$) is proportional with a fixed coefficient (denoted as$\chi_{\nearrow}$).

A strong aspect of GTAP-E for economic analysis is that it is able to trace back the carbon emissions to six categories of sources in each region. These categories are associated with source-specific carbon taxes, i.e. taxes on imported and domestic energy commodities purchased by (i) private households, (ii) by government households; and (iii) by producers as intermediate.

In carrying out the quota-policy, the ability of GTAP-E to trace carbon emissions back to the six categories of sources results in the model's flexibility to calculate the marginal reduction costs per ton of CO_2 for various sources. In GTAP-E, this is done by deriving the marginal reduction costs for domestic produced traded commodities or for imported intermediate commodities. Furthermore, this marginal reduction cost could be applied to be: (i) uniform across all regions, or (ii) uniform within the regions. In addition, it is also possible to specify these uniform taxes for: a) every kind of energy commodities; b) the domestic energy demands; or (c) the imported demands. In addition, the GTAP-E model is able to deal with the following situations:

- the CO_2 emission at the world level is exogenously given. Then, a uniform CO_2 tax over all regions will result;
- the CO_2 emission for each region is exogenously given. Then, the level of carbon tax may vary across regions; and
- the CO_2 emission for Annex I countries is exogenously given, but is variable for non-Annex I countries. In this case, the model gives the level of regional carbon tax for the Annex I countries and the CO_2 emission in the non-participating regions.

13.3 Behavioral Rules for Clean Development Mechanism

13.3.1 Introduction

As known, from a GE-perspective, policies or changes in one single market would also influence the equilibrium on other markets, as general equilibrium takes secondary feedback effects into account. For this reason, it is important to discuss how CDM affects the economic structure of a single market, and the impact on the economy as a whole. This section will explore necessary theoretical extensions which should be made for modeling CDM in a general spatial equilibrium model. In the first instance, CDM activities would affect producer's behavior. This will be treated in subsection 13.3.2. In addition, CDM needs other extensions of the basic model, such as how to link the producers. This will be discussed in subsection 13.3.3.

13.3.2 Producer's Behavior

As Annex I countries need not necessarily ratify the Kyoto Protocol (KP), this chapter will use the term `non-participation regions' to refer to a subset of the Annex I regions that are assumed not to carry out the relevant policies needed to achieve the emission reductions agreed in the KP. For this reason, we have three types of producers behavior under CDM activities. This will be discussed in the following subsections.

Producer's behavior in the non-participating countries as reference situation. To begin with, we discuss representative producer's behavior in the non-participating countries. This is also the reference behavior for producer's facing emission restrictions and CDM possibilities. In this reference situation, the producers will not take emissions into account in their decisions. The representative producer is assumed to minimize the costs under restriction of a standard constant-returns-to-scale production function. Furthermore, other standard assumptions of a CGE model apply, i.e. the representative producer is a price-taker, and there are markets which determine the supply of inputs for the representative producers as well as the demand for their outputs.

In a simplified economy of two inputs (economic renewable good … and natural resource $\nearrow$), we have that a producer's $\circlearrowleft$ minimizes the costs under the constraint of the production possibilities as follows:

$$\vdots\circlearrowleft\text{�}\ \coprod_{\circlearrowleft}\ll\|_{\nearrow}\nearrow_{\circlearrowleft}\quad\|_{\ldots}\cdots_{\circlearrowleft} \tag{13.4}$$

subject to:

$$Q_j = A_j\left(\alpha_j K_j^{\phi_j} + (1-\alpha_j)Z_j^{\phi_j}\right)^{1/\phi_j} \tag{13.5}$$

where P_K and P_Z stands for the prices of the inputs K_j (economic input) and Z_j (ecological input), Q_j for output, A_j for the technological parameter, α_j for the share of each representative producer's inputs in the production, and finally, ϕ_j for the substitution parameter.

Producer's behavior in Annex 1 countries. In case of CDM, producers in Annex 1 countries would be affected by 2 behavioral changes, in comparison with the reference situation like that of the producers in non-participating countries. These changes are due to the imposed emission reductions and due to CDM. The minimization problem of the producer becomes:

$$\vdots\circlearrowleft\text{�}\ \coprod_{\circlearrowleft}\ll\|_{\nearrow}\nearrow_{\circlearrowleft}\quad\|_{\ldots}\cdots_{\circlearrowleft}\quad\tau_{\circlearrowleft\smile}\cdots_{\circlearrowleft}\quad\nearrow_{\circlearrowleft\frown} \tag{13.6}$$

subject to Equation (13.5) and

$$\sum\nolimits_Z \chi_Z Z_j \le E_j^{\max} + E_j^{cer}, \tag{13.7}$$

$$E_j^{cer} = E_j^{cer}(V_j^{\tau}), \tag{13.8}$$

where subscript ৹ stands for individual producer ৹, τ for the rate of self-imposed tax, $E_j^{\max}$ stands for the emission restriction target, E_j^{cer} for the CERs and V_j^{τ} for the monetary value of self-imposed tax. It should be noted that Equation (7) consists of two element. First, there is an emission restriction target of $E_j^{\max}$. Second, there is a mitigation effect of the amount of CERs which could be obtained from CDM activities in the order of E_j^{cer}. The calculation of the amount of CERs will be discussed in the next subsection.

In order to understand the impact of CDM activities for the producers in the Annex 1 countries, Figure 1 presents the iso-cost curve under various regimes, i.e. the Business-as Usual (BaU) case, the emission restriction case and the CDM case. It clarifies the impacts of CDM for the producer’s behavior in terms of demand for inputs. In the BaU case, let ’s assume that the equilibrium is reached at point ↨, with a production level ⁖↨↦-, and the corresponding demand for …↨↦- and ⎵↨↦- . Price ratio of … and ↗ corresponds to the marginal productivity of both inputs at the level of ⁖↨↦- . Now, let us consider a situation wherein an emission restriction target is effective and let ‘s assume that the level of use of … does not change. Then, in this emission restriction case, the output will be ⁖↤⊙↤. In addition, a correction is taken into account for the price ratio of … and ↗. The CDM case will, in a ceteris paribus situation, have an impact that will lie between the BaU and the restriction case. This is due to the mitigation effects of CERs on the emission restriction targets.

Representative producers in the non-Annex I countries and conditional technological change. CDM affects technological change in the non-Annex 1 countries. For analyzing technological change of CDM, we have the empirically-based abatement cost literature (Jackson, 1995, Begg *et al.*, 2001, de Leeuw and van Ierland, 2003, Gerlagh *et al.*, 2000, Kremers *et al.*, 2000) and the theoretically-based literature on induced technological change (ITC). The literature on ITC focuses on endogenizing technology in economic models (Weyant and Olavson, 1999, Lőshel, 2002), including technological transfer, technological spillover and endogenous growth approaches. Technology is, however, a broad concept (for an overview, see, e.g., Grübler, 1998; Jaffe *et al.*, 2000). It can, for example, encompass hardware such as machines or tools (Grübler, 1998), and software such as knowledge for making R&D (Romer, 1986; and Lucas, 1988), or knowledge of using the hardware (learning-by-doing, Solow, 1957). So, ITC resulting from

CDM investment can be a transportation of more energy-efficient cars to other countries, the transmission of knowledge on the use of other products (like solar energy), or both.

In modeling the relationship between technological change in the non-Annex I countries and the incentive from the Annex 1 countries, i.e. the volume of CDM investments, this chapter applies a 'conditional technological change' approach. In other words, we only concentrate on that part of technological change in the non-Annex 1 countries that is directly assumed to be affected by CDM investments. This treatment is in line with the basic step in the endogenous growth literature, i.e. the assumption that additional investment in technology affects the level of technology (Aghion and Howitt, 1998). It should, however, be noted that the endogenous growth theory is formulated in a dynamic setting and is more complex.

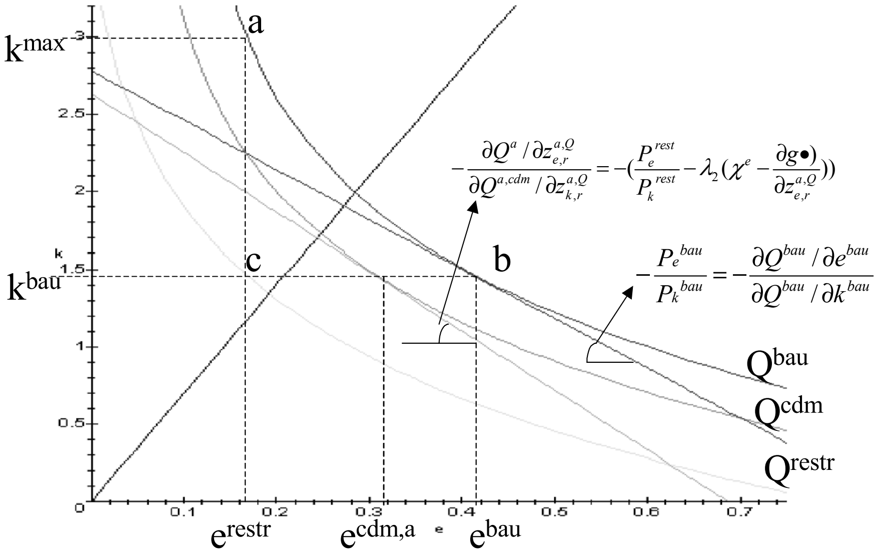


Fig. 13.1. Iso-cost curve under various regimes.

The conditional relationship between CDM investment (V^{τ}) and technological change is formulated as follows:

$$dA_Z = \xi_Z(V^{\tau}) \tag{13.9}$$

where dA_Z stands for technological change (improvement of efficiency or productivity) of input Z, and $\xi_Z(\bullet)$ represents the technological change function, which is conditioned on the CDM investments (V^{τ}).

Given this conditional technological change function, the cost minimization problem of the representative producer in non-Annex 1 countries becomes: ⁚○� ∐₀≪‖↗↗₀ ‖…‥₀ , subject to: Equation (5) and Equation (9). It should be noted that the representative producer take Equation (9) as given, i.e. it is exogenous.

The impact of this extra condition (or possibility because of better technologies) could be clarified by Figure 13.2, which presents the iso-cost curve in case technological change. Figure 13.2 shows that, as a result of new technologies, the producer's conditional demand and production frontier is slightly changed. In this figure, the conditional demand and iso-cost curve is depicted as ·.·‡↦–. Technological change may occur in two ways. First, the total factor productivity is increased. This is depicted by curve ·.·↦. Second, technological change is biased, i.e. productivity increases if a specific factor (i.e. ↗) is increased. This is depicted by curve ·.·ξ.

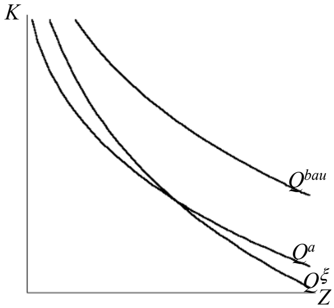


Fig. 13.2. Iso-cost curve in case technological change.

The curves ·.·‡↦–, ·.·↦ and ·.·ξ depict the same amount of outputs, i.e. ·.·‡↦– ≪ ·.·↦ ≪ ·.·ξ. The curve ·.·↦ is a shift of the curve ·.·‡↦– towards the origin, as an increase in total factor productivity implies a decrease of both inputs to the same extent. ·.·ξ is more biased towards ↗, as we assume a biased technological change here. This means that, for a small amount of use of the ecological input ↗, an increase in total factor productivity would result in less use of ecological input, while, in the case of a large amount of use of the ecological input ↗, biased technological change would lead to less use of the ecological input. The indifference point is the crossing of curves ·.·ξ and ·.·↦, which depends on the parameters of the production function, i.e. α ξ and the conditional technological change function, i.e. ○≙ or ○≙↗.

13.3.3 Linking the Producers: CDM Coördinator

CDM coördinator. In this section, we will discuss the model representations of financial flows from Annex I countries to a representative producer in a non-Annex I country. As a matter of simplification, first, we assume that there is an international coordinator who (i) collects all CDM investments from Annex I countries and (ii) distributes these investments to representative producers in non-Annex I countries. This is analogous to the role of the global bank in the case of savings. Secondly, we assume that this CDM coordinator administrates CERs according to the baseline and accounting rule, and then attributes these CERs to representative producers in the Annex I countries. Thirdly, we assume that this coordinator distributes CDM investments according to a certain rule of thumb, which will be discussed next.

Tax coordination. This international coordinator collects all CDM tax revenues from Annex I countries. The share of CDM investments from representative producer j in Annex I countries r in the total of CDM investments ($\theta_{j,r}$) is equal to region r′s contribution to the total CDM investment, i.e.:

$$\theta_{j,r} = \frac{V_{j,r}^{\tau,cdm}}{V^{T}} \tag{13.10}$$

where $V^{T} = \sum_{j}\sum_{r} V_{j,r}^{\tau,cdm}$ stands for the total amount of CDM investments.

This international coordinator distributes these CDM investments to various sectors in Annex I countries by maximizing the amount of CERs, under the constraint of the budget (i.e. total CDM investments – V_{s}^{τ}). From this formulation, the distribution rule for a CES-production function and a decreasing marginal productivity function for conditional technological change can be derived, as follows:

$$\theta_{Z,j,s} = \frac{\chi_i Z_{i,j,s}^{1} A_{i,j,s}^{1}}{\sum \chi_i Z_{i,j,s}^{1} A_{i,j,s}^{1}}$$

This distribution rule states that the distribution of CDM investments for conditional technological change of input Z in sector j in region s is a share of the emissions of the emission-generating inputs, corrected by parameters which are responsible for the effectiveness of the investment to generate technological changes.

Note that this distribution rule is efficient in terms of the objective of maximizing the amount of CERs (which is equivalent to minimizing emissions in the non-Annex I countries). However, in the case of taking other objectives, e.g. other welfare measures, like consumption or levels of production, into account, the distribution rule would become different. In other words, in the case of other objectives given to the CDM coordinator, the above-mentioned distribution rule would no longer be optimal.

Furthermore, as we do not assume transaction costs, entrance barriers, or other forms of organizational costs for initiating CDM projects, there is balance of costs and revenues for this international coordinator. Thus, by definition, we have:

$$\sum_j \sum_r \theta_{j,r} V^T = \sum_i \sum_j \sum_s \theta_{i,j,s} V^T$$

which states that CDM investments collected from Annex I countries equal the total CDM investments distributed to sectors in non-Annex I countries.

CERs: How to locate emission reductions due to technological change? A complicating factor of CDM policies is the emission accounting in the presence of the CERs. In particular, the participating Annex I countries could mitigate their voluntarily-agreed emission reduction target by CERs obtained from CDM investments in the non-Annex I countries, which are able to emit unintended outputs into the ecological system without any constraint, as no emission restriction is imposed. We could clarify this issue by discussing accounting rules.

For the calculation of the CERs, a distinction should be made between the actual emissions and the accounted emissions. For reasons of simplicity, this discussion assumes that countries are representative actors in the accounting of emissions. The same reasoning is applied to the level of individual input in the implementation of the model. Ideally, the accounted emissions should be equal to the emissions in the case of emission restriction target, as that will be the level of emissions without CDM policies. In the emission restriction regime, it would be agreed that, for each individual country $r \in \{\text{Annex I}\}$, the total domestic emissions (E_r) should not exceed the voluntarily-agreed target level of emissions ($E_r{}^{\max}$), i.e.: $E_r \leq E_r{}^{\max}$.

Of course, emission reductions in the non-Annex I countries could have multiple causes, e.g. autonomous technological change, demand shifts, supply shifts, structural change, etc. Therefore, for the attribution of CERs, additional assumptions should be made. First, only those emission reductions in the non-Annex I countries that are a result of technological change which could be traced back to the Annex I countries' contribution, would be attributed as CERs. In addition, we assume that the calculation should be based on the difference between the emissions in the case of technological change and in the case of no technological change, given the same level of production. In particular, we assume that the baseline is set at the ex-post level of the conditional demand for emission generating inputs measured by the ex-ante technological level, i.e.:

$$\bigcirc\llcorner_{\leftarrow}{}^{\Updownarrow\bigcirc\mp} \ll \Sigma\ \chi\ \swarrow \nearrow \smile \triangleq \frown \qquad \nearrow \smile \triangleq \frown \searrow \tag{13.11}$$

which states that the emission reductions as a result of technological change (dE^{cdm}) are calculated as the difference between the ex-post level of demand for ecological inputs ($Z^1(A^1)$), and the baseline ($Z^1(A^0)$). The ex-post level of demand depends, among other things, on the level of technology (A^1). The baseline is calculated as the ex-post level of demand for ecological inputs at the technological level of BaU (A^0), i.e. no technological change.

In a constant returns to scale production function, we have that the term dE^{cdm} would always be negative in case of technological change, as $A^1 > A^0$ results in $Z^1(A^1) < Z^1(A^0)$.

It should, however, be noted that, though the $dE^{cdm}/dA < 0$ holds, this formulation does not exclude the situation that $Z^1(A^1) < Z^0(A^0)$, i.e. that emissions in the non-Annex I countries would, in spite of technological change, be higher in the ex-post situation. One reason for the existence of this situation may be due to leakage and spillover in international trade.

13.4 A General Sketch of Economic Impacts as a Result of CDM

13.4.1 Introduction

This section gives a general sketch of major economic processes that will be affected by the introduction of CDM and some theoretical, plausible impacts of these processes on a number of key economic variables. In discussing the economic processes that will be affected by CDM, and the discussion will focus on the following key aspects of CDM: (i) the emission restriction, (ii) the self-imposed tax, and (iii) technological change. We will discuss the impacts of these three elements on economic processes and focus on the impacts on (i) GDP as a proxy for economic impacts for the individual regions; (ii) terms of trade; and (iii) trade balance.

In the tables, we distinguish two kinds of commodities (which could be endowments, intermediates and outputs), i.e. emission-generating commodities and other commodities. The emission-generating commodities are coal, gas and petroleum, while the other commodities are oil, electricity and the other non-energy commodities. In GTAP, the inputs are, except for the endowments, also intermediates. In addition, we assume ↑ to denote an increase of the initial effect and ↓ a decrease. For other secondary mechanisms, '+' denotes a positive expected effect and '-' a negative one.

It should be noted that this section only considers the expected major direct effects of policy interventions. For CGE models where also substitution effects operate between sectors and regions, the effects presented in this section will not be complete, as general equilibrium feedbacks are ignored. This will be shown in Section 13.5 when the simulation results are presented.

13.4.2 Economy-wide Impact of Emission Restriction

Emission restrictions in a subset of regions, say participating Annex 1 countries, would have the following impacts. In table 13.1, a schematic presentation is provided of the possible effects of emission restrictions on the rest of the economy. To begin with, emission restrictions would result in increases of the shadow prices of intermediates, in particular, the emission-generating commodities (the market price, in contrast, may fall due to reduced demand). In CGE models like GTAP-E, the increases in prices would have an impact for the demand for the emission-generating commodities as well as for other commodities in both the participating and the non-participating regions.

Table 13.1. Possible economy wide effects of emission restriction in GTAP-E.

	Price effect		Quantity effect	
	ecological output	econ. output	ecol. output	econ. output
Ecological inputs				
domestic	(i) ↑↑	(i) ↑↑	(ii) – – – –	(ii) – – –
import	(i) ↑↑	(i) ↑↑	(ii) – – – –	(ii) – – –
Economic inputs				
domestic			(iii) +	(iii) +
import			(iii) + +	(iii) + +
Sector aggregate	(iv) + + +	(iv) + + +	(v) – –	(v) – –
GDP	(vi) + + +		(vi) –	
Export to other regions				
ecological intermediates	(vii) + +	(vii) + +	(viii) – –	(viii) – –
economic intermediates	(vii) +	(vii) +	(viii) –	(viii) –
Trade balance	(ix) + +	(ix) + +	(x) + +	(x) –

Table 13.1 shows that a shadow price increase for emission-generating intermediates (i) would directly affect the demand for these intermediates in the production process. As a result of the Armington-structure of imported and domestic nest of intermediates, price increases in the higher nest would, depending on the substitution parameters, reduce the demand for domestic emission-generating commodities and for imported emission-generating commodities (ii). The substitution-effect between the emission-generating input and other input would, in the first instance, result in increases of demands for the other, non-emission generating, inputs (iii). However, this effect is not certain, because, the feedback-effects described by (iv) and (v) could decrease the demand for the economic input. In the end, the price of outputs (also the non-emission generating intermediates) would increase, as the price of the output is assumed to be weighted according to the prices of inputs, and the production depends on the use of the inputs. Of course, these effects would, relatively speaking, not be as large as for the demand of the input, as other substitution processes work in the economy as well. For example, an increase in the price of emission-generating intermediates would result in more demand for other intermediates.

Furthermore, emission restriction would also affect (iv) the price of output, (v) the demand, (vi) the aggregate regional output, i.e. GDP. Another expected general outcome is that, as a result of decreasing demand for emission-generating commodities in the participating countries, the outputs from these sectors will also be lower; while the outputs in the non-participating countries would be higher.

For the spatial-economic interaction, price increase in output would result in a loss of competitiveness (vii), therefore, a decrease in export will result (viii). However, from a volume-aspect of the net trade effect, a surplus may be the result of emission restriction. This 'counterintuitive' effect results from the decrease of demand for imports (ii). The results show that, given the parameters in GTAP-E, the import effect dominates. For the other, non-emission generating, outputs, the net trade effect (export (viii) - import (iii)) is negative for the participating regions. This result on spatial-economic interaction may, in practice, lead to a relocation of production, i.e. the producers in the emission-generating sectors may shift to the non-participating countries. Alternatively, production expansion in the non-participating countries may also occur via their own capacity-building mechanisms in the regions themselves.

13.4.3 CDM: Self-imposed Tax and Mitigation Effects of CERs

The CERs from CDM investments would mitigate the price-increasing effects of emission restriction targets that should be carried out domestically. The reason is that the amount of domestic reductions would be lower, which would result in a lower shadow price for emissions. This means that, as a result of CDM, the quantitative impacts of effects which are discussed in table 13.1 would be mitigated.

The self-imposed tax, however, of course has a price-increasing impact. This is presented in table 13.2. To begin with, the price of output commodities would increase in the Annex 1 countries. To some extent, the effects would be analogous to the price-increasing effects of carbon taxes or emission restrictions. The difference is that, in the case of carbon tax, it is the prices of emission-generating inputs that are largely affected, while the price of the output is only slightly affected as a result of substitution possibilities. In the case of output tax, the price and the demand for the inputs are all affected directly.

Table 13.2 A general sketch of impact of self-imposed tax for participating regions

	Price effect		Quantity effect	
	ecological output	econ. output	ecol. output	econ. output
GDP	↑		– –	
Sector aggregate	↑	↑	– –	– –
Export				
ecological intermediates	+ +	+ +	– –	– –
economic intermediates	+	+	– –	– –
Trade balance	+ +	+ +	–	–

The combination of self-imposed tax and the mitigation effects of CER would mean that the shadow prices of emission-generating inputs would be lower in comparison with solely the emission restriction. The prices of non-emission-generating inputs would, in comparison with emission restriction, be higher. This is due to the price-increasing effect of the self-imposed tax, which is absent for the non-emission generating commodities in case of emission-restriction. As a result, the leakage-effect would also be slightly be affected, as the prices of emission-generating intermediates are changed on the world market. However, the net aggregate impact of these effects is not certain as it depends on the elasticity and other parameters in the model.

13.4.4 Economic Impact of Technological Change

CDM results in technological change and would have direct impact for demand for inputs by the producers in the non-Annex 1 countries. Though we do not explicitly model the 'development additionality,' it will be a result of CDM investments. A general sketch of the impacts is presented in table 13.3. To begin with, technological change in the efficiency of ecological inputs in the non-Annex 1 countries lowers the relative price of these inputs. The price of outputs would also decrease.

Table 13.3. A general sketch of impact of technological change.

	Price effect		Quantity effect	
	ecological output	econ. output	ecol. output	econ. output
Technological change in non-Annex 1 countries				
Ecological inputs				
domestic intermediates			↓↓	↓↓
import intermediates			↓↓	↓↓
Economic inputs				
Sector aggregate	↓↓	↓↓	+ + +	+ +
Export				
ecological intermediates	– –	– –	+	+
economic intermediates	– –	– –	+	+
GDP	– – –		+ +	
Trade	– – – –	– – – –	+ + + + +	+ + + + +
Impact on other countries				
Ecological intermediates				
domestic intermediates				
import intermediates	– –	– –	+	+
Economic intermediates				
domestic intermediates				
import intermediates	– –	– –	+	+
Sector aggregate	–	–	+	+

In turn, these price-effects on input and output would affect the conditional demand for intermediates by the producers' from other sector. First, under a ceteris paribus condition, we would expect that technological change in sector ? , for example, would result in a decrease in sector ? 's demand for both domestic and imported intermediates per unit of output, as a more efficient use of the inputs requires less input for the same level of output. Secondly, the demand for products in ROW would be higher because of (i) feedback effects from lower output prices (resulting from better efficiency, and thus lower factor use); and (ii) substitution effects from policies in the Annex I regions, i.e. emission restriction policy and CDM tax. A higher price in the participating Annex I regions results in lower output in these countries, but higher output in the non-participating regions. These substitution effects also result in a difference between the domestic and imported energy intermediates, as both are imperfect substitutes.

For other regions, the precise impacts of technological change depend on, among other things, the elasticities and other (calibrated) parameters in the model. A lower price of imported intermediates as a result of technological change in the non-Annex 1 countries would, in the first instance, result in lower demand for domestic intermediates as a result of substitution. However, a lower price for imported intermediates would also result in a lower price for the output, while a lower price of output may, in turn, result in higher demand for domestic intermediates.

13.4.5 Emission Accounting

In GTAP-E, given the same amount of emission reduction, a comparison between two regimes would reveal the cost-effectiveness of some policy instruments. On the other, if the results for emission reductions from two possible policy regimes differ, the evaluation between the trade-offs between extra units of emission reductions and extra 'units of utilities' should be made explicitly. The reason is that the feedback from ecological system to the economic system is not modeled.

For this reason, this chapter will show the results for the ecological system in presenting the results of the impact of CDM on the rest of the economy. As a result of price impacts of emission reduction targets on the output in the participating Annex 1 countries (i.e. increased output price), spatial-economic interaction via international trade will affect the export, as other regions would import less commodities from the participating Annex 1 regions. For the non-participating Annex 1 countries, there would be, as a result of the production structure according to Armington-approach, (i) substitution between the imported intermediates from other regions, and (ii) substitution between domestic intermediates and aggregate imports.

Generally, this will result in (a) an increase in domestic demand; (b) an increase in imports from regions without emission restriction, and (c) a decrease in imports from regions with emission restriction. This increased demand for domestic intermediates in the non-participating regions is also called `leakage effect', meaning that a part of emission reduction target is cancelled out by increased emissions abroad, as a result of the substitution effects.

The net-effect of technological change would, as may be deduced from table 13.3, result in lower emissions as a result of decreased demand for ecological intermediates. However, in some settings, what is known as the rebound-effect would prevail. In these cases, the income effect as a result of technological change is larger than the lower demand as a result of increased efficiency.

The overall effect of CDM for the actual emission reduction on the world level is, however, not certain before hand. The reason is that this overall effect depends on many model parameters representing the responsiveness of the economic subjects to changes resulting from technological change.

13.5 Numerical Calibration

13.5.1 Introduction

This section presents a method for calibration of relevant parameters for modeling CDM extenions in a general spatial equilibrium setting. The resulting GTAP-CDM model will be used for a numerical analysis of a few CDM regime possibilities. The numerical analysis are needed as the general sketch of economic impacts shows that it is very difficult to pinpoint the overall economic effects, as these overall effects depend on spill-over and other feedback mechanisms in a CGE model.

13.5.2 Algorithm for Finding the Equilibrium in GTAP-CDM

This subsection introduces an iterative algorithm for finding the optimal rate of the self-imposed tax. The equilibrium conditions could not directly be implemented, as, in the benchmark equilibrium, there are no CDM activities. This is related to the fact that CDM activities are new instruments. An attempt to fully endogenize this instrument will result in an outcome of zero activity, as the latter is an equilibrium as well[2].

This iterative algorithm for finding the equilibrium condition for CDM activities is shown by figure 13.3. For this algorithm, we have that (i) technological change is endogenously captured in the GTAP-E model; but (ii) this technological change is conditioned on model variables that are exogenously determined, i.e. a self-imposed tax for generating investments; however, (iii) the optimal level of this exogenous variable, i.e. the optimal rate of the self-imposed tax, and the rate of technological change (which is derived from the conditional technological change function) can be determined by equilibrium conditions through a stepwise algorithm.

---

[2] It should be noted that, after obtaining the calibrated parameters using the semi-endogenous algorithm, it is possible to fully endogenize the CDM activities. This may be done in future research.

In this algorithm, we start with representative producers in Annex I countries that exogenously choose a self-imposed tax. This gives, through mechanisms of technological change and allocation by an international coordinator, an amount of CERs that may be attributed to representative producers. The algorithm stops when the CDM market is in equilibrium, i.e. when the monetary compensation of CERs equals the tax revenue for CDM activities, i.e.

$$V_j^{cer} = V_j^{\tau},$$

where V_j^{cer} is the monetary compensation of the CERs and is defined as the amount of CERs (E_j^{cdm}) times the price of the CERs (P_{cer}), i.e.: $V_j^{cer} = P_{cer} E_j^{cdm}$ is the tax revenue of CDM activities. The exogenous rate of tax would be increased if $V_j^{cer} > V_j^{\tau}$, and decreased if otherwise.

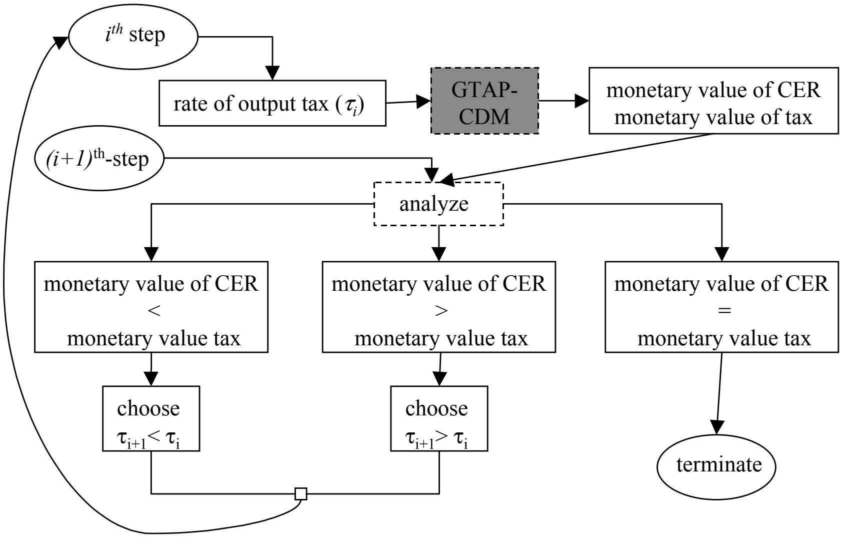


Fig. 13.3. Iterative algorithm for finding the optimal rate of self-imposed CDM tax.

13.5.3 The Conditional Technological Change Function

The functional form of the conditional technological change function. The parameter values to be calibrated are β_i, C_i and σ_i of the conditional technological change function ($A_i^n = \beta_i \left[C_i + V_i \right]^{\sigma_{ii}}$, for each ⟳. The subscript ⟳ stands for each emission-generating input (i.e. coal, gas and petroleum) in each sector (i.e. coal, gas, oil, petroleum, electricity, non-energy commodities) in the non-Annex I regions (i.e. ROW). In addition, a superscript �≪ denotes the benchmark equilibrium, a superscript �≪‿ the state of economy as if the relevant policy has been successfully implemented.

The intuition of the parameters of the conditional technological change function is as follows. First, we may perceive the term ∐⟳ △⟳, which is within the brackets of the conditional technological change function, as an aggregate measure for all capital and knowledge that help in improving the efficiency of the use of input ⟳. Then, given the general skills which are present in the economy, a given investment △⟳ would result in an efficiency increase by a certain amount. As △⟳ is expressed in monetary terms, it is assumed that the existing knowledge, reflected in ∐⟳ , is measured in units that make it comparable to money. Although within our GE framework, no actual investments in ∐⟳ are made. This parameter, β⟳ , may differ for each input ⟳ in order to reflect the assumption that a bigger sector would need a higher stock of capital investment to achieve the same relative increase in efficiency. For this reason, the β⟳-parameter will be assumed to be lower when the sector is bigger.

Conditions from the benchmark database. The benchmark equilibrium in the GTAP(-E) model imposes also conditions on the parameter values for the conditional technological change functions, as the initial value for technologies is normalised in the calibration of the benchmark equilibrium of GTAP-E. In particular, the normalisation involves the calibration of the production function $Q_i^n = A_i^n \left(\left[\left(\alpha_i^n K_i^n \right) \right]^{\phi_i} + \left[\left(1 - \alpha_i^n \right) Z_i^n \right]^{\phi_i} \right)^{1/\phi_i}$, where ∵ stands for output, … and ↗ for the inputs, ≙ for the level of technology, α for the share of each individual input in the production, and φ for the substitution parameter; subscript ⟳ stands for sector ⟳, and superscript �≪ for the ex-ante situation, and �≪‿ for the ex-post situation. For �≪ , the GTAP(-E) has normalised $A_i^0 = 1$ for all ⟳. GTAP-model needs to use this normalization in order to use the information on the physical inputs and outputs to calibrate the parameters α to represent the benchmark equilibrium.

This normalisation imposes a constraint on the calibration of the parameter values of the conditional technological change function. Given that $A_i^0 = 1$, and $V_i^0 = 0$, we have that the conditional technological change function should satisfy $1 = \beta_i \left[C_i \right]^{\sigma_i}$. In other words, the benchmark equilibrium requires that Ci be defined as:

$$C_i = \beta_i^{-1/\sigma_i}$$

for each ⟳. This means that we have two unknown parameters for the conditional technological change function for each input ⟳, as the third parameter will be determined by this constraint. The determination of these unknown parameters will be dicussed next.

Assumption about the nature of technology. There are no data available on the parameter values of the technological change functions to be used. Therefore, reasonable assumptions on the numerical values of these model parameters need to be made. The assumptions will partly reflect the assumed nature of technologies. In this study, we impose different values of the β_i and/ or C_i parameters and an equality of σ's to reflect the private character of technology. The equality of the σ's is simply to reflect a comparable `curvature' of the technological change functions.

In addition, we calibrate C_i, such that the size of the `existing knowledge' is proportional to the emissions emitted in the benchmark equilibrium, yielding:

$$\frac{C_i}{\sum_j C_j} = \frac{\chi_i Z_i^0}{\sum_j \chi_j Z_j^0}$$

This is an arbitrary choice, as empirical data on the `existing knowledge' is not yet available. However, this choice has a specific advantage in terms of computational convenience. It enables the use of a closed-form `distribution rule'. After assigning a value to one particular $\coprod_i$, the others thus follow, and given the fixed value of σ_i=0.5 for all i, all β_i's are determined via $C_i = \beta_i^{-1/\sigma_i}$. The fixed value of σ_i=0.5 is chosen to obtain a concave form (decreasing return to scale) for the conditional technological change function, which will result in a convex form of the implied marginal costs function of emission reductions due to CDM (i.e. increasing marginal costs).

13.5.4 Context for the CDM Standard Regime

The data. The data used in this chapter are the GTAP-4E data of 1995.[3] The level of aggregation involves three kinds of regions and three kinds of output commodities. The regional structure in this aggregation comprises: (i) source countries (participating Annex I regions, i.e. the EU and Other OECD-countries (OO)), (ii) host countries (Rest of World (ROW)) and (iii) non-participating Annex I regions (i.e. the USA and Economies in Transition, EIT). The structure for output commodities comprises: (i) emission-generating energy commodities (i.e. coal, gas and petroleum); (ii) non-emission-generating energy commodities (i.e. oil and electric energy); and (iii) non-energy commodities (i.e. other final goods).

In the simulations of the GTAP-CDM, the closure rules indicate the variables which should be kept exogenous. Each regime has a specific set of variables which should be kept exogenous, i.e., the closure rules differ between the regimes.

---

[3] This study does not use the GTAP-E data for the year 1997, as these data were only released recently (2002). Since this study only aims at offering a framework for analyzing the CDM, the 1995 data are sufficient to highlight the impacts of CDM.

Except for the variables specifically applicable to CDM policies, the closure rules follow the intra-regional emission trade regime (which follows Truong, 1999).

Emission restriction targets. The KP defines the commitments for Annex I Parties to reduce their overall greenhouse gas (GHG) emissions by an average of 5.2% below their 1990 levels in the commitment period 2008-2012. The commitments differ among the Annex I countries. In table 13.4, the assigned amounts are recalculated according to the regional aggregation used in this study for CO_2 emissions. It should be noted that GHGs also include other gasses such as methane. In addition, the 1995 levels of emissions are calculated according to the GTAP-E 1995 data set. According to these data, the emissions of the USA are 14.7% more than the assigned amount for the commitment period, while the EU emits 7.1% more, and the OO countries 12.4%. In contrast, the Economies in Transition (EIT) emit 27.4% less than their assigned amounts. This result is due to economic decline in these countries, related to transition problems after 1991.

In the literature, the determination of the emission reduction targets for the period 2008-2012 in static models is typically based on historical data (e.g. 1995), and involves the projection of data for the year 2010 (see, for example, IPCC 2000). The issue that then emerges is the choice among the many possible scenarios for the future, and consequently many different model results. As this study mainly aims at offering a framework for analyzing CDM policies, the simulations in this chapter will apply an emission reduction target of 5% for the year 1995 for the participating regions. The motivation for using the relatively low value of 5% is that countries would not be required to restrict their emissions at once, but can spread out the attainment of the target over many years. In this way, the shocks to the economic system would be softened, compared with those that would occur when just applying target to the economy in its base-year equilibrium. In other words, if the regions were already commited to implement the targets for the period 2008 -- 2012 in 1995, their reduction targets for that period would not necessarily be as high as the percentages shown in table 13.4. This is because the participating regions would be able to spread out the reductions over many years to minimize the shocks to the economic system. Applying a 5% reduction to 1995 data may reflect such a softened approach. We are aware that this percentage as such is quite arbitrary. However, the underlying economic processes affected by the CDM policies are not likely to be much different whatever percentage is chosen.

Table 13.4. Data on emissions and reduction targets for Annex I regions.

Regions	1990 emissions	Assigned amounts[a] (2008-12)		1995 emissions[b]	
	Mt CO_2	Mt CO_2	% of 1990 emissions	Mt CO_2	% of 1990 emissions
USA	4791	4456	93	5110	114.7
EU	3222	2964	92	3174	107.1
OO	1841	1772	96.3	1993	112.4
EIT	4155	4097	98.6	2972	72.6

Sources: [a] Aggregation based on Kuik, 2001; [b] Aggregation based on GTAP-E 1995 data.

Contribution of CDM Policies to the Emission Reduction Targets. To determine the particular $\coprod_{\circlearrowright}$, we use the empirical literature on the expected share of CDM activities in total emission reductions. Given the overall target of a 5% reduction, and given the choice of $\sigma_{\circlearrowright}$=0.5, different levels of the factors of $\coprod_{\circlearrowright}$ will imply different shares of CDM activities in total emission reductions achieved by Annex I regions, and by varying the `free' $\coprod_{\circlearrowright}$, any target (aggregate) share can be simulated.

The relevant literature shows that the estimates of the potential CDM projects vary from less than 30 megatonnes (Mt) of CO2$ to more than 1.6 gigatonnes of CO2 (Sijm *et al.*, 2000). However, the potential of the CDM projects is largely determined by the emission reduction targets as well as by the projection of the economic activities in the future. Another way is to use the size of the CDM market in terms of volumes of emission reductions. For this purpose, we refer to Zhang (2000), who cited a number of studies which estimated the size of the CDM market in 2010. The results from various studies are listed in table 13.5. From this table, the size of CDM market appears to differ significantly over studies.

Finally, as the absolute level of emissions in the equilibrium differs due to various model settings, e.g. emission reduction targets, we may use information on the share of the CDM in the total emission reduction, which in this study is determined as 25% of the emission reduction target by the participating regions. This share falls, according to table 13.5, in the low range of the available estimates. The models in table 13.5 are mainly based on projections for the year 2010 (e.g. the Kyoto targets and other assumptions on economic growth). We think a share of 25% for the CDM's contribution to achieving the emission reduction target is plausible, as the CDM standard regime is based on the 1995 data and has a relatively low emission restriction target.

Table 13.5. Estimates of the size of the CDM market involume.

	Size of the CDM market in (MtC)	contribution of CDM as a percentage of total emission reduction
EPPA	723	55
Haites	265-575	27-58
G-Cubed	495	45
GREEN	397	31
SGM	454	43
Vrolijk	67-141	10-21
Zhang	132-358	21-58

EPPA is based on Ellerman and Decaux (1998), Haites on Haites (1998), G-Cubed on McKibbin *et al.* (1999), GREEN on van der Mensbrugghe (1998), SGM on Edmonds *et al.* (1998), Vrolijk on Vrolijk (1999) and Zhang on Zhang (2000).

Source: Zhang, 2000

13.5.5 Calibration Results

Simulations with GTAP-CDM were carried out to find the parameter value for the `free' $C_{↻}$, which gives an equilibrium for the CDM standard regime. For the determination of the particular value of $C_{↻}$ consistent with this target, we may begin with a random value, for example, $C_{↻}$ =3.5. Then, simulations with GTAP-CDM will give the price of emissions. In combination with the amount of the CERs, the value of CDM activities can be determined. Given the value of CDM activities, we may determine the value of $C_{↻}$ such that the emission reduction target of 5% for the participating regions will be satisfied. This step involves many runs of simulations with the GTAP-CDM model to achieve the equilibrium value, i.e. (i) the emission reduction target is obtained; and (ii) the budget constraint is satisfied -- the CDM tax revenues equal the monetary value of the CERs.

Table 6 presents the parameter values for the CDM standard regime resulting from the calibration. The constants ($C_{↻}$) are proportional to the initial levels of emissions generated by the relevant inputs. In addition, given the predetermined value of $\sigma_{↻}$=0.5, the value of $\beta_{↻}$, is calculated. For the CDM standard regime, in which (i) the context is a reduction target of 5% for the participating regions and a CDM contribution of 25%; and (ii) the parameter setting is σ=0.5 and the assumption on the ratio between the parameter $C_{↻}$ for each ↻‖↺, we then have that ∐↻ ≪⌣ ⌣*χz↗↻ gives equilibrium for the CDM standard regime.

As the calibrated parameter values are affected by the model environment -- i.e. the reduction target, the share of CDM contribution in the total emission target of the participating regions, and the parameter value of σ -- this study also carried out two sensitivity analyses to examine the effects of these predetermined model settings. These two sensitivity analyses focus on the share of the CDM contribution and the parameter value of σ. The sensitivity analyses show that parameter choices do not affect the qualitative conclusions of the simulations, while the effects on quantitative results are small.

Table 13.6. Parameter values for the CDM standard regime.

			β?		
	oil	gas	petroleum	electricity	other
Coal	0.13	0.19		0.02	0.02
Gas	0.51	0.09	0.34	0.05	0.06
Petroleum		0.34	0.07	0.05	0.02
			C?		
Coal	62.20	26.67	0.00	2187.79	2031.15
Gas	3.81	113.92	8.43	365.43	319.36
Petroleum	0.00	8.47	218.61	446.84	2244.13

Model environment: reduction target 5%; CDM contributions: 25%; σ=0.5

Note: rows denote intermediate inputs; columns denote the sectors

13.6 Simulation Experiments

13.6.1 Introduction

Given the developments in the international negotiations on climate change policies, in particular the scenario of USA non-participation, it is clear that the global environmental regime is still fragile. For analyzing the policy impacts of global climate policies, this fragility of the climate regime leads us to consider the policy impacts of various possibilities for a future climate policy regime. This section constructs three variants to analyze the impacts of CDM policies. These are: (i) the CDM standard regime, (ii) the cap regime, and (iii) the USA participation regime.

13.6.2 Key Results

The CDM standard regime is based on the situation in which only the EU and the Other OECD countries (OO) participate as source countries of CDM activities. The USA are assumed not to have an emission target and not to participate in CDM activities.

Annex I regions' choice of the self-imposed CDM tax. Table 13.7 presents the market seize of emission reductions in the case of CDM standard regime, a cap on CDM and the intra-regional emission trade regime. In the intra-regional emission trade regime, the participating regions comply with the emission reduction targets solely by domestic emission reductions. Table 13.7 shows that in the intra-regional emission trade regime, the costs of emission reduction is US$ 842.3 million for the Annex I producers (US$ 471.5 million for the EU, and US$ 370.8 million for the OO countries).

As expected, CDM possibilities reduce the amount of emissions that should be achieved from intraregional emission trade. In the CDM standard regime, the amount of emissions achieved within the Annex I regions is 193.8 Mt CO_2, as compared with 258.4 Mt CO_2 in the intra-regional emission trade regime. The other 64.6 Mt CO_2 of the voluntarily-agreed target of emission reductions are achieved by CDM activities (that are calculated based on the assumed baseline method). In addition, CDM activities reduce the shadow price of the emissions because of the `where'-flexibility of CDM policy, i.e. to achieve the emission reduction targets through CDM investments in non-Annex I regions. The cap regime forms, as it partly includes CDM activities, a midway between the intra-region

Table 13.7. Market size of emission reductions.

	Intraregional emission trade			CDM Cap			CDM Standard		
	volume	price	Value	volume	price	value	volume	price	value
Market for emission permits									
Eu	158.70	2.97	471.52	150.66	2.79	420.73	129.81	2.38	308.75
Oo	99.67	3.72	370.79	75.41	2.79	210.59	63.97	2.38	152.15
subtotal	258.37		842.30	226.07		631.32	193.78		460.91
CDM market									
EU				8.04	2.32	18.69	28.90	2.38	68.74
OO				24.25	2.32	56.39	35.70	2.38	84.91
subtotal				32.29		75.07	64.59		153.64
Total	258.37		842.30	258.36		706.39	258.37		614.55

An interesting result is that the CDM activities, as modeled in this study, would result in an equalization of the shadow prices of emissions between the participating Annex I regions. This is an indirect form of efficiency which is comparable with interregional trade, because domestic marginal abatement costs will be equalized to the common $\|_{\Updownarrow\odot\rightarrow}$ in each of the participating Annex I regions.

Table 13.7 shows that the uniform price of US$ 2.38 per tonne of CO2 in the CDM standard regime, and US$ 2.79 in the cap regime, is lower than in the intra-regional emission trade regime, which is US$ 2.97 per tonne of CO2 for the EU and US$ 3.72 per tonne of CO2 for the OO countries. As a result, the CDM standard regime gives lower costs for the Annex I producers than the intra-regional emission trade regime. Table 13.7 shows that the costs for the Annex I producers are US$ 614.6 in the CDM standard regime and US$ 706.39 in the cap regime. These costs are lower than the US$ 842.3 in the intra-regional emission trade regime.

Thus, a first conclusion would be that the CDM standard regime is more cost-effective for the Annex I producers than the intra-regional emission trade regime. However, as we will discuss in the next section, leakage effects may alter this first conclusion.

The impact of CDM investment on non-Annex I regions. Table 13.8 shows the impact of CDM on demand for intermediates in ROW in the CDM standard and CDM cap regime. This results from two opposite effects. First, under a ceteris paribus condition, we would expect that technological change would result in lower demand for the inputs, as a more efficient use of the inputs requires less input for the same level of output. Secondly, the demand for products in ROW would be higher because of: (i) feedback effects from lower output prices (resulting from better efficiency, and thus lower factor use); and (ii) substitution effects from policies in the Annex I regions, i.e. emission restriction policy and CDM tax. A higher price in the participating Annex I regions results in lower output in these countries, but higher output in the non-participating regions. These substitution effects also result in a difference between the domestic and imported energy intermediates, as both are imperfect substitutes.

Table 13.8. Impact of CDM on ROW (non-Annex I regions).

	CDM standard regime					CDM cap regime				
	oil	gas	Petroleum	electricity	other	oil	gas	petroleum	electricity	Other
change domestic intermediates (in %)										
Coal	-0.95	-1.60	-0.49	0.11	0.09	-0.53	-1.25	-0.52	0.54	0.50
Oil	-0.35	-1.28	-0.45	0.45	0.42	-0.28	-1.05	-0.14	0.39	0.33
gas	-0.95	-1.60	-1.04	-0.61	-0.64	-0.45	-1.08	-0.37	-0.14	-0.19
petroleum	-0.33	-1.57	-1.02	-0.54	-0.57	-0.27	-1.07	-0.34	-0.07	-0.12
electricity	-0.35	-1.33	-0.49	-0.08	0.12	-0.28	-1.08	-0.17	-0.06	0.12
change imported intermediates (in %)										
Coal	-0.95	-1.60	-0.49	0.10	0.08	-0.09	-0.81	-0.08	0.98	0.94
Oil	-0.42	-1.35	-0.53	0.37	0.34	-0.34	-1.10	-0.19	0.34	0.28
gas	-0.95	-1.60	-1.04	-0.61	-0.64	-0.39	-1.02	-0.31	-0.08	-0.13
petroleum	-0.43	-1.67	-1.12	-0.64	-0.66	-0.36	-1.15	-0.43	-0.16	-0.21
electricity	-0.54	-1.52	-0.68	-0.27	-0.07	-0.49	-1.28	-0.37	-0.27	-0.08

Impacts on regional economies. As discussed in Section 13.4, the aggregate impact of CDM policies on the economic system and on emissions results from the three key aspects of the CDM as well as from the feedback effects in a CGE-model. These effects are confirmed by the simulation results. Table 13.9 shows the result of the CDM activities on GDP and sectoral disaggregation in the CDM standard and the CDM cap regime. Real GDP decreases for the participating Annex I regions, while it increases for other regions. A disaggregation to sectoral output shows that all the sectors in the participating Annex I regions would be faced with declining output. For the other countries, the basic energy sectors (i.e. coal, oil, gas and petroleum) are, except for petroleum in the USA, faced with lower production; while the other sectors (electric energy and `other', non-energy commodities) are gaining in production levels.

Table 13.9 Regional economic performance (in percentage change)

	CDM Standard					CDM cap				
	USA	EU	OO	EIT	ROW	USA	EU	OO	EIT	ROW
Price GDP	0.07	0.03	0.06	0.00	0.00	0.08	0.04	0.07	0.00	0.03
Real GDP	0.00	-0.03	-0.02	0.00	0.04	0.00	-0.04	-0.03	0.00	0.02
Sectoral desaggregation										
COL	-0.62	-1.26	-2.03	-0.61	-1.25	-0.67	-1.46	-2.26	-0.67	-1.26
OIL	-0.12	-0.16	-0.24	-0.20	-0.23	-0.10	-0.11	-0.22	-0.17	-0.19
GAS	-0.09	-0.74	-2.42	-0.91	-1.21	-0.08	-0.85	-2.79	-1.02	-0.98
P_C	0.19	-0.15	-0.81	-0.10	-0.51	0.15	-0.23	-1.00	-0.12	-0.18
ELY	0.07	-0.67	-0.01	0.49	0.06	0.09	-0.76	0.00	0.56	0.07
OTH	0.01	-0.03	-0.03	0.04	0.04	0.01	-0.04	-0.03	0.05	0.03
CGDS	0.26	-0.25	-0.03	0.33	0.22	0.28	-0.36	-0.06	0.35	0.25

Impacts on emissions. Finally, we look at the impacts on emissions. As a result of CDM policies and the related baseline calculations, there are four emission levels which are interesting to compare: (i) `actual' emissions in the case of no additional policies; (ii) ‘actual’ emissions under the emission restriction regime; (iii) ‘actual’ emissions under the CDM regime; and (iv) baseline emissions under the CDM regime. Table 13.10 shows emissions of (ii), (iii) and (iv) relative to the benchmark equilibrium (i).

Table 13.10 shows that the CDM regime is slightly more effective in reducing the world level of emissions than the emission restriction regime in Annex I regions alone. The real emission reduction in the world as a whole, which is 0.81% for the CDM standard regime, is lower than the `calculated' emission reduction (1.11%) according to the `baseline' method, but higher than in the case of the emission restriction regime. In the emission restriction regime, the reductions for the world as a whole is 0.80%. The calculated emission reduction under the CDM regime for the world as whole is larger than the actual emission reduction under the CDM regime, as the CERs attributed to the participating Annex I regions are not deducted from the actual emissions in the non-Annex I regions. Without CDM (i.e. in the intra-regional emission trade regime), emissions in the ROW would increase.

Table 13.10 Changes in emissions under various regimes

	USA	E_U	O_O	EIT	ROW	World
emission restriction	20.78	-158.70	-99.67	19.83	50.22	-167.53
	0.41%	*-5.00%*	*-5.00%*	*0.67%*	*0.64%*	*-0.79%*
CDM cap	19.45	-150.66	-75.41	18.35	20.19	-200.37
	0.38%	*-4.75%*	*-3.78%*	*0.62%*	0.26%	*-0.80*
Baseline cap	19.45			18.35	52.48	-32.29
	0.38%				*0.67%*	*-0.95*
cdm standard	18.92	-129.81	-63.97	16.46	-11.59	-169.94
	0.37%	*-3.98%*	*-3.13%*	*0.55%*	*-0.21%*	*-0.81%*
Baseline CDM standard	18.92			16.46	53.05	-234.53
	0.37 %			*0.55%*	*0.68%*	*-1.11%*

This result confirms the basic assumption for the introduction of CDM as an effective instrument for global climate change policies. In addition, this result corresponds to the models of induced technological change in which the positive effects of cleaner technologies are stronger. The reason is that, in our model, we have biased technological change towards more efficient technologies for emission generating inputs. However, the models with exogenous technological change -- in which technological change results in economic development and thus demand for emission-generating commodities -- give other results (for a review, see, e.g., Grubb 2000; Löschel 2002). In that case, the scale effect as a result of technological change is larger than the lower demand for emission-generating commodities as a result of increased efficiency.

13.6.3 CDM and US Participation

The likely scenario of USA non-participation shows the fragility of the KP in particular and climate change policies in general. The simulation in this section investigates whether the decision of USA non-participation is rational. As a criterion for rationality, this section focusses on the following two indicators: (i) total costs of emission reductions, and (ii) real GDP. The reason for using real GDP as an indicator is that the 'self-interested' state would only cooperate if it made a gain from the participation in a treaty. In this study, a gain is expressed by indicators describing regional economic performance, i.e. real GDP. It should be noted that this is an narrow indicator for 'rationality,' as it neglects the feedback effects of the ecological system on the economic activities, i.e. the valuation of the environmental aspects.

Impact on participating Annex I regions. In order to illustrate the impact of US participation, we will compare three regimes. These three regimes are: (i) intra-regional emission trade US_P, i.e. the USA, the EU and the OO countries carry out emission restriction policies through intra-regional emission trade; (ii) US_P regime, i.e. emission restriction policies and CDM activities by the USA, the EU and the OO countries; and (iii) the CDM standard regime, in which the USA does not participate.

Table 13.11 Market size of emission reductions in US_P

		intraregional			cdm US_P	
	volume	shadow price	costs	volume	shadow price	costs
costs of intra regional trade						
USA	255.53	2.50	638.65	247.30	2.38	589.04
EU	158.70	3.21	509.61	117.98	2.38	281.01
OO	99.67	4.00	398.91	56.40	2.38	134.34
subtotal	513.89		1547.17	421.68		1004.39
Costs of CDM						
USA				8.22	2.38	19.59
EU				40.73	2.38	97.01
OO				43.26	2.38	103.05
subtotal				92.22		219.65
Total	513.89		1547.17	513.89		1224.04

Table 13.11 presents the costs of emission reductions for the Annex I regions. As expected, USA participation involves more total costs of emission reductions (=US$ 1224.0 million) in comparison with the CDM standard regime (=US$ 614.6 million). In addition, CDM would lower the costs in comparison with the intra-regional emission trade regime with USA participation (total costs are US$ 1547.2 million). In the US_P regime, the volume of CERs for the US is low (8.2 Mt CO2). This indicates that, for the US, the CDM is not a very relevant policy alternative. This is reflected in the relatively equal price of emission reductions of US$ 2.38 per tonne CO2 between the US_P and the CDM standard regime (the price in the US_P regime is slightly higher than that in the CDM standard regime, but this difference is not visible due to rounding).

The lower costs in the CDM standard regime in comparison with the US_P regime are mainly due to the volume effect. The lower volume of emissions in the CDM standard regime is evident as only the EU and the OO countries are committed to emission reduction targets. The slightly lower price in the CDM standard regime (not visible in table 13.11 due to rounding) is due to the decreasing marginal productivity characteristics of the maginal costs function of emission reduction through CDM activities. The reason why the price of emission reductions through CDM activities is not much higher than US$ 2.38 per tonne of CO2 in the CDM US_P regime, is that, for the calculation of the marginal costs, not only the size of CDM investments but also the level of emissions in the new equilibrium is relevant. For a comparison, in the cap regime, halving the CDM investments results in a decrease from US$ 2.38 to US$ 2.32 per tonne CO2.

Impact on the non-Annex I regions. Table 13.12 presents the effects of the USA-participation regime for the non-Annex I regions. The results are shown relative to the CDM standard regime, i.e. the results from the US_P regime divided by the results from the CDM standard regime. We present the results to three decimal places in order to highlight the differences between the distribution of CDM investment and the CERs. In this table, we see that the impacts in the US_P region are, in general, larger than the CDM standard regime, as the amount of CDM investment is higher. However, because of the mentioned substitution effects, demands for some goods are different. Especially with relation to coal as intermediate commodity, we see substitution of the demand for the imported coal in the electricity and other, non-energy commodities for domestic coal as intermediates, as well as for the use of imported coal in other sectors.

Table 13.12. Impact on ROW (US_P relative to CDM standard).

	Oil	gas	petroleum	electricity	other
demand for domestic intermediates					
Coal	1.962	1.490		0.336	0.672
Oil			1.053	2.293	2.484
Gas	1.664	1.279	1.310	1.530	1.482
Petroleum		1.279	1.320	1.216	1.176
Electricity				2.442	1.275
demand for imported intermediates					
Coal	0.593	0.683		12.918	16.780
Oil			1.233	2.279	2.511
gas		1.306	1.338	1.601	1.550
petroleum		1.280	1.318	1.228	1.193
electricity				1.756	1.836

Table 13.13. Impacts on GDP.

	real GDP		price of GDP	
	USA_P	CDM standard	USA_P	CDM standard
USA	-0.01	0.00	0.09	0.07
E_U	0.02	-0.03	0.09	0.03
O_O	-0.01	-0.02	0.10	0.06
EIT	0.01	0.00	0.03	0.00
ROW	0.06	0.04	-0.01	0.00

Impacts on regional economies and their emissions. Table 13.13 shows the impact on real GDP as a result of the USA-participation regime and the CDM standard regime. In terms of regional economic performance, the USA would be worse-off in the US_P regime compared with the CDM standard regime. The real GDP decreases by 0.01% instead of an increase very slightly more than 0.00%. This result implies that, for the USA, participation in the CDM regime involves slightly more costs in terms of real GDP. In addition, all other regions benefit from USA participation.

Table 13.14 shows the disaggregated sectoral output in the US_P regime relative to the CDM standard regime. The results are diverse; in general, US_P enlarges the substitution effects, which are present in the CDM standard regime. So there is less demand for coal, oil, gas and petroleum because of the reduced emission reductions for the world as a whole due to US participation. In addition, it is interesting to note that USA participation results in shifts of competitiveness within the participation regions. This is the case for the petroleum sector and capital goods in the EU and electricity and capital goods in the OO countries. For these sectors, the negative effects in the CDM standard regime become positive in the US_P regime.

Table 13.14. Sectoral disaggregation.

	USA	EU	Other OECD	EIT	ROW
Relative to CDM standard					
Coal	6.00	1.36	1.46	1.72	1.55
Oil	6.10	3.28	2.60	2.33	2.54
Gas	58.96	1.09	1.05	1.00	1.19
Petroleum	-8.34	-0.75	0.72	0.45	1.20
Electricity	-2.79	1.06	-3.05	1.33	0.79
Other	-3.78	0.40	0.69	1.53	1.78
Capital goods	-0.83	-0.14	-2.02	1.43	1.30

Table 13.15. Impact on emissions.

	USA	E_U	O_O	EIT	ROW	World
CDM standard	0.37	-4.09	-3.21	0.55	-0.15	-0.81
US_P	-4.84	-3.72	-2.83	0.78	-0.15	-1.94

Table 13.15 presents the impacts of USA participation on emissions of the world as whole, as well as the contribution of each region. It can be seen that in the USA participation case, the world emissions would be reduced by -1.94%, while if only the EU and the OO countries participate, the world emissions would only be reduced by -0.81%. However, the difference in the emission reductions for the world as a whole is largely due to the emission reductions in the USA.

13.7 Conclusions

This chapter has discussed and motivated the assumptions that are required to operationalize CDM policies in a general equilibrium context. After the modeling of CDM extensions in a general spatial equilibrium model, this chapter further focused on the processes and variables within the economic system and the carbon emissions in the ecological system, which will be affected by the introduction of CDM policies.

First, we discussed the relevant part of the economic system in which major behavioral changes would take place due to the introduction of CDM policies. This concerns producer's behavior in both the participating Annex I and the non-Annex I regions. From a policy perspective, the following relevant assumptions were made. First, we assumed a multilateral institutional structure, by using a CDM coordinator for distributing the investments and CERs. This is analogous to the collection of savings by the global bank in GTAP-E model. Secondly, we did not explicitly take what are called 'additionality principles' into account (though implicitly, it is expected that technological change would increase the economic output and welfare). Thirdly, we restrict ourselves to the modeling of carbon emission targets, while other CDM-opportunities, such as other greenhouse gasses and reforestation are not included. Fourthly, the baseline emissions used for calculating the CERs awarded to the Annex I producers are calculated according to a pragmatic (but realistic) rule. Fifthly, no transaction costs are assumed for compliance and verification mechanisms.

From a modeling point of view, the following assumptions are made in addition to those in GTAP-E. First, it is assumed that Annex I countries finance CDM investments through self-imposed taxes on outputs. Secondly, technological change for the producers in the non-Annex I countries are assumed to be conditional on CDM investments from the Annex I countries. Thus, producers in non-Annex I countries cannot affect this conditional part of technological change. Thirdly, the transformation of the amount of investment into a change in technology can be described by a conditional technological change function. It is necessary to be aware that, in specifying the conditional technological change formulation, an assumption is made concerning the question: "How much investment is needed for a certain amount of technological progress in non-Annex I countries?" Fourthly, the CDM coordinator distributes these investments via an efficient rule.

As a final comment to the modeling aspect, this study is confined to a static analysis of one-directional ecological-economic interaction, as the aim is to analyze the impacts of climate change regimes on the regional distribution of wealth. For this reason, we implicitly assumed that the emission restrictions that are imposed on the economic system are a solution from an intertemporal optimization and thus are optimum. In this study, we take this restriction as given, and are unable to quantify the welfare effects of meeting this constraint.

Secondly, this chapter has presented the impacts of CDM policies by discussing the simulation results of a number of climate change policy regimes. This is done by presenting results concerning various parts of the economic system and carbon emissions, using the GTAP-CDM model. As spatial interaction in a CGE model takes place in terms of inputs and outputs, the results are presented in terms of inputs, sectoral outputs and regional outputs. In addition, except for the presentation of costs of emission reductions, these results are presented in quantitative measures. This shows that a synthesis between both the mainstream economic perspective and the ecologically-oriented economic perspective is possible. Finally, this study did not present a welfare evaluation of the climate change regimes, as, in terms of the spatially-extended E-E interaction framework, the GTAP-CDM model only takes one-way interaction between the E-E systems into account. This means that a welfare evaluation is not easy to give, because we do not have full information on the feedback of the ecological system on the economic system. In other words, a welfare evaluation may otherwise even be misleading due to the spatial-economic environmental externalities that are not taken into account by the economic subjects.

From the simulations discussed in this chapter, the general result is that CDM policies lower the total costs of emission reductions in comparison with policies focussed only on intra-regional emission reduction targets. A closer look at various parts (in terms of inputs, sectors and regions) of the economic system shows that the impacts are diversified. First, emission restrictions increase the relative price and decrease the demand of the emission-generating inputs in the participating regions. The self-imposed output taxes increase the prices and thus decrease the demand of the outputs. Secondly, the CERs will partly mitigate the price increases of the emission-generating inputs in comparison with the intra-regional emission reduction policy. Thirdly, technological change resulting from CDM activities decreases both the price and the demand for emission-generating inputs in the non-Annex I regions.

Each of these processes has different impacts on the various sectors and parts of the economic system, as well as on the emissions. In addition, spatial interaction takes place by three groups of countries. The interplay of each of these three processes would have different impacts for these groups of countries. Finally, the feedback effects, which are present in a general equilibrium model of the economic system, complicate the interpretation of the results.

In the CDM standard regime, it is found that the non-Annex I regions would, in terms of real GDP, gain most from CDM policies as a result of technological change. The participating regions would lose as a result of the emission reduction targets. The non-participating regions would slightly gain. In addition, the results for the cap regime show that it is midway between the CDM standard regime and the intra-regional emission trade regime. Thus, in the cap regime, the impacts on technological change and demand in ROW are lower than the impact on demand in the CDM standard regime. For all regions, the impacts of CDM policies on real GDP are lower in the cap regime in comparison with the CDM standard regime. US participation enlarges the impacts of (i) the relative prices of the emission-generating inputs; (ii) the self-imposed output taxes; and (iii) technological changes in the non-Annex I regions in comparison with the CDM standard regime. Moreover, this regime highlights the substitution within the participation regions. The negative impacts of CDM policies in the CDM standard regime for the EU and OO are lessened by US participation. The interplay of these three major processes associated with the CDM activities are shown more in detail by the disaggregated results on sectoral output.

In addition, it is important to see how the ecological system, represented as emissions only, will be affected. First, it is shown that the ‘baseline’ calculation for the CDM standard regime suggests a greater emission reduction than the ‘real’ reductions, compared with the case of the emissions restriction policy only. The reason is that the ‘baseline’ calculation is based on the old level of technology, with which more inputs would be needed to produce the `new' amount of output than with the new level of technology. Secondly, we may conclude that, though the participating Annex I regions would reduce their emissions less domestically, the level of emissions for the world as a whole would, in the case of the CDM regime, be less than in the case of the intraregional trade regime. The reason is that, in the non-Annex I regions, the leakage effects from emission reductions in the participating Annex I regions are offset by the technology effects from the CDM investments. This indicates that CDM policies are indeed a promising complementary alternative for emission restriction policy.

An interesting feature of CDM policies, at least in our model setting, is that they indirectly equalize the price of emissions in the participating Annex I regions. This follows from the assumption that the CDM activities are coordinated by the CDM coordinator, so that the trade-off between emission reductions through CDM policies and emission reductions through domestic intra-regional trade becomes equal for the representative producers in the participating Annex I regions.

By incorporating conditional technological change in a computable general equilibrium model, this study gives more detailed insights into the spatial economic impacts of a new instrument of climate change policies which involves technological change, i.e. the CDM policies, on economic activities. Indeed, the results confirm the intuitive notion that technological change will have a positive impact on economic performance in comparison with emission restriction policy only. In the first place, this holds for the non-Annex I regions. In the second place, via spatial-economic interaction, the positive impact spills over to other countries. However, as shown, whether the results will be positive or negative for a particular region largely depends on the regime setup and interplay of various processes within the global economy. Finally, it should be noted that the numerical results are sensitive to the model environment which is chosen for the CDM standard regime.

References

Aghion,P, Howitt P (1998) Endogenous growth theory. MIT Press, Cambridge

Begg KG, Jackson T, Parkinson S (2001) Beyond joint implementation-designing flexibility into global climate policy. Energy Policy 29:17-27

Bollen J, Gielen A, Timmer H (1999a) Clubs, ceilings and CDM. Macroeconomics of compliance with the Kyoto Protocol. The Energy Journal 20:177-206.

Bollen J, Manders T, Timmer H (1999b) Kyoto and carbon leakage. simulations with WorldScan. In: Pan JH, van Leeuwen N, Timmer H, Swart R (1999b), Proceedings of IPCC expert meeting on economic impact of mitigation measures. CPB, The Hague, pp. 93-116.

Bosi M, Ellis J (1999) Implications of multi-project emissions baselines for CDM projects -- examples from the electricity generation sectors in Brazil and India. In: Baron R, Bosi M, Lanza A (eds) Emission trading and the clean development mechanism: resource transfers, project costs and investment incentives. IEA

Chen W (2003) Carbon quota price and CDM potentials after Marrakesh. Energy Policy 31:709-719

Duic N, Alves LM, Chen F, da Gracca Carvalho M (2003) Potential of Kyoto protocol clean development mechanism in transfer clean energy technologies to small island developing states. Case study of Cape Verde. Renewable & Sustainable Energy Reviews 7:83-98

Fichtner W, Goebelt M, Rentz O (2001) The efficiency of international cooperation in mitigating climate change. Analysis of joint implementation, the clean development mechanism and emission trading for the Federal Republic of Germany, the Russian Federation and Indonesia. Energy Policy 29:817-830

Forner C, Jotzo F (2002) Future restrictions for sinks in the CDM. How about a cap on supply? Climate Policy 2:353-365

Geres R, Michaelowa A (2002) A qualitative method to consider leakage effects from CDM and JI projects. Energy Policy 30: 461-463

Gerlagh R, Dellink R, Hofkes M, Verbruggen H (2000) An applied general equilibrium model to calculate a Sustainable National Income for the Netherlands. In: Verbruggen H (ed) Final report on calculations of a sustainable national income according to Hueting's methodology. Institute for Environmental Studies, Amsterdam, pp 97-124

Ginsburgh, V., M. Keyzer (1997). *The Structure of Applied General Equilibrium Models.* Cambridge, MIT Press.

Gruebler A (1998) Technology and global change. Cambridge University Press, Cambridge

Halsnaes K (2002) Market potential for Kyoto mechanisms—estimation of global market potential for co-operative greenhouse gas emission reduction policies. Energy Policy 30:13-32

Harberger AC (1962) The incidence of the corporation income tax. Journal of Political Economy 70:215-240

Hardner JJ, Frumhoff PC, Goetze DC (1999) Prospects for mitigating carbon, conserving biodiversity, and promoting socioeconomic development objectives through the clean development mechanism. Mitigation and Adaptation Strategies for Global Change 5:61-80

Hertel TW, Tsigas ME (1997) Structure of GTAP. In: Hertel TW (1997) Global Trade Analysis: Modeling and Applications. Cambridge University Press, Cambridge, pp 9-71

Hertel TW (1997) Introduction and Overview. In: Hertel TW (ed) Global trade analysis: Modeling and applications. Cambridge University Press, Cambridge, pp 3-12

IISD (2002) Earth Negotiation Bulletin 12:209 http.//www.iisd.ca/linkages/vol12/enb12209e.html

IPCC (2001) Climate change: the IPCC scientific assessment, intergovernmental panel on climate change. Cambridge University Press, Cambridge

Jackson T (1995) Joint implementation and cost-effectiveness under the framework convention on climate change. Energy Policy 19:35-47

Jaffe AB, Peterson StR, Portney PR, Stavins R (1995) Environmental regulation and the competitiveness of US manufacturing. Journal of Economic Literature 33:557-572

Johansen L (1960) A multi-sector study of economic growth. North-Holland Publishing Company, Amsterdam

Jotzo F, Michaelowa A (2002) Estimating the CDM market under the Marakech Accords. Climate Policy 2:179-196

Kremers H, Wang S, Nijkamp P (2002) Modelling issues on climate change policies—A discussion of the GTAP-E model. Journal of Environmental Systems 28:217-241.

de Leeuw R, van Ierland EC (2003) CDM in climate policies in the Netherlands. A promising tool? In: Faure M, Gupta J, Nentjes A (eds) Climate change and the Kyoto protocol: The role of institutions and instruments to control global change. Edward Elgar, Cheltenham, pp 71-94

Loeschel A (2002) Technological change in economic models of environmental policy: A survey. Ecological Economics 43:105-126

Lucas Jr. RE (1988) On the mechanics of economic development. Journal of Monetary Economics 22:3-42

Michaelowa A (2003) The Kyoto protocol and its flexibility mechanisms. International Society for Ecological Economics: Internet Encyclopaedia of Ecological Economics, http//www.ecologicaleconomics.org/publica/encyc_entries/Kyotoandflex.pdf.

Noble I, Scholes RJ (2001) Sinks and the Kyoto protocol. Climate Policy 1:5-25

Nordhaus WD, Yang Z (1996) A regional dynamic general-equilibrium model of alternative climate change strategies. American Economic Review 86:741-765

Parkinson S, Begg K, Bailey P, Jackson T (1999) JI/CDM crediting under the Kyoto Protocol: Does ‘interim period banking' help or hinder GHG emissions reduction? Energy Policy 27:126-136

Romer PM (1986) Increasing returns and long-run growth. Journal of Political Economy 94:1002-1037

Scarf H (1967) On the Computation of Equilibrium Prices. In: Fellner W *et al.* (eds) Ten Economic Studies in the Tradition of Irving Fisher. John Wiley, New York, pp 207-30

Schwarze R, Niles JO (2000) The long-term requirement for clean development mechanism: Forestry and economic liability. Journal of Environment and Development 9:384-404

Shoven JB, Whalley J (1984) Applied general-equilibrium models of taxation and international trade: An introduction and survey. Journal of Economic Literature 22: 1007-1051

Shoven JB, Whalley J (1992) Applying General Equilibrium. Cambridge University Press, Cambridge

Sijm JPM, Ormel F, Martens JW, van Rooijen SNM, Voogt MH, van Wees MT, de Zoeten-Dartenset C (2000) Kyoto mechanisms: The role of joint implementation, the clean development mechanism and emissions trading in reducing greenhouse gas emissions. Netherlands Energy Research Foundation ECN, Petten.

Solow RM (1957) Technical change and the aggregate production function. The Review of Economics and Statistics 39:312-320

Toman M, Hourcade J (2000) From Bonn to The Hague: Many questions remain. Resources 138:14-16

Truong TP (1999) GTAP-E. Incorporating Energy Substitution into GTAP Model. Center for Global Trade Analysis, Purdue University, West Lafayette, IN.

Wang S, Nijkamp P, Verhoef ET (2001) Modelling externalities between ecological and economic systems, Journal of Environmental Law and Policy 26:395-415

Wang S, Nijkamp P, Kuik O (2003) Global environmental change regimes: Impact assessment on the basis of an extended GTAP model. In: Faure M, Gupta J, Nentjes A (eds) Climate change and the Kyoto protocol: The role of institutions and instruments to control global change. Edward Elgar, Cheltenham, pp 248-276

Weyant JP, Olavson T (1999) Issues in modeling induced technological change in energy, environmental, and climate policy. Environmental Modeling and Assessment 4:67-85

Weyant JP (1999) Economics of the Kyoto Protocol. Energy Journal 20:25-72

Woerdman E (2000) Implementing the Kyoto protocol: Why JI and CDM show more promise than international emissions trading. Energy Policy 28:29-38

Zhang Z-X (2000) The design and implementation of an international trading scheme for greenhouse gas emissions. Environment and Planning C: Government and Policy 18:321-37

Zhang Z-X (2001) An assessment of the EU proposal for ceilings on the use of Kyoto flexibility mechanisms. Ecological Economics 37:53-69

14 An Environmental Socioeconomic Framework Model for Adapting to Climate Change in China*

Bin Li[1] and Yoshiro Higano[2]

[1] Department of Biology, University of Jiaxing, China
[2] Institute of Agricultural and Forest Engineering, University of Tsukuba, Japan

14.1 Introduction

Global warming will be a significant common issue facing society in the 21st century. At the Third Conference of Parties to the United Nations Framework Convention on Climate Change (COP3) held in Kyoto, it was agreed by 161 participating nations to do their best in the period from 2008 to 2012 to curtail global warming gases as per the Kyoto Protocol to the United Nations Framework Convention on Climate Change (UNFCCC). The Kyoto Protocol came into force on 16 February 2005.

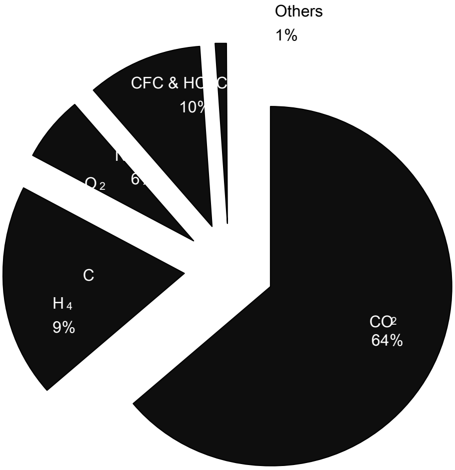
China ratified the Kyoto Protocol on September, 2003. China is the second largest energy producer and the second largest energy consumer in the world just ranked after the US. China's energy consumption pattern largely depends on its energy resources, distribution and transmission. Generally, carbon dioxide (CO_2) makes the largest contribution to the global warming among greenhouse gases (GHGs; figure 14.1). This gas is released into the environment mostly by combustion of fossil fuel for energy use. Among the fossil fuels, coal is the fuel that discharges the most $CO_2$when burned. Because the carbon content per calorific value of coal is on average 30% higher than that of oil, coal is a major cause for the increase in the amount of CO_2 discharged, especially within China where the energy dependence on coal was 66.1% in 2000 (China National Bureau of Statistic (NBS), 2003). Therefore, from the viewpoint of taking measures against global warming, it is desirable not only to aim at achieving efficient energy consumption but also to shift from coal to other energy sources, especially those that discharge

---

* We would like to thank Professor Raymond J.G.M. Florax, Professor Jacques Poot; Dr. Pawel Bartoszczuk and Professor A. Stewart Fotheringham for their valuable comments.

smaller amounts of CO_2 than coal, such as nuclear power or other forms of alternative energy.

In this chapter, the focus is on a nation-wide carbon tax and the advancement of alternative energy sources. This chapter consists of seven sections. We introduce the purpose briefly in section 14.1, while the literature review follows in section 14.2. The model, simulation results and interpretations are described in sections 14.3, 14.4 and 14.5, while some considerations for future issues is provided in section 14.6.



Source: Intergovernmental Panel for Climate Change (IPCC), 1996a.

Fig 14.1 Contribution to global warming of each greenhouse gas from the Industrial Revolution to 1992.

14.2 Literature Review: China-Specific Environmental Models

We reviewed China-specific environmental models focusing on their use of data related to pollutants and natural resources. For example, the World Bank model (Martin 1993; Research Institute for Development and Finance of Japan Bank for

International Cooperation, 2000) is a dynamic general equilibrium model (GEM) of the post-reform Chinese economy, centering on the extent to which prices affect the allocation of resources in China's 'planned commodity economy.' The initial source of data in the model was the World Bank input-output (I-O) table for 1981, the latest available when the model was constructed.

The model of Zhang (Zhang 1997) using the 1987 I-O table is another GEM which 'pays particular attention to modeling the energy sector and linkages to the rest of the economy.' Carbon emissions are calculated from energy consumption by applying energy-specific emission coefficients. In this sense, energy consumption and CO_2 emissions are decoupled from economic growth in two ways.

In an optimum transportation model (Li and Higano 2001), China was divided into seven zones and multi-regional, multi-sectional socio-economic model was built. Based on social environment and economy of China in 1997, the changes in economic growth are analyzed into the 21st century under harsher measures for the amount of discharge of contaminants of CO, SO_x and NO_x. In particular, as an experimental study, the change in the regional economy of the western part of China was analyzed. This has become a burning issue recently due to the expansion of the transportation system as a result of significant infrastructure investment. The model was calibrated on the 1997 I-O table.

3Es-Model (Li 2003) is an integrated econometric model consisting of macroeconomic sub-model, energy sub-model and environment sub-model to perform a long-term simulation study for China without using an I-O table. The environment sub-model is simply designed to generate the energy-related productions matrices and emissions matrices of both SO_2 and CO_2, following the energy balance table. Case-5, the carbon tax case, assumes that a carbon tax of 10US$/t-c will be imposed on energy consumption from 2011. This means that the rate of carbon tax is an exogenous variable.

In the next section, the model developed for the analysis will be presented.

14.3 Eco-conscious Socioeconomic Framework Model

The whole model can be divided into two parts. The first part estimates the cost of alternative energy based on the theory of learning by doing. As is well known, the learning curve based on this theory has the characteristic that the production cost of a new product with advanced technology falls exponentially. It is assumed that the cost of alternative energy is determined by using the information on the rate of increase in the production of output, the feature of the learning curve, and the price of competing energy. The production cost of the alternative energy was estimated based on the logistic curve and penetration ratio. The second part of the model is specified mainly based on the idea of an environmental value-added tax (Higano 1996). The whole model is a GEM and markets are mutually linked through the duality of commodity and value balance, and are in equilibrium.

14.3.1 Modeling of Cost of Alternative Energy

The cost of alternative energy reflects the price of output of an alternative energy sector, under the assumption of exponentially declining production costs and a market penetration ratio that characterizes the competitiveness of a new product.

The Empirical Equation of Cost of Wind Power. Based on numerous studies published elsewhere, the following assumptions were made. It was assumed that a wind-power enterprise could use imported windmills with given purchase prices for the machine, and given cost of construction and maintenance expenses. It was then possible to estimate the unit price of power generation, depending on the location, with average wind velocity of five to six meters per second, using the following formula.

$$220 \cdot \left(\frac{(1+a)^{t-1986} - 1}{a} \right)^{-b} \cdot \frac{WPI_t}{WPI_{1986}} \quad \text{(Unit: Japanese yen (JPY)/ kilowatt-hour (kWh))}, \tag{14.1}$$

where,

a rate of increase in production a year;

b the rate of reduction of the cost accompanied by accumulated production;

WPI_t wholesale price index in year *t*.

Moreover, the future cost can be set by referring to the spread of wind power in the West. The cost of wind power generation as of 1986 in Japan was 220 JPY/kWh.

The Empirical Equation of Cost of Solar Power. The price of solar power systems differs according to their scale. Supposing that for the household sector, a small solar cell of 3-4 kW is used and for the industrial sector, a larger-sized one exceeding 10 kW is used. Hence, the power costs according to the expense of installation and maintenance in various areas are estimated. Annual utilization is assumed to be 1460 hours, i.e. one day 4 hours × 365 days.

$$242 \cdot \left(1 + 0.3565 \cdot \frac{(1+a)^{t-1986} - 1}{a} \right)^{-b} \cdot \frac{WPI_t}{WPI_{1986}} \quad \text{(Unit: JPY/kWh)}. \tag{14.2}$$

The value of b is calculated from the estimates of power cost in the past. The cost of solar power generation as of 1986 in Japan was 242 JPY/kWh. Here, the interpretation of α, b and WPI_t is the same as for equation (14.1).

The power cost in the model is divided into two types; one is a trend type and the other is a breakthrough type. The trend type relates to traditional production systems while the breakthrough type is based on anticipated progress in technical developments in the future. It was felt that the former would be more appropriate for application in China.

The cost of the alternative energy in 1997 was calculated using the parameter of the trend type (table 14.1). The currency rate in 1997 is calculated as: 1 CNY = 15.68 JPY (NBS, 1999b). The cost data is used as P_{new} in the equation (7) of the model.

Table 14.1. Parameters in cost model of trend type and cost of alternative energy in China.

Alternative energy	α	b	Cost (CNY/kWh)
Wind power	0.4	0.3219	2.5
Solar power	0.2	0.3546	0.75

14.3.2 Industrial Classification

Using a classification of 32 sectors, the sector of Electric Power, Gas and Heat Supply is sub-divided into Electric Power, and Gas and Heat Supply. Adding two Alternative Energy sectors, Solar Power and Wind Power and one sector of Pollution Abatement, which removes CO_2 with cleaner technology such as methane fermentation, a total of 36 sectors were specified for the Chinese economy (see Table 14.2).

Table 14.2. Sectoral classification.

Index	Sector	Index	Sector
1	Agriculture	19	Water
2	Mining	20	Whole sale and retail service
3	Food products	21	Finance and insurance
4	Textiles and clothing	22	Real estate services
5	Paper and wooden products	23	Transport
6	Chemical products	24	Communications
7	Coal and petroleum products	25	Public works
8	Non-metallic mineral products	26	Education and research services
9	Iron and steel	27	Medical, health and social insurance services
10	Non-ferrous metals	28	Other public services
11	Fabricated metal products	29	Miscellaneous business services
12	General machinery	30	Miscellaneous personal services
13	Electric machinery	31	Office supplies
14	Transport equipment, automobiles	32	Not elsewhere classified
15	Precision apparatus	33	Electricity
16	Other manufactures	34	Solar power
17	Building and construction	35	Wind power
18	Gas	36	Pollution abatement

The input coefficients of Solar Power sector and Wind Power sector were estimated using the I-O structure of the Household Electric Appliances sector and Generators sector, respectively. In addition, considering the social cost of CO_2 that has been discussed extensively, the Pollution Abatement sector was estimated using the I-O structure of the General Machinery sector. The cost of Abatement sector was estimated at 100 CNY/t in the research. The abatement industry is a virtual one.

14.3.3 The Assumptions of the Model

An assumption of the model is that there is a surplus of labor to substitute for capital; that tradable emission quotas are allocated by the government or another organization commissioned by the government under an emission trading system in a quasi-market; that there exist 36 commodity markets, 36 capital markets and 33 consumption markets, and that they are in equilibrium as of 1997.

14.3.4 Framework of the Multi-Sectoral Model Including I-O Table

The model consists of the following 13 groups of equations. Each will be described briefly.

Commodity Flow Balance in the Usual Sectors. Each usual sector would produce enough to meet all demand. Namely, the product of an industry is equal to or larger than the total demand for the product. The total demand consists of the sum of intermediate input, final consumption, fixed capital formation and the export-import balance.

$$\begin{aligned}&A_{11}X_1(t)+A_{12}X_2(t)+A_{13}X_3(t)+A_{14}X_4(t)+C_1(t)+G_1(t)+B^p\Delta K^p(t)+\\&+B^g\Delta K^g(t)+E(t)-M(t)\le X_1(t)\end{aligned}, \qquad (14.3)$$

$$X(t)=\begin{bmatrix}X_1(t)\\X_2(t)\\X_3(t)\\X_4(t)\end{bmatrix},\quad K(t)=\begin{bmatrix}K_1(t)\\K_2(t)\\K_3(t)\\K_4(t)\end{bmatrix},\quad \Delta K^p(t)=\begin{bmatrix}\Delta K_1(t)\\\Delta K_2(t)\\\Delta K_3(t)\\\Delta K_4(t)\end{bmatrix},$$

$$C(t)=\begin{bmatrix}C_1(t)\\C_2(t)\end{bmatrix},\quad G(t)=\begin{bmatrix}G_1(t)\\G_2(t)\end{bmatrix},\quad B^p=\begin{bmatrix}B_{11} & B_{12} & B_{13} & B_{14}\end{bmatrix},$$

and where,

$X(t)$ a column vector of the total production of all the sectors in term t;

$X_1(t)$ a column vector of the total production of the usual sectors in term t;

$X_2(t)$ the total production of the energy sector in term t;

$X_3(t)$	a column vector of the total production of the alternative energy sectors in term *t*;
$X_4(t)$	the activity level of the CO_2 abatement sector in term *t*;
A_{11}	an input coefficient matrix between the usual sectors (exogenous);
A_{12}	an input coefficient matrix from the usual sectors to the energy sector (exogenous);
A_{13}	an input coefficient matrix from the usual sectors to the alternative energy sectors (exogenous);
A_{14}	an input coefficient matrix from the usual sectors to the CO_2 abatement sector (exogenous);
$C(t)$	a column vector of the consumption in term *t*;
$C_1(t)$	a column vector of the consumption of the usual sectors in term *t*;
$C_2(t)$	the consumption of the energy sector in term *t*;
$G(t)$	a column vector of the government expenditure in term *t*;
$G_1(t)$	a column vector of the government expenditure on the usual sectors in term *t*;
$G_2(t)$	the government expenditure on the energy sector in term *t*;
B^p	a column vector of the fixed capital formation rate (exogenous);
B_{11}	a column vector of the fixed capital formation rate of the usual sectors (exogenous);
B_{12}	a column vector of the fixed capital formation rate of the energy sector (exogenous);
B_{13}	a column vector of the fixed capital formation rate of the alternative energy sectors (exogenous);
B_{14}	a column vector of the fixed capital formation rate of the CO_2 abatement sector (exogenous);
$\Delta K^p(t)$	a column vector of the private fixed capital formation in term *t* (exogenous);
$\Delta K_1(t)$	a column vector of the private fixed capital formation of the usual sectors in term *t* (exogenous);
$\Delta K_2(t)$	the private fixed capital formation of the energy sector in term *t* (exogenous);
$\Delta K_3(t)$	a column vector of the private fixed capital formation of the alternative energy sectors in term *t* (exogenous);
$\Delta K_4(t)$	the private fixed capital formation of the CO_2 abatement sector in term *t* (exogenous);
B^g	a public fixed capital formation rate (exogenous);
$\Delta K^g(t)$	the public fixed capital formation in term *t*;

$E(t)$ a column vector of export in term *t*;

$M(t)$ a column vector of import in term *t*.

Commodity Flow Balance in the CO$_2$ Pollution Abatement Sector. Total emissions consist of emissions from the usual sectors, the energy sector, the alternative energy sectors, the CO$_2$ pollution abatement sector and the final demand sector. Further, the net emissions are calculated by subtracting the amount of CO$_2$ pollutants abated by the CO$_2$ pollution abatement sector from total emissions.

$$Z(t) \equiv B_{41}X_1(t) + B_{42}X_2(t) + B_{43}X_3(t) + B_{44}X_4(t) + B_{45}[lC(t) + lG(t)] - X_4(t) , \tag{14.4}$$

in which,

$Z(t)$ a column vector of net emissions of CO$_2$ in term *t*;

B_{41} an emission coefficient matrix of the usual sectors (exogenous);

B_{42} an emission coefficient of the energy sector (exogenous);

B_{43} an emission coefficient matrix of the alternative energy sectors (exogenous);

B_{44} an emission coefficient of the CO$_2$ pollution abatement sector (exogenous);

B_{45} an emission coefficient of consumption and government expenditure (exogenous);

l an aggregation row vector whose elements are all one.

Value Flow Balance on each Sector. Value flow balance on the usual sectors is specified as the following equation, which implies the prices of commodities of usual sectors are decided according to the factors of intermediate cost, net value added and the carbon tax rate:

$$P_1(t)\tilde{X}_1(t) = P_1(t)A_{11}\tilde{X}_1(t) + \psi_1\tilde{X}_1(t) + \delta_1\tilde{K}_1(t) + \tau_1^i\tilde{X}_1(t) + T(t)B_{45}\tilde{C}_1(t) + T(t)B_{41}\tilde{X}_1(t) . \tag{14.5}$$

Value flow balance on the energy sector is specified as the following equation:

$$\begin{aligned} P_2(t)X_2(t) = {} & P_1(t)A_{12}X_2(t) + \psi_2X_2(t) + \delta_2K_2(t) \\ & + \tau_2^iX_2(t) + T(t)B_{45}C_2(t) + T(t)B_{42}X_2(t) \end{aligned} \tag{14.6}$$

Value flow balance on the alternative energy sectors is specified in the following equation:

$$P_3(t)(1 - \tau_3^i(t))\tilde{X}_3(t) = P_1(t)A_{13}\tilde{X}_3(t) + \psi_3\tilde{X}_3(t) + \delta_3\tilde{K}_3(t) + T(t)B_{43}\tilde{X}_3(t) . \tag{14.7}$$

Value flow balance on the CO$_2$ pollution abatement sector is specified in the following equation:

$$T(t)X_4(t) \le P_1(t)A_{14}X_4(t) + \psi_4X_4(t) + \delta_4K_4(t) + \tau_4^iX_4(t) + T(t)B_{44}X_4(t) . \tag{14.8}$$

Following the estimated empirical equation of cost of alternative energy, $P_3\left(t\right)$ is bound by the following equation:

$$P_3(t) = \frac{P_{new}(t)}{\theta l P_2(t)}, \tag{14.9}$$

in which,

$\tilde{X}_1(t)$	a diagonalized matrix of $X_1(t)$;
$\tilde{X}_3(t)$	a diagonalized matrix of $X_3(t)$;
$P_1(t)$	a row vector of price of the usual sectors in term *t*;
$P_2(t)$	a row vector of supply price of the conventional energy sector in term *t*;
$P_3(t)$	a row vector of supply price of the alternative energy sectors in term *t*;
$T(t)$	a carbon tax rate in term *t*;
ψ_1	a row vector of rate of gross income in the usual sectors;
ψ_2	a rate of gross income in the energy sector;
ψ_3	a row vector of rate of gross income in the alternative energy sectors;
ψ_4	a rate of gross income in the CO_2 pollution abatement sector;
τ_1^i	a row vector of indirect tax and subsidy rate of the usual sectors (exogenous);
τ_2^i	an indirect tax and subsidy rate of the energy sector (exogenous);
$\tau_3^i(t)$	a subsidy rate for the alternative energy sectors in term *t*;
τ_4^i	an indirect tax and subsidy rate of the CO_2 pollution abatement sector (exogenous);
δ_1	a row vector of rate of depreciation of the usual sectors (exogenous);
δ_2	a rate of depreciation of the energy sector (exogenous);
δ_3	a row vector of rate of depreciation of the alternative energy sectors (exogenous);
δ_4	a rate of depreciation of the CO_2 pollution abatement sector (exogenous);
$K(t)$	a column vector of the private capital stock of all the sectors in term *t* (exogenous);
$K_1(t)$	a column vector of the private capital stock of the usual sectors in term *t* (exogenous);
$K_2(t)$	the private capital stock of the energy sector in term *t* (exogenous);
$K_3(t)$	a column vector of the private capital stock of the alternative energy sectors in term *t* (exogenous);
$K_4(t)$	the private capital stock of the CO_2 pollution abatement sector in term *t* (exogenous);

$P_{new}(t)$ a row vector of the experimental cost trend of the alternative energy sectors in term t (exogenously given in Section 3.1);

θ a factor which adjusts difference of energy price indices (exogenous).

Energy Supply-and-demand Restrictions. These restrictions are as follows:

$$l\gamma_1 X_1(t) + l\gamma_2 X_2(t) + l\gamma_3 X_3(t) + l\gamma_4 X_4(t) + \gamma_5 \{lC(t) + lG(t)\} \leq X_2(t) + lX_3(t), \tag{14.10}$$

in which,

γ_1 a coefficient matrix of energy consumption per production of the usual sectors (exogenous);

γ_2 a coefficient of energy consumption per production of the energy sector (exogenous);

γ_3 a coefficient matrix of energy consumption per production of the alternative energy sectors (exogenous);

γ_4 a coefficient of energy consumption per production of the CO_2 pollution abatement sector (exogenous);

γ_5 a coefficient of energy consumption per unit of the consumption and government expenditure sector (exogenous).

Disposable Income of Household. Disposable income, Y_d, is the income calculated by deducting direct tax from personal income. It is specified in the following equations:

$$Y_d(t) = [1-\tau] \times \left[Y_d^u(t) + Y_d^e(t) + Y_d^n(t) + Y_d^p(t) \right], \tag{14.11}$$

$$Y_d^u(t) = \psi_1 X_1(t), \tag{14.12}$$

$$Y_d^e(t) = \psi_2 X_2(t), \tag{14.13}$$

$$Y_d^n(t) = \psi_3 X_3(t), \tag{14.14}$$

$$Y_d^p(t) = \psi_4 X_4(t), \tag{14.15}$$

in which,

$Y_d(t)$ the disposable income of the household sector in term t;

$Y_d^u(t)$ the disposable income of the usual sectors in term t;

$Y_d^e(t)$ the disposable income of the energy sector in term t;

$Y_d^n(t)$ the disposable income of the alternative energy sectors in term t;

$Y_d^p(t)$ the disposable income of the CO_2 pollution abatement sector in term t;

τ the average rate of the direct tax to personal income (exogenous).

Investment and Saving. The net investment is equal to all savings:

$$I_n(t) + lE(t) - lM(t) = S^p(t) + S^g(t)\,, \tag{14.16}$$

$$I_n(t) = I^p(t) + I^g(t)\,, \tag{14.17}$$

$$I^p(t) = l\Delta K^p(t) - \delta_1 \cdot K_1(t) - \delta_2 \cdot K_2(t) - \delta_3 \cdot K_3(t) - \delta_4 \cdot K_4(t)\,, \tag{14.18}$$

$$I^g(t) = l\Delta K^g(t) - l\delta^g \cdot K^g(t)\,, \tag{14.19}$$

in which,

$I_n(t)$	the total amount of the net investment in term *t*;
$I^p(t)$	the total amount of the net investment of the household sector in term *t*;
$I^g(t)$	the total amount of the net government investment in term *t*;
$S^p(t)$	a saving of the household sector in term *t*;
$S^g(t)$	a government saving in term *t*;
δ^g	a depreciation rate of the government capital (exogenous);
$K^g(t)$	a government capital stock in term *t*.

Consumption and Saving Level of Household. The household allocates disposable income for consumption and saving at fixed rates respectively.

$$P(t)\tilde{C}(t) \le \tilde{\alpha} Y_d(t)\,, \tag{14.20}$$

$$S^p(t) \le \beta Y_d(t)\,, \tag{14.21}$$

where,

$$\alpha l + \beta = 1\,, \tag{14.22}$$

$$P(t) = \begin{bmatrix} P_1(t) \\ P_e(t) \end{bmatrix}, \tag{14.23}$$

$$\alpha = (\alpha_1, \alpha_2)\,, \tag{14.24}$$

$P(t)$	a row vector of consuming price of commodity in the usual sectors and the energy sector in term *t*;
$P_e(t)$	a consuming price of demanding commodity of the energy sector in term *t*;
α	a column vector of the propensity to consume of the household in the usual sectors (α_1) and the energy sector (α_2) (exogenous);
α_1	a column vector of the propensity to consume of the household in the usual sectors (exogenous);
α_2	the propensity to consume of the household in the energy sector (exogenous);
β	the propensity to saving of the household (exogenous).

Production Function. The production function is given as the following equations:

$$\lambda \cdot X(t) \leq K(t), \tag{14.25}$$

$$\lambda^g \cdot X(t) \leq K^g(t), \tag{14.26}$$

where,

λ	a diagonal matrix whose diagonal elements are the elements of capital and production ratio of the total sectors (exogenous);
λ^g	government capital and production ratio of the total sectors (exogenous).

Tax Revenue. The total tax revenue of the government consists of direct tax revenue, indirect tax revenue and the total emission tax revenue. Tax revenue is specified as the following equation.

$$\begin{aligned} Tre^g(t) &= \tau(t)\left\{Y_d^u(t) + Y_d^e(t) + Y_d^n(t) + Y_d^p(t)\right\} + \tau_1^i X_1(t) + \tau_2^i X_2(t) + \tau_3^i(t) X_3(t) + \\ &+ \tau_4^i X_4(t) + T(t)\left\{B_{41}X_1(t) + B_{42}X_2(t) + B_{43}X_3(t) + B_{44}X_4(t) + B_{45}l(C+G)\right\} \end{aligned} \tag{14.27}$$

where,

$Tre^g(t)$	the total tax revenue of the government in term *t*.

Government Expenditure. The government pays out the total tax revenue to government consumption, government saving, the depreciation of the capital, the subsidy for the alternative energy sectors and the countermeasure-related costs for CO_2-reduction.

$$Tex^g(t) = S^g(t) + lG(t) + \delta^g K^g(t) + \tau_3^i(t) X_3(t) + T(t) X_4(t), \tag{14.28}$$

$$Tre^g(t) = Tex^g(t), \tag{14.29}$$

$$S^g(t) = \phi \cdot Tre^g(t), \tag{14.30}$$

where,

$Tex^g(t)$	the total tax expenditure of the government in term *t*;
ϕ	a column vector of the propensity to save of the government (exogenous).

Dynamic Equations of Capital Stock. In the model, one term equals three years. Therefore, the dynamic equations of capital stock are specified as follows:

$$K(t+1)=(I-3\tilde{\delta})K(t)+3\Delta K^{p}(t)\,, \tag{14.31}$$

$$K^{g}(t+1)=(1-3\tilde{\delta}^{g})K^{g}(t)+3\ell\Delta K^{g}(t)\,, \tag{14.32}$$

where,

δ	a row vector of the depreciation rate to the private capital stock of the total sectors (exogenous);
I	a row vector whose elements are all the same as 1.

Total Emission Standards of CO2 Pollution.

$$Z(t)\le \overline{Z}(t)\,, \tag{14.33}$$

where

$\overline{Z}(t)$	the actual level of emissions of the CO_2 Pollution as of 1997 in China (exogenous).

Objective Function. The objective function is to maximize the gross national product (GNP).

$$\max\sum_{t=1}\frac{1}{(1+\rho)^{(t-1)}}GNP(t)\,, \tag{14.34}$$

where

$$\begin{aligned}GNP(t)=&(lP_1(t)\tilde{X}_1(t)\text{-}lP_1(t)A_{11}\tilde{X}_1(t))+(P_2(t)X_2(t)\text{-}P_1(t)A_{12}X_2(t))\\&+(lP_3(t)\tilde{X}_3(t)\text{-}lP_1(t)A_{13}\tilde{X}_3(t))\end{aligned}\quad, \tag{14.35}$$

and,

$GNP(t)$	gross national product in term *t*;
ρ	the average social depreciation rate in the economy (exogenous).

Assuming that Walras's Law obtains, the price of the commodity of each sector in market equilibrium is just expressed as a relative price. Therefore, in the model, the sector of the commodity of the "not elsewhere classified" sector (No.32) in the 1st term is set to be the *numeraire* of the integrated system, that is to say, P32(1) = 1. P32(2) is calculated endogenously according to the model. Other main parameters and the bases for calculation are listed in table 14.3.

Table 14.3. The main parameters used in the model.

Parameter	Symbol	Value
Term	*Time*	2
The adjustment factor	θ	0.500
The average direct tax rate	τ^{a}	0.026
Depreciation rate of the government	δ^{g}	0.048
Savings propensity of the government	φ^{b}	0.351
Proportion of savings in disposable income	β^{c}	0.334
Social depreciation rate	ρ	0.050

Note: Basis for calculation
a, c: Disposable income=Compensation of employees + Operating surplus
τ = (Disposable income - (Total household consumption +
Private domestic gross fixed capital formation)) /(Disposable income)
β = 1 - ((Total household consumption) / ((1 - τ) * Disposable income))
b: φ = (Annual government revenue - Government current expenditure) /(Annual government revenue)

14.4 Simulation

14.4.1 Design of the Scenarios

In the simulation, the focus is on the pollutant of anthropogenic CO_2 which is the main GHG resulting from energy use. The software *LINGO of LINDO Systems* for mathematical programming is used to perform the simulation.

In the Base Case (baseline case), the CO_2 emissions and economic trends in China were estimated by the simulation using the model calibrated on the I-O table of 1997. Here, neither CO_2 pollution abatement sector nor the carbon tax was considered. An emission restriction was not imposed.

In Case-0 and Case-10 (the reference case), the existence of alternative energy sectors was assumed, a carbon tax policy was implemented and the emission restrictions of CO_2 were also set. In the last term, 0% and 10% of upper constraint on the emissions of CO_2 was considered as compared with the amounts of emissions in the Base Case. The simulation of two cases of Case-0 (0% mitigation) and Case-10 (10% mitigation) were also performed.

14.4.2 Sensitivity Analysis

Next, attention was directed to how the results of the model were influenced by different values of the social discount rate ρ and adjustment factor θ (table 14.4). When the value of the social discount rate ranges between 0.025, 0.05 and 0.075, and when the adjustment factor θ changes from 0.5 to 1.5, GNP increases at various rates. This represents the non-linear effect in the model. Furthermore, the result shows that the value of the social discount rate has an important influence on GNP. For example, in the case where the value of the adjustment factor θ is fixed at 0.5, when the value of social discount rate (0.075) becomes larger than the annual rate 0.05, which is used in the model, GNP has a downward tendency and is decreasing 0.6% in Case-10. On the other hand, when the value of the social discount rate (0.025) becomes smaller than the annual rate 0.05, GDP has a tendency to increase and increases by 0.1% in Case-10. Since a high social discount rate would put the future generation at quite a disadvantageous position, the correctness of the non-linear model is also verified.

Table 14.4. Sensitivity analysis based on the change of the social depreciation rate and the adjustment factor (Unit: 0.1 billion CNY).

Index	ρ	θ	GNP in Case-10
1	0.025	0.5	74,931
2	0.025	1	81,901
3	0.025	1.5	101,070
4	0.05	0.5	74,867
5	0.05	1	88,079
6	0.05	1.5	112,301
7	0.075	0.5	74,447
8	0.075	1	94,750
9	0.075	1.5	145,509

14.4.3 Simulation Result: the Integrated Evaluation of Optimal Carbon Tax Policy in China

Optimum Carbon Tax Rate. An optimum carbon tax rate required in order to attain the mitigation target was estimated as 85.25 CNY/t-CO_2 and 100.00 CNY/t-CO_2 (see table 14.5). According to the present ‘Resource Tax Enactment in People’s Republic of China,’ the optimal carbon tax rate is nearly as much as 3 to 17 times the resource taxes on crude oil and coal, respectively. This indicates that the inadequate pricing of natural resources and energy causes inefficient use and environmental damage. At this stage, it is impossible to apply a carbon tax policy solely because the carbon tax rate is optimal. It is also quite difficult to establish a country-wide environmental taxation system in China.

Table 14.6 shows briefly various carbon tax rates in the world. It is derived from various sources (Japan Ministry of the Environment (ME), 2004; IPCC, 1996b, etc.). Assuming that 1 JPY = 0.0662 CNY and 1 US$ = 8.2898 CNY as of 1997 and the average emission coefficient of electricity equals 0.38 kg-CO_2/kWh (NBS, 2003; ME, 2004), the carbon tax rate derived in Case-0 implies a rates of 10 US$/t-$CO_2$ and 0.49 JPY/kWh. Through comparison among various carbon tax rates in the world, this would be the lowest carbon tax rate.

Table 14.5. Optimum carbon tax rate, impact on public welfare by case and electricity generated by alternative energy sectors.

Case	The optimum carbon tax rate ($CNY/t-CO_2$)	GNP (0.1 billion CNY)		Solar power (10,000 kWh)		Wind power (10,000 kWh)	
		1st term	2nd term	1st term	2nd term	1st term	2nd term
Base case	0	75,704	102,398	0.45	0.60	16.5	27.0
Case-0	85.25	75,246	103,418	0.50	0.65	23.6	46.7
Case-10	100.00	74,867	104,419	0.52	0.86	36.2	75.3

Table 14.6. Various carbon tax rates in the world.

Country	carbon tax ($US\$/t-CO_2$)	Carbon tax on electricity (JPY/kWh)	Year Initiated
Finland	392	0.95	1990
Sweden	785	0	1991
Norway	667	0	1991
Denmark	539	1.77	1992
Britain	389	0.85	2001
Switzerland	726	0	2005 (scheduled)
Germany	348	2.69	1999
France	840	0	Illegal judgment in 2000
Netherlands	326	8.59	1988 or 1996
Italy	818	0	1999
Japan	73	0.25	2005 (scheduled)
U.S.	293	0	
Canada	191	0	

Source: ME, 2004; IPCC, 1996b, and other sources.

Impact on the Economy. The impact on economic welfare is shown in table 14.5. GNP in the 1st term for Case-0 shows a slight decrease of 0.6% compared with the GNP in the 1997 base year. It could be concluded that the level of production in this term as compared to the achievements in 1997 was scarcely maintainable due to the introduction of the carbon tax. This results from the delay in bringing on-line the alternative energy sectors and the CO_2 pollution abatement sector which were in the process of establishment. This has resulted in the difference between the result of the simulation and the achievements in 1997 for CO_2 emissions and the production of Electricity sector. In the 2nd term, GNP in Case-10 exceeded the achievements in 1997 base and grew about 1.0% more than that in Case-0. Minimizing the short-term negative impact by the introduction of a carbon tax, the combination of a carbon tax policy and a concrete mitigation target could result in the Chinese economy growing more rapidly.

Table 14.7. Change of products by sector, in 10 thousand CNY.

Industry	Basic Case	Case-0		Case-10	
		1st term	2nd term	1st term	2nd term
1	250,373,327	250,373,327	250,489,294	250,373,327	235,085,904
2	68,283,887	68,283,887	62,629,902	55,880,548	39,855,379
3	137,925,948	151,718,558	183,414,077	151,718,558	178,431,721
4	131,451,437	131,451,437	151,048,147	131,451,437	151,048,147
5	46,787,245	46,522,346	50,778,057	46,197,844	48,685,532
6	110,069,221	110,069,221	119,022,723	110,069,221	118,162,452
7	30,981,911	30,952,310	36,971,637	29,962,384	32,220,191
8	88,074,034	86,633,232	93,208,108	85,823,206	88,255,864
9	54,067,064	53,506,716	52,683,595	54,175,862	50,764,996
10	23,683,242	24,186,982	21,039,063	23,516,977	19,255,452
11	49,832,738	49,316,292	52,983,748	49,088,059	51,802,146
12	82,266,501	80,082,016	78,296,959	74,045,882	68,921,750
13	104,564,055	104,564,074	108,747,575	104,564,074	108,101,446
14	53,138,393	58,415,482	56,627,595	58,415,482	56,627,595
15	5,773,251	5,773,251	6,215,866	5,773,251	6,215,866
16	107,159,365	107,159,365	120,957,574	107,159,365	118,264,094
17	173,855,000	173,855,000	183,633,281	172,904,521	182,384,057
18	2,745,749	2,769,478	3,111,198	2,442,452	2,535,390
19	9,169,614	10,027,649	10,480,637	9,845,504	9,614,949
20	110,485,707	110,485,707	131,604,864	110,485,707	131,604,864
21	35,952,759	35,952,759	46,516,014	35,952,759	46,516,014
22	18,553,568	18,553,568	24,004,786	18,553,568	24,004,786
23	50,662,699	55,625,923	64,309,940	55,625,923	57,339,326
24	19,589,200	19,589,200	21,652,501	19,589,200	22,212,504
25	49,751,546	47,461,508	34,826,086	45,019,284	34,826,086
26	25,783,142	26,405,303	20,445,477	25,747,531	20,445,477
27	17,977,638	19,775,402	19,775,402	19,775,402	19,775,402
28	18,112,972	18,112,972	17,496,872	18,112,972	17,496,872
29	7,193,965	7,193,965	8,026,824	7,138,874	7,813,933
30	50,102,387	50,102,387	64,822,946	50,102,387	64,822,946
31	2,529,125	2,560,493	2,826,991	2,549,080	2,782,365
32	23,808,482	24,026,779	23,125,157	23,892,999	22,417,609
33	37,736,318	38,062,439	42,758,878	33,567,939	34,845,241

Next, the impact at the sector level is evaluated. Table 14.7 shows the change of products in each sector. In the 1st term of Case-0, the winners by the carbon tax introduction are Food Products (No.3), Transport Equipment and Automobiles (No.14), Transport (No.23), and Water (No.19); here, production levels increase 10.0%, 9.9%, 9.8% and 9.4% respectively. The sectors most negatively affected would be Public Works (No.25), General Machinery (No.12), Non-metallic Mineral Products (No.8), with decrease in production of 1.6%, 2.7% and 4.6%, respectively. In the 1st term of Case-10, the winners under a policy mix of a carbon tax introduction and 10% mitigation are Food Products (No.3), Medical, Health and Social Insurance Services (No.27), Transport Equipment and Automobiles (No.14), Transport (No.23), and Water (No.19),with production level increases about 10.0%, 10.0%, 9.9%, 9.8% and 7.4%, respectively. The worst-hit sectors are Mining (No.2), Electricity (No.33), Gas (No.18), General Machinery (No.12), and Public Works (No.25), with production decreases 18.2%, 11.0%, 11.0%, 10.0% and 9.5% respectively. This also shows change of source of energy use. From the 1st term to the 2nd term of Case-0, winners are the Miscellaneous Personal Services (No.30), Real Estate Services (No.22), Finance and Insurance (No.21), Food Products (No.3), Coal and Petroleum Products (No.7), and Whole Sale and Retail Service (No.20), with production level increases 29.4%, 29.4%, 29.4%, 20.9%, 19.4% and 19.1%, respectively. On the other hand, the worst-hit sectors are Public Works (No.25), Education and Research Services (No.26), Non-ferrous Metals (No.10), and Mining (No.2), with production level decreases 26.6%, 22.6%, 13.0% and 8.3%, respectively.

From the 1st term to the 2nd term of Case-10, the winners are the Miscellaneous Personal Services (No.30), Real Estate Services (No.22), Finance and Insurance (No.21), Wholesale and Retail Service (No.20), and Food Products (No.3), with production level increases of 29.4%, 29.4%, 29.4%, 19.1% and 17.6%, respectively. The worst-hit sectors are Mining (No.2), Public Works (No.25), Education and Research Services (No.26), and Non-ferrous Metals (No.10), with production level decreases of 28.7%, 22.6%, 20.6% and 18.1%, respectively.

The Prediction of the Alternative Energy Sectors' Growth. Table 14.5 also indicates the electricity generated by the alternative energy sectors by case. In the two alternative energy sectors, the electricity generated by wind power expands more rapidly. In Case-0, it increases from 236,100 kWh to 467,000 kWh while in Case-10 is the growth is from 362,000 kWh to 672,200 kWh. However, when we talk about the relative introduction rate of the alternative energy sectors, a different perspective is obtained.

When the two kinds of alternative energies are compared from the viewpoint of the introduction of relative rate based on Base Case, the relative rate for the introduction of wind power is the higher, 143.1% in Case-0 and 219.4% in Case-10. This finding suggests that the use of wind power should grow considerably. Concerning the introduction of alternative energy, generally speaking, the smaller is the emission coefficient the faster is the introduction. Therefore, the introduction of solar power should progress more rapidly, but this is not the case according to the results of the simulation. This may result from the effect of the energy cost which is another significant factor. According to equation (14.10) in the model, it is assumed that the alternative energy competes with the existing conventional electricity supply. Although there is little absolute quantity of alternative energy, it can be substituted by the existing conventional electricity resource; further at the same time, substitutability also exists between the two kinds of alternative energies. However, this is not the only case for wind power to be introduced. Since they can be substituted mutually between the existing conventional electricity resource and the alternative energy, the energy costs can be gradually decreased. As previously calculated from the alternative energy cost model, the cost in 1997 of solar power is 2.5 CNY/kWh and that of wind power is 0.75 CNY/kWh. However, in the 2nd term, solar power cost remains the same as 2.5 CNY/kWh, but wind power cost is 0.67 CNY/kWh. For this reason, wind power with a larger emission coefficient but a lower cost progresses faster than solar power as a result of a higher rate of an introduction. It is apparent that the introduction of wind power progresses more because of the balance between both the emission coefficient and energy cost.

As shown in Table 14.5, the electricity generated by the two alternative energy sectors increases with the introduction of a carbon tax. Therefore, a carbon tax could promote the growth of the alternative energy sectors. Further, with the mixture of the policy instruments of not only a carbon tax but also a CO_2 specific mitigation target, the growth of the alternative energy sectors could be accelerated.

Table 14.8. Optimum carbon tax rate by sector in Case-0 (Unit: 10 thousand CNY/ 1 million CNY productions).

NO.	Sectors	Case-0
1	Agriculture	0.0752
2	Mining	1.1649
3	Food products	0.0490
4	Textiles and clothing	0.0965
5	Paper and wooden products	0.5673
6	Chemical products	0.4835
7	Coal and petroleum products	1.4300
8	Non-metallic mineral products	2.3183
9	Iron and steel	1.0192
10	Non-ferrous metals	0.2809
11	Fabricated metal products	0.1122
12	General machinery	0.0977
13	Electric machinery	0.0528
14	Transport equipment, automobiles	0.0512
15	Precision apparatus	0.0595
16	Other manufactures	0.1217
17	Building and construction	0.0298
18	Gas	0.1348
19	Water	0.0026
20	Whole sale and retail service	0.0392
21	Finance and insurance	0.0049
22	Real estate services	0.0021
23	Transport	1.3720
24	Communications	0.0627
25	Public works	0.0605
26	Education and research services	0.0499
27	Medical, health and social insurance services	0.0024
28	Other public services	0.0875
29	Miscellaneous business services	0.0026
30	Miscellaneous personal services	0.0025
31	Office supplies	0.0023
32	Not elsewhere classified	0.1288
33	Electricity	2.5617

Optimum Carbon Tax Rate by Sector. Next, the optimum carbon tax rate by sector in Case-0 is evaluated. Table 14.8 shows the optimum carbon tax rate transition in each sector in Case-0. Since the unit in this table is 10 thousand CNY per million CNY of production, each value can also represent the optimum carbon tax rate for each sector. Although most of the optimum carbon tax rates are smaller than 1%, the optimum carbon tax rates of Iron and steel (No.9), Mining (No.2), Transport (No.23), Coal and petroleum products (No.7), Non-metallic mineral products (No.8) and Electricity (No.33) are quite larger, but the highest optimum carbon tax rate is 2.5617% for Electricity (No.33).

14.5 Conclusions

We have summarized the above results of integrated evaluation of an optimal carbon tax policy in China and present the following conclusions. China's economy has been enjoying a high-growth period since the late 1990s due to the gradual introduction of economic reforms. Especially among the four aspects of basic necessities of life, namely 'clothing', 'diet', 'housing' and 'traveling,' China was able to satisfy demands for 'clothing' and 'diet' with an amazing growth rate. This implies that China has been able to de-couple energy use from the high GDP growth rate because meeting the demands for 'clothing' and 'diet' has required less energy. However, with increasing wealth, more and more people are seeking a higher standard of living, mainly on the latter two aspects of 'housing' and 'traveling.' Increased consumption to satisfy non-basic needs has led to the demand for more energy, and so to more demand for industrial products. It has also led to more waste, including CO_2 emissions. Being fully committed to the regulations of the World Trade Organization (WTO), it would be difficult for China to attain the emission mitigation target through energy-saving measures alone.

The optimal carbon tax rate which attains the mitigation target is estimated to be 85.25 CNY/t-CO_2. The optimal carbon tax rate is nearly as much as 3 to 17 times the current resource taxes on crude oil and coal, respectively in China. Through comparison of various carbon tax rates calculated and discussed in developed countries, it turns out that the optimal carbon tax rate is the lowest in the world while a carbon tax on electricity is not quite so low. Therefore, as one of the countermeasures against global warming, it is expected that it would be quite difficult to establish a country-wide environmental taxation system in China.

The impact of introducing the carbon tax on energy-intensive sectors is significant, especially on such energy-intensive sectors as Electricity, Gas, Water, Mining, General Machinery and Non-ferrous Metals. Therefore, when considering the introduction of a carbon tax, a preferential tax system for energy-intensive sectors might be considered. Therefore, it could be expected that the introduction of a carbon tax in combination with appropriate policies might be implemented in China.

14.6 Future Issues

As to future issues, there are two potential developments that could improve the value of the model. First, nuclear power generation was incorporated into the existing electricity sector in this research. In the foreseeable future, it is difficult to assume another option other than nuclear power as a non-fossil energy source in China. Moreover, a nuclear power utility would be a significant step in the development of an artificial carbon management policy. Therefore, one of the future issues is to conduct a simulation in which nuclear power generation is separated from coal- and oil-fired power generation.

Secondly, because China's CO_2 emissions, SO_2 pollution and acid rain come from coal combustion, it will be useful to integrate the existing environment policy for reducing SO_2 pollution with the introduction of a carbon tax policy. The former policy focuses more attention at the local level while the later is more related with global environmental policy.

References

Brown LR, Flavin C, French H (2000) State of the world: A Worldwatch Institute report on progress toward a sustainable society. W.W. Norton & Company, New York

Bye B (2002) Taxation, unemployment, and growth: Dynamic welfare effects of 'green' policies. Journal of Environmental Economics and Management 43:1-19

Chae Y, Hope C (2003) Integrated assessment of CO_2 and SO_2 policies in North East Asia. Climate Policy 3 (Supplement):S57-S79.

Chen ZL (1999) Studies on the energy demand/supply in China's economy. Doctoral dissertation presented to the Graduate School of Environment and Life Engineering, Toyohashi University of Technology, Japan (in Japanese)

China Renewable Energy Power Generation and Heat Use Research Item Group of State Economic and Trade Commission (1999) China renewable energy technical evaluation. China Environmental Science Press, Beijing, China (in Chinese)

China National Bureau of Statistic (NBS) (1996) China energy statistical yearbook. China Statistics Press, Beijing, China (in Chinese)

China National Bureau of Statistic (NBS) (1999a) China energy statistical yearbook. China Statistics Press, Beijing, China (in Chinese)

China National Bureau of Statistic (NBS) (1999b) China statistical yearbook 1999. China Statistics Press, Beijing, China, (in Chinese)

China National Bureau of Statistic (NBS) (1999c) Input-output table of China 1997. China Statistics Press, Beijing, China (in Chinese)

China National Bureau of Statistic (NBS), (2003). *China Statistical Yearbook 2003*. China Statistics Press, Beijing, China (in Chinese)

Edit Committee of China Farm Village Energy Yearbook (1997) China farm village energy yearbook. China Agriculture Press, Beijing, China (in Chinese)

Higano Y (1996) Distribution of the value added to the input of the environmental goods based on the materials balance principle. Studies in Regional Science 26:181-187 (in Japanese)

Higano Y (1997) Optimal taxes on the emission of air pollutants in Japan. Discussion paper read at session in honor of Professor Hirotada Kohno: Environmental Policy and Sustainability at the 15th Pacific Regional Science Conference, December, Wellington

Intergovernmental Panel for Climate Change (IPCC) (1996a) Climate Change 1995: The science of climate change, Houghton JT et al. (eds). Cambridge University Press, Cambridge, UK

Intergovernmental Panel for Climate Change (IPCC), (1996b). Climate change 1995: Economic and social dimensions of climate change, James PB et al. (eds). Cambridge University Press, Cambridge, UK.

International Energy Agency (IEA) and Organization for Economic Co-operation and Development (OECD) (1999) World energy outlook: Looking at energy subsidies: Getting the prices right. OECD, Paris

International Energy Agency (IEA) and Organization for Economic Co-operation and Development (OECD) (2000). World energy outlook. OECD, Paris

International Energy Agency (IEA) and Organization for Economic Co-operation and Development (OECD) (2001) World energy outlook: Insights: Assessing today's supplies to fuel tomorrow's growth. OECD, Paris

International Energy Agency (IEA) and Organization for Economic Co-operation and Development (OECD) (2004) World energy outlook 2004. OECD, Paris

Kratena K (2004) 'Ecological value added' in an integrated ecosystem–economy model: An indicator for sustainability. Ecological Economics 48:189-200

Li B, Higano Y (2001) Optimum transport investment under environmental restrictions: A simulation of the chinese economy. Studies in Regional Science 31:63-77 (in Japanese)

Li ZD (2003) An econometric study on China's economy, energy and environment to the year 2030. Energy Policy 31:1137-1150

Liu WQ, Ganb L, Zhang XL (2002) Cost-competitive incentives for wind energy development in China: Institutional dynamics and policy changes. Energy Policy 30:753-765

Martin W (1993) Modeling the post-reform Chinese economy. Journal of Policy Modeling 15:545-579

Matsui K (2000) Energy economy and policy theory. Sagano Press, Tokyo (in Japanese)

Ministry of the Environment in Japan (2004) Proceedings summary and minutes of comprehensive policy and an earth environment joint section of the Central Environment Council. *http://www.env.go.jp/council/16pol-ear/yoshi16.html*, Read on 14, November 2004 (in Japanese)

Miyazawa K (1995) Guide to input-output analysis. Japan Economical Newspaper Press, Tokyo (in Japanese).

Mizunoya T, Higano Y (1999) Air pollutants emission reduction and the optimal tax-subsidy policy for Japan. Studies in Regional Science 30(3):85-104

Moffat, I. and N. Hanley (2001). "Modeling Sustainable Development: Systems Dynamic and Input-Output Approaches." *Environmental Modeling & Software*, 16, 545-557.

Mori, S., (1992). *Earth Environment and Resources Problem*. Tokyo, Iwanami Press (in Japanese).

National Research Council, Chinese Academy of Sciences and Chinese Academy of Engineering (2000) Cooperation in the energy futures of China and the United States. National Academy Press, Washington, DC

Ng CT (2004) A Learning-by-doing energy model based on dynamic programming. International Institute for Applied Systems Analysis interim report IR-03-048. *http://www.iiasa.ac.at/Publications/Documents/IR-03-048.pdf*, read on 30, March 2004

Nishiyama T (2001) An earth energy theory. Ohm Press, Tokyo, Japan (in Japanese)

Nijkamp P (1977) Theory and application of environmental economics. North Holland, Amsterdam

Nordhaus DW (1979) The efficient use of energy resources. Yale University Press, New Haven, CT

Organization for Economic Co-operation and Development (OECD) (1999) Environmental taxes: Recent developments in China and OECD countries. OECD, Paris

Research Institute for Development and Finance of Japan Bank for International Cooperation (2000) Energy balance simulations to 2010 for China and Japan. Research Institute for Development and Finance of Japan Bank for International Cooperation, Tokyo (in Japanese)

Sinton EJ, Fridley GD (2000) What goes up: Recent trends in China's energy consumption? Energy Policy 28:671-687

Xu ZM, Cheng GD, Chen DJ, Templet HP (2002) Economic diversity, development capacity and sustainable development of China. Ecological Economics 40:369-378

Yagita H, Kondo Y, Kobayashi M, Inaba A (2000) A comparative evaluation of CO_2 mitigation options for Japan using the NICE model. Japan Energy Society 79:118-128 (in Japanese)

Yan CL (ed) (2002) China energy development report 2001. Jiliang Press, Beijing, China (in Chinese)

Part D: Methodological Advances—Econometric Models

15 Effects of Trade on Emissions in an Enlarged European Union: Some Comparative Dynamics Analyses with an Empirically Based Endogenous Growth Model[*]

Nazmiye Balta-Ozkan[1], Kieran P. Donaghy[2], and Clifford R. Wymer[3]

[1]Policy Studies Institute, London, United Kingdom
[2]Department of City and Regional Planning, Cornell University, Ithaca, NY, USA
[3]International Monetary Fund (Retired), Auckland, New Zealand

15.1 Introduction

On May 1, 2004, the European Union (EU) expanded from 15 member states (EU15) to 25 member states. Both the EU15, as a single entity, and most of the new member states, individually, are signatory parties to the Kyoto Protocol with different levels of greenhouse gas (GHG) emission reduction targets to be achieved by 2012. While the emissions of these new entrants are already below their Kyoto targets, the EU15 has not reached its aggregate target yet. Economic integration will certainly increase the volume of trade between these countries, through which their economic growth rates and emissions will be affected.[1] In this study, we analyze possible environmental and economic impacts of increasing trade between the EU15 and the so-called 'new accession countries' (NAC13).[2] In

---

[*] This study draws extensively upon Balta (2004), a PhD dissertation submitted to the University of Illinois at Urbana Champaign. An earlier version of this chapter was presented at the 1st Atlantic Workshop on Environmental and Energy Economics: Economic Modelling of Climate Change Policies, A Toxa, Spain, September, 2004.

[1] Impacts of trade on the environment from enlargement have been the topic of a workshop, whose proceedings have been edited by Maxwell and Reuveny (2005).

[2] Members of the European Union prior to May 1, 2004, forming the so-called EU15 group, are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom. These countries have been joined by Cyprus, the Czech Republic, Estonia, Hungary, Latvia,

focusing on these issues, we take a different approach than has been followed in the extant literature. We employ a two-bloc representative agent model with both endogenous growth and endogenous technological change whose parameters have been estimated from available time series.

There are in the literature different theories about the interaction of trade and environment. Increasing trade in the EU between the new accession countries and longer standing member states will lead to increases in aggregate income. However, with increased income in the enlarged union will come increased demand for output and increased emissions, the so-called *size* effect. Environmental policies might very well change this process. Porter (1991) has claimed that strict environmental policies will induce firms to innovate, the so-called *technology* effect. These technological advancements will bring an output combination that is less polluting, the so-called *composition* effect. On the other hand, it is also claimed that strict regulations can stimulate firms to move to countries with less stringent environmental policies, with the result being that poorer countries become *pollution havens*.

Interaction of trade and environmental policies has been analyzed via a mix of analytical and econometric approaches in the literature on trade and the environment. (For a very recent survey on climate and trade, see Galeotti and Kemfert 2004). The scope and focus of analytical models have extended from North-South trade in a world of two regions (by, *inter alia*, Chichilnisky 1994, Abraham et al. 1998) to national trade policies in which governments/firms act strategically to form/respond to environmental policies to keep competitiveness of industries (Ulph and Ulph 1996). Econometric studies, in turn, have focused on the effects of environmental policies on the location of firms. (For a recent survey, see Jeppessen et al. 2002.). However, these approaches lack a macroeconomic perspective and do not take into account structural features of the EU15 and NAC13 economies.

Paralleling recent developments in macroeconomic modeling, models employed in research in environmental economics have incorporated endogenous growth (see, *inter alia,* Smulders 1995) and endogenous technical change. Few scholars have used both of them in a single framework (*inter alia*, see Bovenberg and Smulders 1996). While studies with endogenous growth are limited to analytical models, application of models with endogenous technical change to environmental problems has been both analytical and empirical. (For a recent review of studies on environmental-policy-induced technological change, see Heidug and Bertram 2003.) Empirical studies of endogenous technical change are mostly carried out with general equilibrium models (Nordhaus 1997, Buonanno et al. 2000).[3] As such, these models allow regional transactions to be characterized and impacts of environmental policies on the each bloc's economy to be analyzed.

Lithuania, Malta, Poland, the Slovak Republic, and Slovenia. Bulgaria and Romania will start accession negotiations in 2007, and Turkey is a candidate country to join the EU. In this study, all these thirteen countries are referred to NAC13.

[3] For a list of general equilibrium model applications to environmental problems, see Adkins and Garbaccio (1999).

While analyses conducted with these general equilibrium models have advanced our understanding of relationships between the environment and economy, they have not accounted for the evolution of structural characteristics of the economies nor contributed much insight into their workings. Model parameter values, moreover, have usually been set by calibration to certain benchmark years, instead of by econometric estimation with time series data. Consequently, simulations with these models may not capture the historical stylized facts of the economies they represent. Hence, they do not inspire much confidence about future predictions. Also, simulations with these models do not present full transition paths of policy responses.

In light of these considerations, we adopt an intertemporally optimizing representative-agent modeling framework. With this framework, we can take into account structural features of the EU15 and NAC13 regional economies in which future-oriented policy behavior is (hypothetically) optimally determined. In this model, both aggregate growth rates and technical change are endogenously determined. As Donaghy et. al (1999) have shown, the feature of endogenous growth determination is important for capturing effects of policies in other economies transmitted via real exchange rates and trade.

In the next section, we discuss specification of a two-bloc macrodynamic model of the EU15 and NAC13. In the third section, we discuss the details of the model's estimation with aggregate data on the EU15 and NAC13 economies, and the empirical results obtained. In the fourth section, we present results of policy simulations to analyze impacts of changing trade patterns on these economies and their emissions patterns. The last section concludes with some observations about this undertaking.

15.2 Specification of the Representative Agent Model

In the two-region setting we consider, agents may hold regional capital and traded bonds, through which they gain indirect ownership of foreign assets. Because they do not hold foreign assets directly, financial markets are 'incomplete.' Each region specializes in the production of a single good, but agents in the two regions may consume goods that are produced within the region and imported from either the other region or the 'rest of the world.' In each region the representative agent has access to the world bond market but is not free to borrow or lend as much as he or she might wish at the world interest rate. Rather, the effective interest rate faced by each agent is a function of his or her debt- or credit-equity position. The agent's borrowing and lending behavior must moreover satisfy an intertemporal solvency constraint. Adjustment to shocks occurs partly through the effective interest rates, partly through the relative price of foreign and domestic goods, and partly through the adjustment costs associated with the accumulation of physical and intellectual capital. While taxes on the representative agent of each country's economy play an important role in this model, tariffs do not, since they do not factor significantly in trading between countries from the regions examined.

In the EU15 regional economy, output, Y is produced with the physical capital domiciled in the region, K, the regional labor force, L, energy inputs, E, and intellectual (or human) capital, HC. We assume the production technology follows a *Cobb-Douglas* specification in which increasing returns to scale is allowed to be achieved depending on the output elasticity of human capital.

$$Y = \alpha e^{\rho t} HC^{\varepsilon} L^{\mu_1} K^{\mu_2} E^{\mu_3}, \tag{15.1}$$

where $\alpha, \rho, \mu_1, \mu_2, \mu_3, \varepsilon > 0$ and $\mu_1 + \mu_2 + \mu_3 = 1$.

In equation (15.1), α is a scale parameter, and ρ is the Harrod-neutral rate of technological progress. μ_1, μ_2, μ_3 and ε are share parameters, and present output elasticities for each factor of production.

Regional output is used domestically for investment in physical capital, I, investment in human capital, R, and consumption, C. The rest of regional output is exported for foreign consumption, either to the NAC13 region, MGS^*, or the rest of the world, $RXGS$, or invested in foreign assets, *b*.

The representative agent of the EU15 derives utility over an infinite time horizon by consuming the regionally produced commodity, C, imports from the NAC13 region, MGS, and imports from the rest of the world, $RMGS$. We assume that the following isoelastic intertemporal utility function represents utility the EU15 agent derives from consumption of these commodities.

$$\int_0^{\infty} \frac{1}{\gamma} \left(C \cdot MGS^{\eta_1} RMGS^{\eta_2} \right)^{\gamma} e^{-\beta t} dt, \tag{15.2}$$

where $\eta_1, \eta_2 > 0$; $-\infty < \gamma < 1$; $\gamma(\eta_1 + \eta_2) < 1$; and $1 > \gamma(1 + \eta_1 + \eta_2)$.

In objective functional (15.2), the exponent β represents the agent's subjective intertemporal discount rate. His/her intertemporal elasticity of substitution is given by $s = 1/(1-\gamma)$. Concavity of the utility function in its arguments suggests certain constraints be imposed upon the parameters as presented above. Labor is assumed to be supplied inelastically.[4]

The representative agent incurs adjustment costs for investment expenditures to both physical and capital stocks. We assume these costs are quadratic. The cost of adjustment to the physical capital stock is represented by the following function:

$$\Phi(I,K) = I + a_1 \frac{I^2}{2K} = I\left(1 + \frac{a_1 I}{2K}\right). \tag{15.3}$$

Similarly, the cost of adjustment to human capital stock brought about by expenditures on research and development, R, is given by the following function:

---

[4] Although it is arguable that labor should be an argument of the utility function, we have assumed inelastic labor supplies for the two blocs of countries for analytical convenience.

$$\Theta(R,HC) = R + a_2 \frac{R^2}{2HC} = R\left(1 + \frac{a_2 R}{2HC}\right). \tag{15.4}$$

The agent accumulates foreign assets in the form of bonds (and other financial assets), *b*, and receives income at an effective interest rate, *r*, from this investment. However, all of his/her income is subject to taxes. He/she pays taxes on income from factors of production, bond income, and consumption expenditures at rates τ_y, τ_b, and τ_c. These tax revenues are rebated to the agent in the form of a lump-sum transfer payment, T. The representative agent, then, faces the following instantaneous budget constraint:

$$\begin{aligned} \dot{b} &= \left(1-\tau_y\right)Y + \left(1-\tau_b\right)rb - \left(1+\tau_c\right)\left(C + \sigma MGS + \sigma_r RMGS\right) \\ &-I\left(1+\frac{a_1 I}{2K}\right) - R\left(1+\frac{a_2 R}{2HC}\right) + T. \end{aligned} \tag{15.5}$$

This budget equation incorporates taxes from consumption of imported goods. The variables σ and σ_r denote the real exchange rates prevailing in trade between the EU15 and the NAC13 and the EU15 and the rest of the world, respectively. Also embedded in this budget constraint are costs of adjustments to changes in the physical and human capital stocks, specified in equations (15.3) and (15.4).

Both physical and human capital investments are taken to be net of depreciation. Hence, the agent faces the following constraints on accumulation of physical and human capital stocks:

$$\dot{K} = I, \tag{15.6}$$

$$\dot{HC} = R. \tag{15.7}$$

In order to maximize the discounted stream of his/her utility, the representative agent must choose his/her levels of consumption of goods produced in the EU15, the NAC13, and rest of the world, C, MGS, and $RMGS$, rates of investment in physical and intellectual capital, I and R, and the rate of asset accumulation (or disaccumulation), $\dot{b}$, determined by residual, subject to the equations constraining changes in financial, physical, and human capital and initial endowments of these capital stocks,

$$K(0) = K_0; \quad HC(0) = HC_0; \quad \text{and} \quad b(0) = b_0. \tag{15.8}$$

The present-value Hamiltonian[5] for the agent's intertemporal optimization problem is given by the following equation:

---

[5] In setting up this optimization problem, we follow Turnovsky's (1997) approach.

$$\begin{aligned} H = {} & \frac{1}{\gamma}\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma} e^{-\beta t} \\ & +\lambda e^{-\beta t}\left[\begin{array}{l} \left(1-\tau_y\right)Y+\left(1-\tau_b\right)rb-\left(1+\tau_c\right)\left(C+\sigma MGS+\sigma_r RMGS\right) \\ -I\left(1+\dfrac{a_1 I}{2K}\right)-R\left(1+\dfrac{a_2 R}{2HC}\right)+T-\dot{b} \end{array}\right] \\ & +q_1' e^{-\beta t}\left(I-\dot{K}\right)+q_2' e^{-\beta t}\left(R-\dot{HC}\right). \end{aligned} \tag{15.9}$$

In this expression there are three co-state variables. The first, λ, denotes the marginal value (or marginal utility) of wealth and q_1' and q_2' are the marginal values of the agent's stock of physical and human capital, respectively. The analysis can be simplified by using the marginal value of wealth as numeraire, in which case $q_1 = q_1'/\lambda$ and $q_2 = q_2'/\lambda$ become market values of additional units of physical and human capital.

Full details of the model's derivation are provided in Appendix 15.A. The first-order optimality conditions with respect to the control variables $C,\ MGS,\ RMGS,\ I$ and R are as follows:

$$\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma-1} MGS^{\eta_1} RMGS^{\eta_2} = \lambda\left(1+\tau_c\right), \tag{15.10}$$

$$\eta_1\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma-1} C \cdot MGS^{\eta_1-1} RMGS^{\eta_2} = \lambda\left(1+\tau_c\right)\sigma, \tag{15.11}$$

$$\eta_2\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma-1} C \cdot MGS^{\eta_1} RMGS^{\eta_2-1} = \lambda\left(1+\tau_c\right)\sigma_r, \tag{15.12}$$

$$1+\frac{a_1 I}{K} = q_1, \tag{15.13}$$

$$1+\frac{a_2 R}{HC} = q_2. \tag{15.14}$$

By rewriting equations (15.13) and (15.14), we can derive the optimal rates of physical and intellectual capital accumulation in terms of the market values of installed physical capital (Tobin's *q*) and human capital attained.

$$\begin{aligned} \frac{I}{K} &= \frac{\dot{K}}{K} = \frac{q_1-1}{a_1}, \\ \frac{R}{HC} &= \frac{\dot{HC}}{HC} = \frac{q_2-1}{a_2}. \end{aligned} \tag{15.15}$$

Standard intertemporal macroeconomics optimality conditions (the co-state equations) with respect to $b,\ K,$ and HC imply the following arbitrage relationships:

$$\beta - \frac{\dot{\lambda}}{\lambda} = \left(1-\tau_b\right)r, \tag{15.16}$$

$$\frac{\dot{q}_1}{q_1}+\frac{\left(1-\tau_y\right)Y_K}{q_1}+\frac{\left(q_1-1\right)^2}{2a_1q_1}=\left(1-\tau_b\right)r\ , \tag{15.17}$$

where $Y_K=\mu_2\alpha e^{\rho t}HC^{\varepsilon}L^{\mu_1}K^{\mu_2-1}E^{\mu_3}$,

$$\frac{\dot{q}_2}{q_2}+\frac{\left(1-\tau_y\right)Y_{HC}}{q_2}+\frac{\left(q_2-1\right)^2}{2a_2q_2}=\left(1-\tau_b\right)r\ , \tag{15.18}$$

where $Y_{HC}=\varepsilon\alpha e^{\rho t}HC^{\varepsilon-1}L^{\mu_1}K^{\mu_2}E^{\mu_3}$.

The first of these arbitrage relationships is the Keynes-Ramsey consumption rule, equating the marginal return on consumption to the after-tax return on holding a foreign bond. The second and third relationships equate the after-tax rates of return on domestic physical and human capital to the after-tax bond yield. The second condition also highlights the importance of adjustment costs and the market price of capital in equilibrating the rate of return.

It is the fact that the rates of growth in physical, human, and financial capital are determined by changes in λ, q_1, and q_2, which are also accounted for in the structural model, that gives the model its *endogenous-growth* character.

From an environmental perspective, we focus only on emissions of CO_2 as they constitute more than two-thirds of world total GHG emissions, and they are produced in large part by anthropogenic sources. Also, data are more available for these emissions. We assume, that the CO_2 emissions, M, as a byproduct of economic activity of the representative agent, are an increasing function of energy use, E, but a decreasing function of human capital per unit of physical capital. Hence, we adopt the following functional relationship:

$$M=\left(g-\chi(HC/K)^{\zeta}\right)E\ , \tag{15.19}$$

where g is an emissions coefficient, χ a scaling coefficient, and ζ an efficiency parameter.

Energy use is dictated by the static optimality condition, according to which the marginal cost of energy is offset by its marginal product.[6] The cost of energy is measured by the real price of energy, P^e which is exogenously given. Hence, the following relationship holds between energy use and its price:

$$P^e=\partial Y/\partial E=\mu_3\alpha e^{\rho t}HC^{\varepsilon}L^{\mu_1}K^{\mu_2}E^{\mu_3-1}\ . \tag{15.20}$$

To ensure optimality of the solution, the following transversality conditions must also be satisfied:

$$\lim_{t\to\infty}\lambda be^{-\beta t}=0;\quad \lim_{t\to\infty}q_1'Ke^{-\beta t}=0;\quad \lim_{t\to\infty}q_2'HCe^{-\beta t}=0\ . \tag{15.21}$$

The economy of the NAC13 block is modeled with the same assumptions as for the EU15 with corresponding parameters and variables denoted by asterisks.

---

[6] We are assuming consistency between the agent's intertemporal and atemporal optimization.

The present-value Hamiltonian for the NAC13 representative agent's optimization problem is as follows:

$$\begin{aligned} H = {} & \frac{1}{\gamma^*}\left(C^* \cdot MGS^{*\eta_1^*} RMGS^{*\eta_2^*}\right)^{\gamma^*} e^{-\beta^* t} \\ & +\lambda^* e^{-\beta^* t}\left[\begin{array}{l} \left(1-\tau_y^*\right)Y^* + \left(1-\tau_b^*\right) r^* b^* - \left(1+\tau_c^*\right)\left(C^* + \dfrac{MGS^*}{\sigma} + \sigma_r^* RMGS^*\right) \\ -I^*\left(1+\dfrac{a_1^* I^*}{2K^*}\right) - R^*\left(1+\dfrac{a_2^* R^*}{2HC^*}\right) + T^* - \dot{b}^* \end{array}\right] \\ & +q_1'^* e^{-\beta^* t}\left(I^* - \dot{K}^*\right) + q_2'^* e^{-\beta^* t}\left(R^* - \dot{HC}^*\right). \end{aligned} \tag{15.22}$$

As the same approach is taken in deriving the optimality and arbitrage relationships for the NAC13 economy as for the EU15, we will not repeat the exercise. The specification of the emissions equations and transversality conditions are also analogous to those of the EU15.

We assume that both the EU15 and NAC13 face effective interest rates in financial markets that are determined partly by the world interest rate, r^w, and partly by their debt-equity (or credit-equity) positions. The agent's debt-equity (or credit-equity) position is given by the ratio of his/her stock of net foreign assets to physical capital stocks. This assumed relationship of effective interest rate determination can be expressed as follows:

$$r = r^w + \nu(b/K), \tag{15.23}$$

$$r^* = r^w + \nu^*(b^*/K^*). \tag{15.24}$$

For net-borrower regions, one would expect the derivatives $\nu'(b/K)$ and $\nu^{*\prime}\left(b^*/K^*\right)$ be negative, while for net-lender regions these derivatives can take positive or negative signs. This ambivalence arises from the fact that functions (15.24) and (15.25) can represent price-dependent supply functions for net-lender regions when the derivatives are positive. When they are negative, they might represent yield curves for increasingly marginal investments in which r^w is the benchmark yield. We assume that there is a linear relationship between the effective interest rate that agents face and their debt/equity ratios. Algebraically, this relationship is defined as follows:

$$r = r^w + \theta \cdot b/K, \tag{15.25}$$

$$r^* = r^w + \theta^* \cdot b^*/K^*. \tag{15.26}$$

We benchmark the world interest rate, r^w, to the effective United States (U.S.) interest rate. In particular, we assume that the world interest rate is a function of the long-term bond rate in the U.S., r^{US} and the U.S. debt/equity ratio as given below:

$$r^w = r^{US} + \theta^{US} \cdot b^{US}/K^{US}. \tag{15.27}$$

The last specification issue concerns the determination of real exchange rates that agents face in terms of other endogenously determined variables, which is necessary to close the model. We assume that an uncovered interest parity (UIP) relationship holds.[7] Consequently, the proportionate rates of change in real exchange rates are set equal to differences in effective interest rates in the respective economies of the trading regions and in the world. The proportionate rates of change in the three real exchange rates are defined as follows:

$$\begin{aligned}
&\dot{\sigma}/\sigma = r\left(r^{w},b/K\right)-r^{*}\left(r^{w},b^{*}/K^{*}\right),\\
&\dot{\sigma}_r/\sigma_r = r\left(r^{w},b/K\right)-r^{w}\left(r^{US},b^{US}/K^{US}\right), \qquad (15.28)\\
&\dot{\sigma}_r^{*}/\sigma_r^{*} = r^{*}\left(r^{w},b^{*}/K^{*}\right)-r^{w}\left(r^{US},b^{US}/K^{US}\right).
\end{aligned}$$

Further analysis and econometric estimation of the model is most conveniently conducted by defining new variables as in identities (15.29), below, by normalizing with respect to physical capital stocks. These new variables are the EU15 and NAC13 regional rates of consumption of regionally produced and imported goods, stocks of net foreign assets per unit of physical capital, energy and labor inputs per unit of physical capital, units of human capital per unit of physical capital, CO_2 emissions per unit of physical capital and output per unit of physical capital.

$$\begin{aligned}
&c=\frac{C}{K};\ c^{*}=\frac{C^{*}}{K^{*}};\ mgs=\frac{MGS}{K};\ mgs^{*}=\frac{MGS^{*}}{K^{*}};\ rmgs=\frac{RMGS}{K};\\
&rmgs^{*}=\frac{RMGS^{*}}{K^{*}};\ z=\frac{b}{K};\ z^{*}=\frac{b^{*}}{K^{*}};\ n=\frac{E}{K};\ n^{*}=\frac{E^{*}}{K^{*}};\ l=\frac{L}{K};\ l^{*}=\frac{L^{*}}{K^{*}};\\
&h=\frac{HC}{K};\ h^{*}=\frac{HC^{*}}{K^{*}};\ m=\frac{M}{K};\ m^{*}=\frac{M^{*}}{K^{*}};\ y=\frac{Y}{K};\ y^{*}=\frac{Y^{*}}{K^{*}}.
\end{aligned} \qquad (15.29)$$

By differentiating these identities with respect to time, and making use of the optimality and arbitrage conditions we can derive a set of equations that determines the evolution of these economies.[8] Further simplification of the specification can be achieved by defining the following new variables:

$$\begin{aligned}
&k=I/K;\quad k^{*}=I^{*}/K^{*};\quad \kappa=R/HC;\quad \kappa^{*}=R^{*}/HC^{*};\\
&p^{e}=\dot{P}^{e}/P^{e}\ \text{ and }\ p^{e*}=\dot{P}^{e*}/P^{e*}.
\end{aligned} \qquad (15.30)$$

The set of first-order behavioral equations that determine the evolution of $c, mgs, rmgs, z, k, \kappa, n, c^{*}, mgs^{*}, rmgs^{*}, z^{*}, k^{*}, \kappa^{*}, n^{*}, \sigma, \sigma_r, \sigma_r^{*}, y,$ and y^{*} are specified in (15.31) through (15.49) below. The first-order identities deter-

---

[7] As there is no common currency for the NAC13, the UIP assumption is more than one of convenience. It should also be noted that Isard (1995) suggests it tends to fair as well empirically as the covered interest parity (CIP) assumption.

[8] Note that in this specification of the model, and as demonstrated in the Appendix 15.A, the co-state variables have been eliminated by substitution. This is *not* necessary to carry out the estimation of the model but does make the estimation less computationally intensive. And, as discussed below, co-state variables must be present in the specification of the estimated model to impose the tranvsersality conditions in the estimation procedure.

mining h, h^*, K, and K^* are defined in (15.50) through (15.53), while the zero-order relations that determine m, m^*, r, r^*, and r^w are given by (15.54) through (15.58).

$$\dot{c} = c \cdot \left[\frac{1}{(\gamma - 1 + \eta_1\gamma + \eta_2\gamma)} \left(\beta - (1-\tau_b) r + \eta_1 \gamma \frac{\dot{\sigma}}{\sigma} + \eta_2 \gamma \frac{\dot{\sigma}_r}{\sigma_r} \right) - k \right], \tag{15.31}$$

$$m\dot{g}s = \frac{\eta_1}{\sigma}\dot{c} - \eta_1 c \frac{\dot{\sigma}}{\sigma^2}, \tag{15.32}$$

$$r\dot{m}gs = \frac{\eta_2}{\sigma_r}\dot{c} - c\eta_2 \frac{\dot{\sigma}_r}{\sigma_r^2}, \tag{15.33}$$

$$\dot{z} = y + rz - (c + \sigma mgs + \sigma_r rmgs) - k\left(1 + \frac{a_1}{2}k\right) - \kappa h\left(1 + \frac{a_2}{2}\kappa\right) - zk, \tag{15.34}$$

$$\dot{k} = \left(\frac{1}{a_1} + k\right)(1-\tau_b) r - \frac{(1-\tau_y)}{a_1}\mu_2 y - \frac{1}{2}k^2, \tag{15.35}$$

$$\dot{\kappa} = \left(\frac{1}{a_2} + \kappa\right)(1-\tau_b) r - \frac{(1-\tau_y)}{a_2}\frac{y\varepsilon}{h} - \frac{1}{2}\kappa^2, \tag{15.36}$$

$$\dot{n} = \frac{n}{(1-\mu_3)}\left(\rho + \varepsilon\kappa + \mu_1 \frac{\dot{l}}{l} - p^e\right), \tag{15.37}$$

$$\dot{c}^* = c^* \cdot \left[\frac{1}{(\gamma^* - 1 + \eta_1^*\gamma^* + \eta_2^*\gamma^*)} \left(\beta^* - (1-\tau_b^*) r^* - \eta_1^* \gamma^* \frac{\dot{\sigma}}{\sigma} + \eta_2^* \gamma^* \frac{\dot{\sigma}_r^*}{\sigma_r^*} \right) - k^* \right], \tag{15.38}$$

$$m\dot{g}s^* = \dot{c}\eta_1^*\sigma + c\eta_1^*\dot{\sigma}, \tag{15.39}$$

$$r\dot{m}gs^* = \frac{\eta_2^*}{\sigma_r^*}\dot{c} - \eta_2^* c^* \frac{\dot{\sigma}_r^*}{\sigma_r^{*2}}, \tag{15.40}$$

$$\dot{z}^* = y^* + r^* z^* - \left(c^* + \frac{mgs^*}{\sigma} + \sigma_r^* rmgs^*\right) - k^*\left(1 + \frac{a_1^*}{2}k^*\right) - \kappa^* h^*\left(1 + \frac{a_1^*}{2}\kappa^*\right), \\ -z^* k^* \tag{15.41}$$

$$\dot{k}^* = \left(\frac{1}{a_1^*} + k^*\right)(1-\tau_b^*) r^* - \frac{(1-\tau_y^*)}{a_1^*}\mu_2^* y^* - \frac{1}{2}k^{*2}, \tag{15.42}$$

$$\dot{\kappa}^* = \left(\frac{1}{a_2^*} + \kappa^*\right)(1-\tau_b^*) r^* - \frac{(1-\tau_y^*)}{a_2^*}\frac{y^*\varepsilon^*}{h^*} - \frac{1}{2}\kappa^{*2}, \tag{15.43}$$

$$\dot{n}^* = \frac{n^*}{(1-\mu_3^*)}\left(\rho^* + \varepsilon^*\kappa^* + \mu_1^* \frac{\dot{l}^*}{l^*} - p^{*e}\right), \tag{15.44}$$

$$\dot{\sigma} = \sigma\left(r - r^*\right), \tag{15.45}$$

$$\dot{\sigma}_r = \sigma_r\left(r - r^{us}\right), \tag{15.46}$$

$$\dot{\sigma}_r^* = \sigma_r^*\left(r^* - r^{us}\right), \tag{15.47}$$

$$\dot{y} = y\left(\rho + \varepsilon\kappa + \mu_1\frac{\dot{l}}{l} + \mu_3\frac{\dot{n}}{n}\right), \tag{15.48}$$

$$\dot{y}^* = y^*\left(\rho^* + \varepsilon^*\kappa^* + \mu_1^*\frac{\dot{l}^*}{l^*} + \mu_3^*\frac{\dot{n}^*}{n^*}\right), \tag{15.49}$$

$$\dot{h} = h\left(\kappa - k\right), \tag{15.50}$$

$$\dot{h}^* = h^*\left(\kappa^* - k^*\right), \tag{15.51}$$

$$\dot{K} = kK, \tag{15.52}$$

$$\dot{K}^* = k^*K^*, \tag{15.53}$$

$$m = \left(g - \chi h^{\xi}\right)n, \tag{15.54}$$

$$m^* = \left(g^* - \chi^* h^{*\xi^*}\right)n^*, \tag{15.55}$$

$$r = r^w + \theta\cdot z, \tag{15.56}$$

$$r^* = r^w + \theta^*\cdot z^*, \tag{15.57}$$

$$r^w = r^{US} + \theta^{US}\cdot z^{US}. \tag{15.58}$$

Because equations (15.34) and (15.44) fail to capture domestic effects of foreign demands for exports, we introduce domestic supply constraints in the instantaneous budget equations of both the EU15 and NAC13. (See Appendix 15.A for details.) The revised budget equations appearing in the estimated model are, then, as follows:

$$\dot{z} = \tau_y y + \tau_c c + \left(1 + \tau_b\right)rz - \left(1 - \tau_c\right)\left(\sigma mgs + \sigma_r rmgs\right) + \frac{K^*}{K}mgs^* + rxgs - zk, \tag{15.59}$$

$$\dot{z}^* = \tau_y^* y^* + \tau_c^* c^* + \left(1 + \tau_b^*\right)r^* z^* - \left(1 - \tau_c^*\right)\left(\frac{mgs^*}{\sigma} + \sigma_r^* rmgs^*\right) + \frac{K}{K^*}mgs \tag{15.60}$$

$$+rxgs^* - z^*k^*.$$

15.3 Estimation of the Model

To estimate the model, we adopted Wymer's (1997) approach to estimating continuous-time models of intertemporally optimizing agents from discrete observations. This approach enables one to work directly with the complete set of theoretical conditions that characterize a macrodynamic equilibrium solution. Such an equilibrium is attained by solving the state and co-state equations, and taking into account initial endowments, and the transversality conditions that impose the implications of the representative agents' instantaneous budget and capital constraints at every data point. An important feature of this approach is that the intertemporal optimization assumed of the representative agent is incorporated directly into the estimation algorithm. This estimation algorithm entails the following steps:

- For a given set of parameter estimates (or initial values) and estimates of the initial values of the unobserved (co-state) variables whose time paths must satisfy specified (boundary-point and transversality conditions), a Newton shooting algorithm is deployed. With this algorithm, the equilibrium conditions are solved by a variable-order, variable-step Adams method for each data point over a time horizon that (ideally) is sufficiently long for the transversality conditions to be satisfied at an acceptable level of tolerance. For each observation, the solutions must converge. Solutions of the co-state variables are updated at every data point. The residuals are computed for those variables on which there are empirical observations by comparing their dynamic solution values with their observed values and the variance-covariance matrix is formed.

- The natural logarithm of the variance-covariance matrix is minimized by a quasi-Newton method to update parameter estimates. Convergence criteria are then checked, and if not met, another iteration is begun.

Imposing the transversality conditions on the solution and estimation of the model presents a problem, since such conditions embody a limit concept. In some cases, one might reasonably assume that they are satisfied—e.g., if one is confident that the absolute values of the estimates of the agents' subjective discount rates, β and β^* exceed the growth rates of the levels of consumption and the capital stocks. But by taking the conditions, which in the limit must converge to zero, sufficiently far out in time, so that they satisfy a reasonably small tolerance level, one can gain information internal to the estimation process needed to impose a boundary point condition on the trajectories of the unobservable costate variables, and, in turn, estimates of the model parameters. (See e.g. Judd 1998).

Although the first-order equations for costate variables q_1, q_2, q_1^*, and q_2^* have been replaced by equations in $k, k^*, \kappa,$ and κ^*, the costate equations determining the levels of λ and λ^* must be included in the estimated model in order for transversality conditions to be imposed upon the solution of the model. From the arbitrage relationship (15.16) for the EU15 and the analogous one for the NAC13, these are

$$\dot{\lambda} = \lambda\left[\beta - r(z)(1-\tau_b)\right], \tag{15.61}$$

$$\dot{\lambda}^* = \lambda^*\left[\beta^* - r^*\left(z^*\right)\left(1-\tau_b^*\right)\right]. \tag{15.62}$$

The zero-order transversality conditions for the EU15, (15.21), and the analogous expressions for the NAC13 can be rewritten in terms of the model's variables[9] (and collected in single expressions) as follows:

$$\lim_{t\to\infty} \lambda e^{-\beta t}\left(z+1+a_1k+h+a_2\kappa h\right)=0\,, \tag{15.63}$$

$$\lim_{t\to\infty} \lambda^* e^{-\beta^* t}\left(z^* +1+a_1^*k^* +h^* +a_2^*\kappa^* h^*\right)=0\,. \tag{15.64}$$

To estimate the model, we have employed quarterly data published mostly at the country level from 1992:1 to 2001:4. While some data are published for the EU15 as an aggregate, there are no aggregate data for the NAC13 bloc. Thus data on macroeconomic variables and price indices representing the aggregate EU15 and NAC13 economies must be constructed. The time span of the data is dictated by the availability of data for the NAC13 bloc. Since there is no common currency for the NAC13 bloc, it is necessary to create an exchange rate regime to aggregate up macroeconomic quantities for the bloc. Similar to the ECU, the synthetic currency that existed before introduction of the euro, we have developed a normalized trade-weighted real exchange rate, which we have called the NACCU (an acronym for New Accession Country Currency Unit).[10]

Exogenous variables also require special treatment, because the solution algorithm entails solving the model over a time horizon beyond the period for which observations are available. Hence, a zero-order forcing function of time is introduced for each exogenous variable other than time. These forcing functions of time are estimated separately from the structural model, assuming that the values of their coefficients should be determined independently of the model's parameters. All the aggregate time series were converted to real terms and deseasonalized around the trend, following Durbin (1963).[11] Appendix B presents the details of the data aggregation procedure.

To obtain suitable initial values of the parameter estimates, we estimated the model first using a one-period-forward solution approach without the co-state equations and transversality conditions. In this and subsequent estimation involving the dynamic-boundary-point solution approach and all theoretical conditions the model's solution must satisfy, we imposed constraints on the values that several parameters could assume. The parameters of the intertemporal utility and production functions are constrained to satisfy the curvature and homogeneity conditions assumed. The tax parameters are allowed to lie between 0.0 and 0.4, their historical range, and the cost of adjustment parameters not to exceed 23.50, a magnitude suggested by Barro and Sala-i-Martin's (1995) discussion of quadratic adjustment costs and practical experience.

The model was estimated consistent with these theoretical and empirical assumptions about the representative agents. However, preliminary investigations of policy simulations that introduce changes in the specification of key equations

---

[9] See Appendix 15.A for details.

[10] For details of data construction for each variable at the bloc level, please see Appendix 15.B.

[11] Wymer's program TRANSF was used to prepare the data for estimation. All estimation work was carried out with his program ESCONA.

have lead to unstable solutions from which not much can be learned. Thus other limits on the values of influential parameters have been experimented with in order to promote stability in policy simulations.[12] For this purpose, we have bounded the values that estimates of $\theta, \theta^*, \theta^{US}, \rho, \rho^*, \varepsilon$ and ε^* can assume. The parameters θ, θ^* and θ^{US} are arguments of budget, and physical and human capital formation equations as presented in equations (15.35), (15.36), (15.42) and (15.43) through r, r^* and r^{US}. The parameters ε and ε^* also affect human capital formation directly. The output-capital ratios, y and y^*, are expected to vary in a small range based on stylized facts of the developed economies. However, as can be seen in (15.48) and (15.49), this ratio depends on the values of $\rho, \rho^*, \varepsilon$ and ε^* which also affect n and n^* directly. Consequently, the values of these parameters were bounded to prevent them from becoming very large, and causing capital/output ratios to reach unreasonable values.

In the estimation, the representative agents in the model were assumed to adopt a relative optimization horizon of 11.5 years, i.e., 46 quarters. (This was the longest relative horizon over which the model could be solved for every data point in the estimation work.) The estimation has yielded a reasonably good fit of the model to the historical time series. Table 15.1 presents the means and standard deviations of the endogenous variables and the root-mean-square-errors (RMSEs) of the in-sample dynamic forecasts produced by the model. Proportionate RMSEs of the variables range from 0.00007 to 0.11399. Because of the dearth of data, we did not hold out observations for out-of-sample fit assessment.

The fit of the model for several of the variables is also exhibited in plots below. As can be observed in Figures 15.1 and 15.2, the estimated model tracks well the debt/equity (credit/equity) ratios of the both the EU15 and the NAC13, which were volatile during the estimation period and, as determined by residual through the budget constratints, are difficult to model.

Proportionate growth rates of physical and human capital stocks also fluctuated greatly between 1992:1 and 2001:4. While in the EU15 the proportionate rate of growth in the physical capital stock was lower than that of the human capital stock until the late 1990s, after 1999 the physical capital stock grew faster. In the NAC13, both capital and human capital formation proceeded at a brisker pace than in the EU15, with the former growth rate nearly reaching parity with the latter at the end of the decade after being nearly half its value. As portrayed in Figures 15.3 and 15.4, the estimated model tracks these dynamics for the EU15, while manifesting some systematic variation in the in-sample forecast error.

The estimated and observed values of the EU15's emissions (per unit of physical capital) are plotted in Figure 15.5 where as those for the NAC13 are in Figure

---

[12] One of the features of the model under discussion—and indeed the entire genre of models based on the assumption of intertemporally optimizing representative agents—is the 'knife's edge' character of its (saddle-path) solution. At this stage it is not known to what extent policy analyses can be conducted with such a model based on empirical data under different parameterizations. Microfoundations of this ilk apparently come with a price.

15.6. The estimated model catches the increase and then decline in EU15 emissions over the last part of the decade and the more steadily declining trend in NAC13 emissions over the entire period.

Table 15.1. In-Sample Forecasting Errors of the Estimated Model

	Endogenous Variables y(t)						
	Observed		Estimated		Error in estimated y(t)		Proportionate
	Mean	Std. Dev'n	Mean	Std. Dev'n	Mean	Std. Dev'n	RMSE
c	0.01893	0.00091	0.01889	0.00088	0.00004	0.00018	0.00956
mgs	0.00035	0.00008	0.00032	0.00008	0.00003	0.00003	0.08782
$rmgs$	0.00828	0.00069	0.00829	0.00068	-0.00001	0.00033	0.04024
z	0.01004	0.00456	0.01016	0.00454	-0.00012	0.00107	0.10690
k	0.00253	0.00038	0.00235	0.00038	0.00017	0.00010	0.03921
κ	0.00272	0.00014	0.00252	0.00014	0.00021	0.00005	0.01911
n	0.00654	0.00007	0.00651	0.00008	0.00003	0.00005	0.00764
c^*	0.00707	0.00028	0.00695	0.00026	0.00013	0.00023	0.03295
mgs^*	0.00263	0.00082	0.00256	0.00083	0.00007	0.00018	0.06923
$rmgs^*$	0.00169	0.00031	0.00164	0.00031	0.00006	0.00008	0.04900
z^*	0.00490	0.00160	0.00434	0.00196	0.00057	0.00054	0.11011
k^*	0.00423	0.00076	0.00393	0.00080	0.00030	0.00023	0.05420
κ^*	0.00514	0.00030	0.00502	0.00027	0.00012	0.00016	0.03152
n^*	0.00783	0.00052	0.00837	0.00061	-0.00055	0.00014	0.01738
σ	1.32458	0.43252	1.29860	0.43835	0.02598	0.08397	0.06339
σ_r	1.17703	0.15397	1.16540	0.14807	0.01163	0.03519	0.02990
σ_r^*	1.03927	0.17445	1.04224	0.18191	-0.00298	0.06774	0.06518
y	0.03249	0.00166	0.03261	0.00161	-0.00012	0.00029	0.00883
y^*	0.01086	0.00039	0.01090	0.00039	-0.00005	0.00032	0.02938
m	0.00330	0.00007	0.00329	0.00008	0.00001	0.00002	0.00576
m^*	0.00498	0.00036	0.00532	0.00042	-0.00034	0.00009	0.01888
K	5.38816	0.14721	5.38771	0.14718	0.00046	0.00044	0.00008
K^*	1.08352	0.05278	1.08334	0.05279	0.00018	0.00015	0.00014
h	0.05703	0.00016	0.05703	0.00016	0.00000	0.00000	0.00007
h^*	0.00377	0.00003	0.00377	0.00003	0.00000	0.00000	0.00027

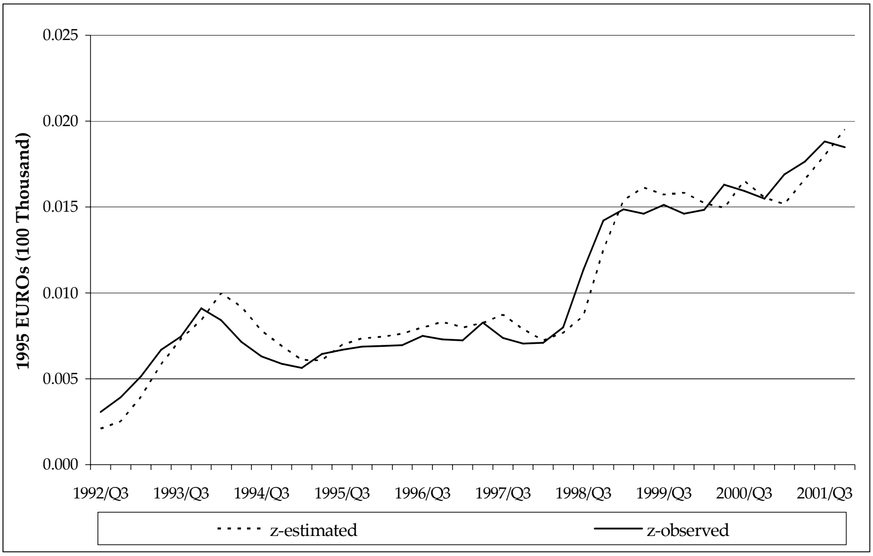


Fig. 15.1. Estimate of the EU15 net foreign assets, 1992-2001.

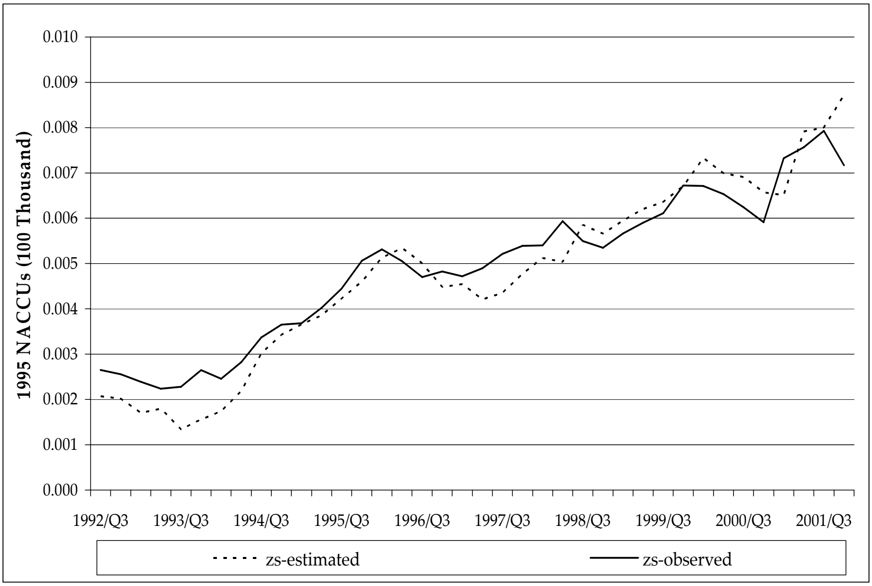


Figure 15.2. Estimates of the NAC13 net foreign assets, 1992-2001.

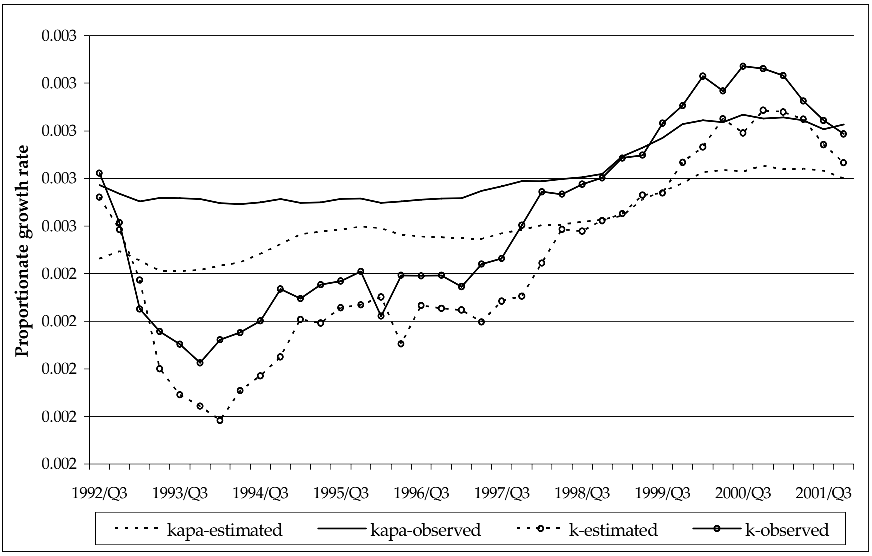


Figure 15.3. Estimates of the EU15 physical and human capital growth rates, 1992-2001

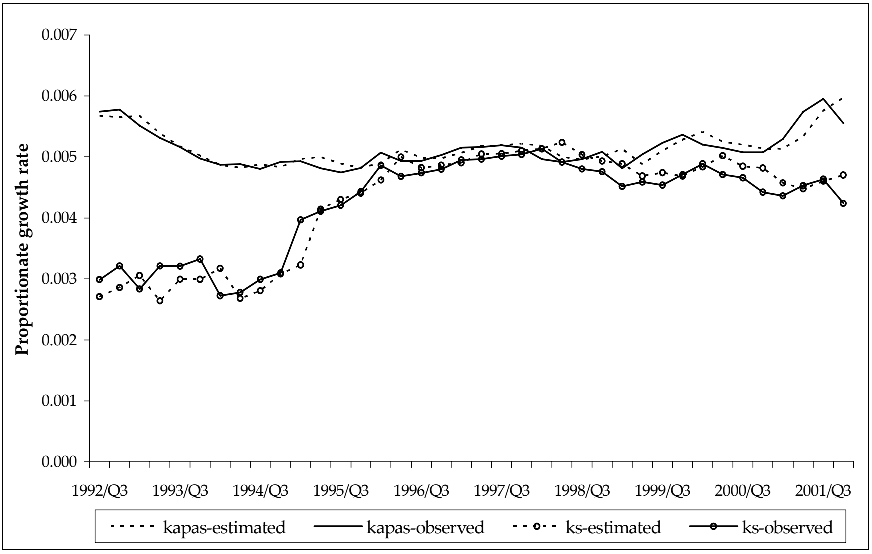


Fig.15.4. Estimates of the NAC13 physical and human capital growth rates, 1992-2001.

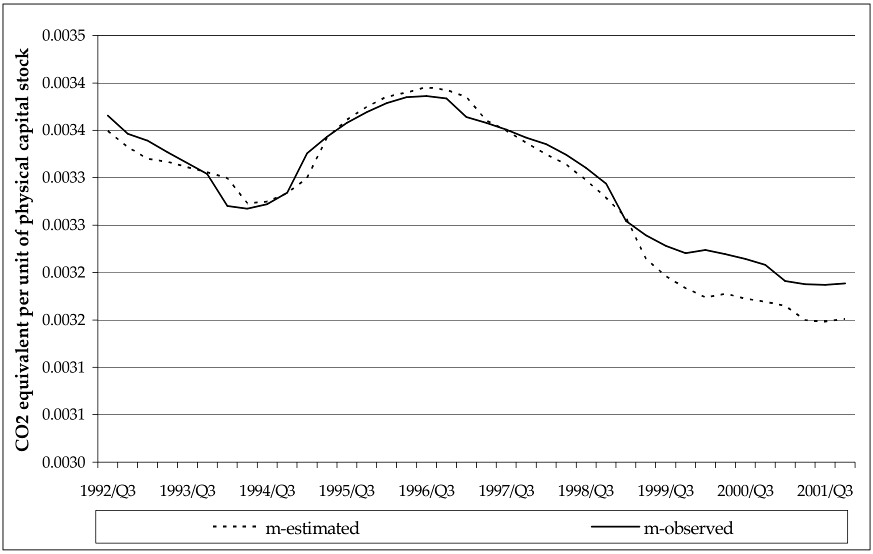


Fig.15.5. Estimates of EU15 emissions, 1992-2001.

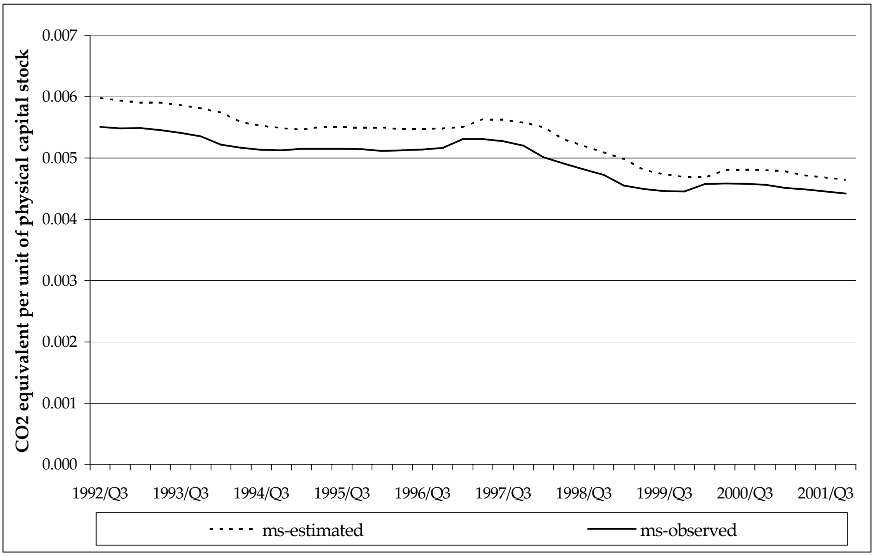


Fig. 15.6. Estimates of NAC13 Emissions, 1992-2001.

With an eye toward the discussion of the effects of trade on emissions in the next section, we offer some historical observations about trade between the blocs.

The relative importance to the EU15 and NAC13 of imported goods and services from the other bloc was markedly different over the sample period. Whereas the NAC13 purchased more than half of its total imported goods from the EU15, the EU15 purchased a percentage in single digits. This amount did increase steadily, however, from 2% in 1992 to 6% in 2001. Figure 15.7 portrays the evolution of each bloc's share of total imports accounted for by purchases from the other bloc from 1992:1 to 2001:4.

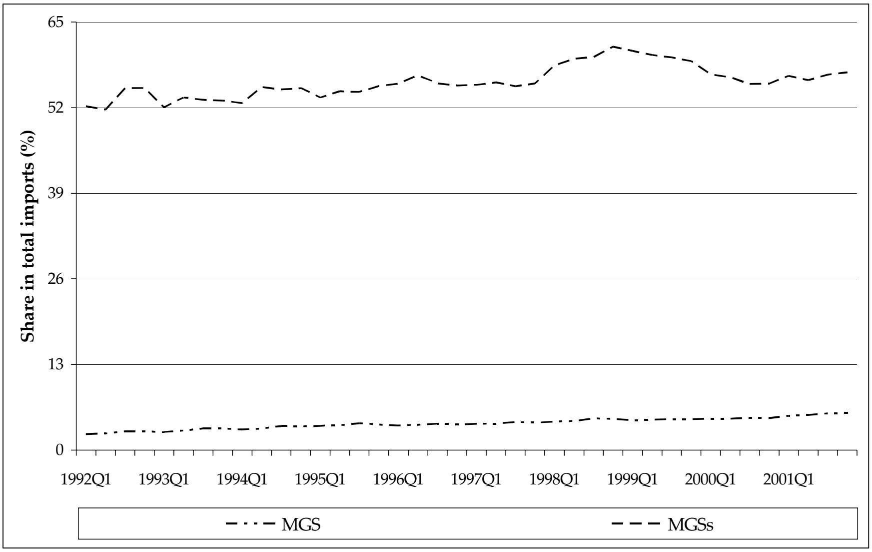


Fig.15.7. The relative importance of imports from the other bloc, 1992-2001.

From both Table 15.1 and Figures 15.1 through 15.6 we can infer that the solution of a two-bloc model based on the micro-behavioral assumptions of representative agents is generally consistent with the historical data. Of course, other models based on other assumptions might also be consistent with the data and a full and fair assessment of the present model would entail comparison with results of other models purporting to render an account of the same empirical phenomena. But such a comparison is beyond the scope of this study.

The estimates of the parameters and their asymptotic standard errors (ASE) are given in Table 15.2. All estimates are statistically discernible from zero at the 0.05 level of significance except τ_y and τ_c, taxes on income and consumption in the EU15. The significant estimates of ε and ε^* indicate the presence of increasing returns to scale in both economies, while the magnitude of the estimate is larger in the case of the EU15 than the NAC13. As expected, the estimate of the scale parameter for the EU15 production function, α, is larger than that for the

NAC13, α^*. The estimated rate of technological progress, ρ is also higher in the EU15 than in the NAC13.

The economies of the two blocs are characterized by different output elasticities of factors of production. In the EU15, the output elasticity of labor is higher than that of the capital, while the opposite holds in the NAC13 case. The estimates of γ and γ^* imply disparate intertemporal elasticities of substitution for the EU15 and the NAC13 of 1.123 and 1.176. The temporal discount rate of the EU15 representative agent is slightly lower than that of the NAC13 agent. As the estimate of β is at its lower bound, it is reasonable to expect that if a smaller bound was allowed, the estimate might have been lower. Both agents would appear to derive more utility from the goods imported from the rest of the world than from the goods they import from each other.

Table 15.2 Quasi-FIML Estimates of Parameters and their Asymptotic Standard Errors

Parameter	Parameter Estimate	Estimate of ASE	Parameter	Parameter Estimate	Estimate of ASE
α	0.06349	0.00096	α^*	0.01196	0.00010
ρ	0.00710	0.00001	ρ^*	0.00500	0.00001
ε	0.02000	0.00001	ε^*	0.02500	0.00001
μ_1	0.63000	0.00001	μ_1^*	0.38000	0.00008
μ_2	0.35000	0.00001	μ_2^*	0.60000	0.00008
μ_3	0.02000	0.00001	μ_3^*	0.02000	0.00001
η_1	0.40022	0.00443	η_1^*	0.08000	0.00011
η_2	0.99931	0.00007	η_2^*	0.59998	0.00001
γ	0.10957	0.00024	γ^*	0.14990	0.00005
β	0.00500	0.00001	β^*	0.00581	0.00107
a_1	23.49756	0.00110	a_1^*	10.50000	0.00471
a_2	20.49921	0.00035	a_2^*	21.49967	0.00006
τ_b	0.40000	0.00001	τ_b^*	0.39998	0.00001
τ_y	0.00000	0.00000	τ_y^*	0.00002	0.00001
τ_c	0.00000	0.00000	τ_c^*	0.07433	0.05225
θ	0.00700	0.00003	θ^*	0.01533	0.00004
g	9.09424	1.56488	g^*	5.63698	0.99479
χ	20.31805	3.70174	χ^*	22.14662	4.40495
ς	0.30061	0.00007	ς^*	0.26664	0.00011
θ^{US}	0.00638	0.00055			

15.4 Simulations of Changes in Trade Patterns

In this section, we analyze potential effects of increasing trade between the EU15 and the NAC13. In doing so, we solve the estimated model over 40 quarters, i. e. 10 years, out of sample from 2002:1 to 2012:4, unless otherwise stated, with the initial values of the endogenous variables taken from the last data point of 2001:4. Quarterly results were obtained with Wymer's program APREDIC by using the dynamic simulation algorithm. The solutions obtained are in terms of physical capital units, in which the data were denominated. As one can not confidently guess the future behavior of the exogenous variables, we made certain assumptions about their values.

We fixed the values of the world interest rate and the U.S. net foreign assets to their values at the last data point, 2001:4. We also set the price of energy to its value in 2001:4. We assumed, perhaps unrealistically, that the labor force of each bloc will grow by 0.8% quarterly and that exports of each bloc to the rest of the world will increase by 1% quarterly. For completeness, solution paths of all endogenous variables are presented, even though we will only highlight the paths of variables of particular interest. For purposes of presentation, we group variables with similar values.

Both the historical trends in Figure 15.7 and historical increases in the amount of trade within the EU following member states' joining indicate that it is reasonable to assume trade will increase between the EU15 and NAC13. To examine potential impacts of increasing trade between these blocs on their emission levels with no policy intervention, we perform a comparative dynamics analysis.

Increases in trade might take place in three different ways, among others. First, the volume of trade might increase steadily, perhaps through trade agreements. Second, the preferences of the representative agent of one country bloc for goods from the other bloc may change in favor of these goods, leading to a reduction in demand for goods from the rest of the world, a so-called trade diversion effect. Such changes in trade might occur unilaterally or bilaterally. A third way in which increases in trade might come about would be through changes in the effective exchange rate between the two blocs. We consider the first two scenarios here.[13]

15.4.1 Effects of Steady Increases in Trade

The specifications of the imports equations can be altered as follows to reflect steady increases in trade by fixed amounts, Δ and Δ^* :

$$m\dot{g}s = \frac{\eta_1}{\sigma}\dot{c} - \eta_1 c\frac{\dot{\sigma}}{\sigma^2} + \Delta\,, \tag{15.65}$$

---

[13] See Balta (2004) for a discussion of the third alternative.

$$m\dot{g}s^* = \sigma\eta_1^* \dot{c}^* + c\eta_1^* \dot{\sigma} + \Delta^* . \tag{15.66}$$

We consider four cases, reflecting alternative combinations of values of Δ and Δ^*. The first case is the baseline or 'business-as-usual' solution of the model as estimated. In Case 2, both the EU15 and the NAC13 increase the amount of goods they import from each other by 2% over the 2001:4 level. In Case 3, only the EU15 increases its imports, mgs, by 2% over the 2001:4 level, and mgs^* is left as business-as-usual. In the last case, Case 4, only the NAC13 increases its imports from the EU15 by 2% over the 2001:4 level.

Table 15.3. Mean values of the variables of the model with increasing trade scenarios (2002:1-2014:2)

	Means				Ratio to BAU			Change from BAU		
	Case 1	Case 2	Case 3	Case 4	Case 2	Case 3	Case 4	Case 2	Case 3	Case 4
c	0.02069	0.02087	0.02088	0.02069	1.00863	1.00895	0.99962	+	+	-
mgs	0.00054	0.00084	0.00086	0.00053	1.54030	1.57348	0.96765	+	+	-
$rmgs$	0.00781	0.00793	0.00795	0.00780	1.01525	1.01729	0.99790	+	+	-
z	0.15929	0.15522	0.14422	0.17035	0.97446	0.90542	1.06942	-	-	+
c^*	0.00582	0.00527	0.00582	0.00527	0.90466	1.00030	0.90501	-	+	-
mgs^*	0.00345	0.00524	0.00345	0.00524	1.51967	0.99890	1.52069	+	-	+
$rmgs^*$	0.00093	0.00063	0.00092	0.00064	0.67250	0.98872	0.68322	-	-	-
z^*	0.05328	0.06224	0.07687	0.03882	1.16816	1.44277	0.72871	+	+	-
σ	1.91131	1.90783	1.90083	1.91836	0.99818	0.99451	1.00369	-	-	+
σ_r	1.50906	1.50865	1.50736	1.51036	0.99973	0.99887	1.00086	-	-	+
σ_r^*	1.08586	1.08758	1.09075	1.08275	1.00158	1.00450	0.99714	+	+	-
n	0.00829	0.00829	0.00829	0.00829	1.00004	1.00020	0.99985	+	+	-
n^*	0.00816	0.00816	0.00815	0.00817	0.99961	0.99884	1.00077	-	-	+
m	0.00451	0.00451	0.00451	0.00451	1.00010	1.00045	0.99965	+	+	-
m^*	0.00502	0.00502	0.00503	0.00501	1.00070	1.00210	0.99860	+	+	-
y	0.04424	0.04424	0.04425	0.04423	1.00004	1.00020	0.99985	+	+	-
y^*	0.01337	0.01337	0.01336	0.01338	0.99961	0.99884	1.00077	-	-	+
U	0.89070	0.89137	0.89145	0.89062	1.00076	1.00085	0.99991	+	+	-
U^*	0.33859	0.33449	0.33867	0.33440	0.98789	1.00023	0.98763	-	+	-
K	5.49212	5.49183	5.49081	5.49315	0.99995	0.99976	1.00019	-	-	+
K^*	1.35033	1.35168	1.35439	1.34766	1.00100	1.00300	0.99803	+	+	-
HC	0.30850	0.30848	0.30842	0.30856	0.99994	0.99973	1.00021	-	-	+
HC^*	0.00522	0.00522	0.00523	0.00522	1.00051	1.00152	0.99900	+	+	-

Table 15.3 presents the mean solution values of all endogenous variables in the model in comparison to the 'business-as-usual' (BAU) solution. In these simulations, the instantaneous levels of utility the EU15 and NAC13 agents derive from consumption of domestic and imported goods are also explicitly computed, and

they are denoted respectively by U and U^*. The results for physical and human capital, K, K^* and HC, HC^*, are in trillions of 1995 EUROs and NACCUs. The results for the other variables are in terms of physical capital units. The first four columns exhibit the means of the variables under these scenarios. The second part of the table compares the results of these scenarios to the BAU results. In the last part of the table, the results are presented as direction of changes from the BAU case.

From the solutions of the model in these scenarios and the specification of the model, one can make the following observations and inferences.

1. If the representative agents were to increase the amounts of goods they import from each other, their credit/equity positions (z) would be affected through their budget constraints. The size of the effect would be larger in the case of the NAC13, however. These changes would in turn affect the real exchange rate between the blocs, inducing further changes in their consumption patterns. Changes in debt/equity positions would also affect the formation of physical and human capital in both blocs, ratios of human to physical capital, and hence emissions.

2. If a bilateral increase in trade were to occur, it would affect the EU15 and NAC13 economies differently. With a bilateral increase in imports of 2% above 2001:4 levels, the credit/equity ratio of the EU15 would decrease, while that of the NAC13 would increase. These changes would in turn lower the EURO/NACCU exchange rate (σ), making exports from the EU15 less expensive than domestic goods in the NAC13. Consequently, the EU15 would increase its exports to the NAC13 (mgs^*) more than its imports from both from the NAC13 (mgs). Since, as a consequence of the increase in bilateral trade, the NACCU would appreciate against the US Dollar (σ_r^*) less than it would against the EURO, the NAC13 would increase its imports from the EU15 (mgs^*) at the expense of its imports from the rest of the world ($rmgs^*$). As human capital per unit of physical capital would decrease in both the EU15 and NAC13, emissions would increase. Even though the NAC13 would ostensibly increase its human and physical capital stocks to meet foreign sales demand, increases in physical capital would exceed those in human capital, hence the ratio of human capital to physical capital would decline.

3. If the EU15 were to increase its imports from the NAC13 unilaterally, a picture similar to the previous case emerges with some differences. Domestic consumption of the NAC13 would increase and imports from the EU15 would decrease in comparison to the BAU case, as there would be no income transfer through trade. In order to purchase more goods from the NAC13, the EU15 would have to increase both its energy use and output. As human capital per unit of physical capital decreases in both blocs, emissions would increase.

4. If the NAC13 were to increase its imports from the EU15 unilaterally, it would consume less of its domestically produced goods and less imports from the rest of the world. The credit/equity position of the EU15 would improve and that of the NAC13 would decline. Consequently, the EURO would appreciate—albeit, marginally—against the NACCU and the US dollar, inducing a decline in EU15 imports by the NAC13 and the rest of the world. Changes in credit/equity positions would also affect the effective interest rate faced in each bloc, the combined effect of which would be to induce increases in the ratios of human capital to physical capital—h and h^*—and hence small decreases in emissions per unit of capital relative to the BAU case.

5. Although not directly comparable, the effects of increased trade on utility levels that the EU15 and NAC13 agents derive under these scenarios are also not symmetric. Solutions of the model suggest that a bilateral increase in trade (case 2) would lead to a welfare improvement for the EU15 but not the NAC13 and that a unilateral increase in the imports by the NAC13 from the EU15 (case 4) would lead to a deterioration in welfare for both blocs. A unilateral increase in EU15 purchases from the NAC13, however, would result in a ‘win-win’ situation.

Overall, the exercise indicates that the impacts of increasing trade would depend on the direction of trade and whether changes are unilateral or bilateral. Under bilateral increases in trade, output and emissions of the EU15 would increase, indicating the presence of a so-called size effect. However, this does not lead to composition nor technology effects in the EU15. If the EU15 upped its imports from the NAC13 unilaterally, the same effects would prevail. In these cases, due to changes in exchange market, the NAC13 would reduce its output as its consumption patterns change. If, however, the NAC13 were to increase its imports from the EU15, these foreign sales would induce the EU15 to increase its stocks of human and physical capital. Hence, the so-called technology and composition effects would be evident in the EU15 economy. The findings of the exercise also suggest that reductions in aggregate emissions are not likely to be coincident with increases in trade if there are not associated structural changes in preferences or technology in the trading blocs, to whose consideration we now turn.

15.4.2. Effects of Changes of Preferences

With the integration of the NAC13 countries into the larger economy of the European Union, one might expect both blocs of countries considered to develop preferences for each other’s goods. It has been observed that as the pre-2004 EU expanded, member states would purchase more from each other and less from the rest of the world. A similar phenomenon has been documented for the North American Free Trade Agreement as well. (See Romalis 2004, *inter alia.*) In the trade literature, this so-called trade diversion is explained by the absence of tariffs between trade partners (Romalis 2004). However, in the European case, this diversion might be motivated by the increasing consumer preferences for European brands as well as by the unimpeded movement of goods.

In order to mimic such a phenomenon, we will increase the values of the parameters of the agents' utility functions associated with products from the other blocs, η_1 and η_1^* and lower the values of parameters of associated with products from the rest of the world, η_2 and η_2^*. However, since we do not wish to effect arbitrary changes in utility levels, we will adjust the values of the parameters in question so as to keep the agents welfare-neutral—i.e., keep measures of derived utility constant for sample averages of goods traded between the blocs and between the blocs and the rest of the world.

Considering the case of the EU15, we note that the isoelastic intertemporal utility function, which is assumed to represent utility the EU15 derives from consumption of domestically produced and imported goods, has been defined as follows:

$$\int_0^\infty \frac{1}{\gamma}\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^\gamma e^{-\beta t} dt . \tag{15.67}$$

This utility function is rewritten in terms of the variables used in the estimation of the model below where $mgs = MGS/K$ and $rmgs = RMGS/K$.

$$\int_0^\infty \frac{1}{\gamma}\left(c \cdot mgs^{\eta_1} rmgs^{\eta_2} K^{1+\eta_1+\eta_2}\right)^\gamma e^{-\beta t} dt . \tag{15.68}$$

In order to introduce change in preference to consume more of mgs at the expense of $rmgs$, the following ratio is assumed to hold where $M\bar{G}S$, $R\bar{M}GS$ and $\bar{K}$ are the sample averages and η_1' and η_2' are new parameter values.

$$\frac{M\bar{G}S^{\eta_1} R\bar{M}GS^{\eta_2}}{\bar{K}^{\eta_1+\eta_2}} = \frac{M\bar{G}S^{\eta_1'} R\bar{M}GS^{\eta_2'}}{\bar{K}^{\eta_1'+\eta_2'}} . \tag{15.69}$$

An analogous assumption is made in the case of the NAC13 economy.

For this exercise, two different alternatives for the values of η_1', η_2', $\eta_1'^*$, and $\eta_2'^*$ are considered. In the baseline case (Case 1), the values of η_1, η_2, η_1^*, and η_2^* are the estimates obtained, 0.40, 0.99, 0.08 and 0.59. In Case 2, a bilateral change in preferences in analyzed and the values of η_1', η_2', $\eta_1'^*$, and $\eta_2'^*$ are set equal to 0.43, 0.95, 0.086 and 0.594. In Case 3, a preference change is introduced only in the EU15 and $\eta_1'^*$, and $\eta_2'^*$ are set equal to their values in the baseline case. Similarly, in Case 4, the EU15 parameters are kept as in the baseline case and the NAC13 is allowed to change its trade preferences. In Case 5, the values of η_1', η_2', $\eta_1'^*$, and $\eta_2'^*$ are set equal to 0.47, 0.89, 0.093 and 0.587, assuming that trade diversion will take place mutually in both blocs. In Cases 6 and 7 unilateral changes in trade patterns are introduced.

The results are presented in Table 15.4 for mean solution values of the endogenous variables in the model in comparison to the BAU. The results are presented in the same units as in the previous section. While U and U^* denote utilities the agents derive from consumption of domestic and imported goods, the results for physical and human capital, K, K^*, HC and HC^*, are in trillions of 1995 EUROs and NACCUs. The results for the rest of the variables are in terms of physical capital units.

The first part of Table 15.4 presents the means of the variables under these scenarios. The second part of the table compares the results of these scenarios to the BAU results. In the last part of the table, the results are presented as direction of changes from the BAU case.

Changes in the values of η_1, η_2, η_1^*, and η_2^* affect not only the imports equations, but also consumption equations, through which debt/equity and credit/equity positions of the economies are influenced. Through these variables the effects are transmitted to other variables in the model. From the solutions of the model in these scenarios and inspection of the model's specification, one can make the following observations and inferences.

1. Even though the effects of changing these parameters bring about results at different magnitudes, direction of changes are indifferent to both the values of these parameters and the direction of trade, if changes are unilateral or bilateral. Some results change in Cases 3 and 7.

2. Any preference change leading to increased trade between the two blocs would result in an increase in the credit/equity ratio of the EU15 and a reduction in that of the NAC13. These changes would be accompanied by higher levels of physical and human capital stocks, and, in the case of the EU15, a higher ratio of human capital to physical capital. Even though the levels of capital stocks in the NAC13 decrease, human capital per unit of physical capital stock would increase as decreases in human capital stock are lower than those in physical capital stocks. Consequently, *both economies would experience lower emissions*, although reductions would be marginal *vis a vis* the BAU case.

3. Changes in credit/equity positions would also affect exchange rates. With increasing z and decreasing z^*, the EURO would appreciate against the NACCU and marginally against the US dollar. On the other hand, NACCU would depreciate against US Dollar.

Table 15.4. Mean values of the variables of the model under structural change in trade scenarios (2002:1-2014:2)

	Averages							Ratio to BAU						Change from BAU					
	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
c	0.02069	0.02067	0.02067	0.02069	0.02064	0.02064	0.02069	0.99882	0.99885	0.99997	0.99748	0.99754	0.99995	-	-	-	-	-	-
mgs	0.00054	0.00054	0.00054	0.00054	0.00054	0.00054	0.00054	0.99560	0.99778	0.99797	0.99011	0.99491	0.99593	-	-	-	-	-	-
$rmgs$	0.00781	0.00780	0.00780	0.00781	0.00779	0.00780	0.00781	0.99874	0.99878	0.99995	0.99763	0.99772	0.99991	-	-	-	-	-	-
z	0.15929	0.15951	0.15948	0.15932	0.15970	0.15964	0.15934	1.00139	1.00121	1.00017	1.00258	1.00219	1.00034	+	+	+	+	+	+
c^*	0.00582	0.00593	0.00582	0.00593	0.00604	0.00582	0.00604	1.01866	1.00000	1.01866	1.03738	0.99999	1.03739	+	-	+	+	-	+
mgs^*	0.00345	0.00345	0.00345	0.00345	0.00345	0.00345	0.00345	1.00015	1.00000	1.00014	1.00033	1.00000	1.00032	+	+	+	+	+	+
$rmgs^*$	0.00093	0.00100	0.00093	0.00100	0.00107	0.00093	0.00107	1.07377	1.00001	1.07376	1.14805	1.00002	1.14802	+	+	+	+	+	+
z^*	0.05328	0.05197	0.05323	0.05203	0.05064	0.05317	0.05077	0.97553	0.99907	0.97651	0.95049	0.99790	0.95288	-	-	-	-	-	-
σ	1.91131	1.91178	1.91134	1.91175	1.91225	1.91137	1.91218	1.00024	1.00002	1.00023	1.00049	1.00003	1.00046	+	+	+	+	+	+
σ_r	1.50906	1.50908	1.50908	1.50906	1.50909	1.50908	1.50907	1.00001	1.00001	1.00000	1.00002	1.00002	1.00000	+	+	+	+	+	+
σ_r^*	1.08586	1.08560	1.08585	1.08561	1.08534	1.08584	1.08536	0.99976	0.99999	0.99977	0.99952	0.99998	0.99954	-	-	-	-	-	-
n	0.00829	0.00829	0.00829	0.00829	0.00829	0.00829	0.00829	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	-	-	-	-	-	-
n^*	0.00816	0.00816	0.00816	0.00816	0.00816	0.00816	0.00816	1.00006	1.00000	1.00006	1.00012	1.00000	1.00012	+	+	+	+	+	+
m	0.00451	0.00451	0.00451	0.00451	0.00451	0.00451	0.00451	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	-	-	-	-	-	-
m^*	0.00502	0.00502	0.00502	0.00502	0.00502	0.00502	0.00502	0.99989	1.00000	0.99990	0.99978	0.99999	0.99979	-	-	-	-	-	-
y	0.04424	0.04424	0.04424	0.04424	0.04424	0.04424	0.04424	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	-	-	-	-	-	-
y^*	0.01337	0.01337	0.01337	0.01337	0.01338	0.01337	0.01338	1.00006	1.00000	1.00006	1.00012	1.00000	1.00012	+	+	+	+	+	+
U	0.89070	0.89039	0.89039	0.89069	0.88969	0.88970	0.89069	0.99965	0.99966	1.00000	0.99887	0.99888	0.99999	-	-	-	-	-	-
U^*	0.33859	0.33917	0.33859	0.33917	0.33971	0.33859	0.33971	1.00171	1.00000	1.00171	1.00330	1.00000	1.00330	+	-	+	+	-	+
K	5.49212	5.49213	5.49213	5.49212	5.49213	5.49213	5.49212	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	+	+	+	+	+	+
K^*	1.35033	1.35012	1.35033	1.35013	1.34992	1.35032	1.34993	0.99985	1.00000	0.99985	0.99969	0.99999	0.99970	-	-	-	-	-	-
HC	0.30850	0.30850	0.30850	0.30850	0.30850	0.30850	0.30850	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	+	+	+	+	+	+
HC^*	0.00522	0.00522	0.00522	0.00522	0.00522	0.00522	0.00522	0.99992	1.00000	0.99992	0.99984	1.00000	0.99985	-	-	-	-	-	-

4. The model indicates that for all changes in preferences considered the EU15 would lower its consumption of both domestic and imported goods. As these parameters enter the model non-linearly, a plausible explanation why this should be the case is difficult to make out. On the other hand, the model indicates that for some changes in preferences consumption of domestic goods by the NAC13 would increase and for all changes in preference consumption of imported goods would increase. As it is consuming more than in the BAU case, the representative agent of the NAC13 would derive more utility.

15.5 Concluding Remarks

The research presented in this chapter has made intensive use of recent developments in international macroeconomic theory to provide a formal and empirical framework to examine the likely impact on EU emissions of economic integration of the EU15 and the NAC13. The specification of the model is based on solid microbehavioral assumptions. The intertemporal optimization assumed of representative agents that underlies the model's specification brings considerations of intertemporal trade-offs explicitly into view. And the endogenous determination of the accumulation of physical, financial, and human capital within the model allows for the interplay between these distinctly different factors of growth.

The present research goes beyond analytical exercises conducted with similar CGE frameworks in that the parameters have been estimated with available data and transversality conditions on the theoretical equilibrium solution of the model have been imposed in the estimation process. The estimated model then lends itself to examining not only the environmental impacts of increased economic integration in the EU but also alternative policies that might be employed to achieve targeted emissions reductions efficiently and equitably *when intertemporal optimization is assumed to be operative*. (For discussion of latter policies, see Balta, 2004.)

The estimated values of the model's parameters would appear to be consistent with underlying technical and behavioral assumptions and the estimated model fits well the historical database developed expressly for this research. Simulations of the model in which a progressive increase in trade occurs between the two blocs of EU member states suggests that such a development will favor the EU15 more so than the NAC13. Importantly, the simulations also suggest that emissions by both blocs will decrease if consumers in each bloc develop preferences for each others goods..

The research discussed in this chapter is not without problems, however, and one of the purposes of undertaking it has been to explore just how far this line of inquiry can be profitably pursued. The model employed in the formal, empirical, and speculative analyses has, by its nature, a knife's-edge equilibrium growth-path solution that is saddle-path stable at most. In order to promote stability in policy simulations, limits were necessarily imposed on the values that influential parameters could assume.

A second difficulty has to do with the treatment of real exchange-rate determination within the model. As the model is presently configured, time rates of change in real exchange rates are determined through a standard covered-interest-parity (CIP) relationship, which, it is assumed, is not taken into account by the representative agents in their intertemporal optimization decisions. Consequently there are no co-state equations associated with the exchange rates and no formal feedbacks through the optimality conditions determining the levels of the agents' control variables. On the face of it, this is probably a reasonable assumption to make, in view of how poorly exchange rate dynamics are understood (Sarno and Taylor 2002). But the absence of feedbacks may be a contributing factor to the instability of the knife's edge solution.

In order to examine if this is the case, we have pursued estimation of a more general specification of the model in which additional co-state equations and feedback relations were introduced. However, we have encountered difficulties in the estimation of the more general specification. This is an area in which further work is needed. A formal analysis of the model's stability properties and a sensitivity analysis of stability properties with respect to key parameters also need to be performed in order to understand more fully properties of the model.

One final difficulty to mention lies in the development of the empirical database to enable the estimation of the model. Data were collected from diverse sources and had to be aggregated and disaggregated to achieve appropriate geo-political and temporal scales. To produce quarterly time series, extrapolations and interpolations were carried out where necessary. Problems with geo-political scale were overcome in part by taking into account the relative size of each country's economy in its aggregate bloc. A further issue existed in the case of the NAC13 countries. Since the economies of Central and Eastern European Countries were liberalized only very recently, a consistent set of time series could be constructed only from 1992 onwards. For most of the variables there were some missing observations, whose values were estimated by exponential or linear interpolation procedures. The dearth of consistent time series on macroeconomic aggregates for all the NAC13 countries prior to 1992 and the fact that this period represents one of structural transition might be another factor contributing to model instability. For these reasons, one cannot expect to draw inferences from the estimated model with confidence concerning long-term growth properties of the NAC13 economy.

In spite of the difficulties and limitations noted above, we hope to have demonstrated some of the potential of an empirically based model of endogenous growth to investigate the effects on emissions of increased trade between more closely integrated blocs of nations and to have identified where further investigation is needed. We think the findings reported here provide grounds for cautious optimism on the fronts of both theoretically driven analysis and potential emissions abatement.

References

Abraham R, Chichilnisky G, Record R (1998) North-South trade and the dynamics of the environment. In: Chichilnisky G, Heal GM, Vercelli A (eds) Sustainability: Dynamics and uncertainty. Fondazione Eni Enrico Mattei, Kluwer Academic Publishers, The Netherlands

Adkins LG, Garbaccio RF (1999) A bibliography of CGE models applied to environmental issues. US Environmental Protection Agency, Office of Policy, Office of Economy and Environment

Balta N, Donaghy K (2002) Harmonization of environmental policies between the EU-15 and new accession countries: An empirical framework for the examination of issues, Second World Congress of Environmental and Resource Economists, Monterey, California, June 24-27

Balta N (2004) Climate change policy in an enlarged European Union: Institutions, efficiency and equity. PhD dissertation, submitted to the University of Illinois at Urbana-Champaign in May

Bovenberg AL, Smulders S (1996) Transitional impacts of environmental policy in an endogenous growth model. International Economic Review 37:861-893

Buonanno P, Carraro C, Castelnuovo E, Galeotti M (2000) Emission trading restrictions with endogenous technological change. Nota di Lavoro 43.2000, Fondazione Eni Enrico Mattei, Milan

Chichilnisky G (1994) North-south trade and the environment. American Economic Review 84(4):851-873

Chichilnisky G, Heal GM (1994) Who should abate carbon emissions? An international viewpoint. Economics Letters 44:443-449

Donaghy KP, Federici D, Wymer CR (1999) An empirical two-good, two-country representative-agent model with endogenous growth. Paper presented at the 16th meetings of the Pacific Regional Science Conference Organization, Seoul, July

Durbin J (1963) Trend elimination for the purpose of estimating seasonal and periodic components. In: Rosenblatt M (ed) Time Series Analysis. Wiley, New York, pp 3-16

Galeotti M, Kemfert C (2004) Interactions between climate and trade: A survey. Nota di Lavoro 88.2004, Fondazione Eni Enrico Mattei, Milan, available online at http://www.feem.it/Feem/Pub/Publications/WPapers/default.htm

Heal GM (1999) Emission constraints, emissions permits and marginal abatement costs. In: Boman M, Brannlund R, Kristrom B (eds) Topics in environmental economics. Kluwer Academic Publishers, Dordrecht

Heidug WK, Bertram R (2003) Environmental policy, induced technological change, and economic growth: a selective review. Manuscript.

Isard P (1995) Exchange Rate Economics. Cambridge University Press, Cambridge.

Jeppessen T, List JA, Folmer H (2002) Environmental regulations and new plant location decisions: Evidence from a meta-analysis. Journal of Regional Science,42:19-49

Maxwell JW, Reuveny R (2005) Trade and environment: Theory and policy in the context of EU enlargement and economic transition, Elgar Publishing Ltd

Nordhaus WD (1997) Modeling induced innovation in climate-change policy. Paper presented at the IIASA workshop on induced technological change and the environment, Laxenburg

Smulders S (1995) Environmental policy and sustainable economic growth: An endogenous growth perspective. De Economist 143:163-195

Sarno L,Taylor MP (2002) The economics of exchange rates. Cambridge University Press, Cambridge

Turnovsky S J (1997) International macroeconomic dynamics. MIT Press, Cambridge

Ulph A, Ulph D (1996) Trade, strategic innovation and strategic environmental policy—a general analysis. In Carraro C, Katsoulacos Y, Xepapadeas A (eds) Environmental policy and market structure. Kluwer Academic Publishers, Dordrecht

Wymer CR (1997) Structural non-linear continuous-time models in econometrics. Macroeconomic Dynamics 1:518-548

Wymer CR (2001) Systems estimation and analysis programs. Unpublished mimeo, 31 December 2001

Appendix 15.A: Derivation of the Model

In this appendix we present the derivation of the equations of the model for only the EU15, since the same approach is taken in deriving the equations of the model for the NAC13. We express the intertemporal optimization problem for the EU15 by using a present value Hamiltonian using λ, $q_1^{'}$, and $q_2^{'}$ as the co-state variables associated with the three constraints on capital accumulation:

$$\begin{aligned} H = {} & \frac{1}{\gamma}\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma} e^{-\beta t} \\ & +\lambda e^{-\beta t}\begin{bmatrix} \left(1-\tau_y\right)Y+\left(1-\tau_b\right)rb-\left(1+\tau_c\right)\left(C+\sigma MGS+\sigma_r RMGS\right) \\ -I\left(1+\dfrac{a_1 I}{2K}\right)-R\left(1+\dfrac{a_2 R}{2HC}\right)+T-\dot{b} \end{bmatrix} \\ & +q_1' e^{-\beta t}\left(I-\dot{K}\right)+q_2' e^{-\beta t}\left(R-\dot{HC}\right). \end{aligned} \tag{15.A.1}$$

15.A.1. Optimality Conditions

The optimality conditions for the control variables C, MGS, $RMGS$, I, and R are as follows:

$$H_C = 0 \;\Rightarrow\; \left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma-1} e^{-\beta t} MGS^{\eta_1} RMGS^{\eta_2} = \lambda e^{-\beta t}\left(1+\tau_c\right), \tag{15.A.2}$$

$$H_{MGS} = 0 \Rightarrow$$

$$\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma-1} e^{-\beta t}\eta_1 C \cdot MGS^{\eta_1-1} RMGS^{\eta_2} = \lambda e^{-\beta t}\left(1+\tau_c\right)\sigma, \tag{15.A.3}$$

$$H_{RMGS} = 0 \Rightarrow$$

$$\left(C \cdot MGS^{\eta_1} RMGS^{\eta_2}\right)^{\gamma-1} e^{-\beta t}\eta_2 C \cdot MGS^{\eta_1} RMGS^{\eta_2-1} = \lambda e^{-\beta t}\left(1+\tau_c\right)\sigma_r, \tag{15.A.4}$$

$$H_I = 0 \Rightarrow \quad \lambda e^{-\beta t}\left(1+\frac{a_1 I}{K}\right) = q_1' e^{-\beta t}, \tag{15.A.5}$$

$$H_R = 0 \Rightarrow \quad \lambda e^{-\beta t}\left(1+\frac{a_2 R}{HC}\right) = q_2' e^{-\beta t}. \tag{15.A.6}$$

As equations (15.A.2) through (15.A.4) have C, MGS, and $RMGS$ in common, this three-equation system can be solved for these variables as follows:

$$MGS = \frac{\eta_1 C}{\sigma}, \tag{15.A.7}$$

$$RMGS = \frac{\eta_2 C}{\sigma_r} \text{ and} \tag{15.A.8}$$

$$C^{\gamma-1+\eta_1\gamma+\eta_2\gamma} = \frac{\lambda(1+\tau_c)\sigma^{\eta_1\gamma}\sigma_r^{\eta_2\gamma}}{\eta_1^{\eta_1\gamma}\eta_2^{\eta_2\gamma}}. \tag{15.A.9}$$

Once the time derivative of C has been defined in terms of physical capital stocks, then solution paths of *MGS* and *RMGS* are very straightforward. In finding the solution path of *C*, one needs to differentiate both sides of equation (15.A.9) logarithmically with respect to time. By subtracting multiples of *log K* in coefficients of *log C* from both sides of this equation, one obtains the following equation that determines the evolution of consumption per unit of physical capital stock, c [14]:

$$\dot{c} = c \cdot \left[\frac{1}{(\gamma-1+\eta_1\gamma+\eta_2\gamma)}\left(\frac{\dot{\lambda}}{\lambda}+\eta_1\gamma\frac{\dot{\sigma}}{\sigma}+\eta_2\gamma\frac{\dot{\sigma}_r}{\sigma_r}\right)-k\right]. \tag{15.A.10}$$

In above equation, $\dot{\lambda}/\lambda$, will later be substituted for by an expression defined in terms of observable variables and parameters of the model and which follows from an intertemporal arbitrage relationship.

In similar fashion, one may obtain expressions for the time rates of change in imports from the other bloc and the rest of the world, mgs and $rmgs$:

$$m\dot{g}s = \frac{\eta_1}{\sigma}\dot{c} - \eta_1 c\frac{\dot{\sigma}}{\sigma^2}, \tag{15.A.11}$$

$$r\dot{m}gs = \frac{\eta_2}{\sigma_r}\dot{c} - c\eta_2\frac{\dot{\sigma}_r}{\sigma_r^2}. \tag{15.A.12}$$

In the last two first-order optimality conditions, $e^{-\beta t}$ is eliminated from both sides of equations (15.A.5) and (15.A.6), and newly defined variables, q_1, q_2, k, and κ are substituted, yielding

$$q_1 = 1 + a_1 k, \tag{15.A.13}$$

$$q_2 = 1 + a_2\kappa. \tag{15.A.14}$$

---

[14] When a new variable, consumption in this case, is defined in terms of physical capital stocks, $c = C/K$, the following relationship is implied in their time rates: $\dot{c}/c = \dot{C}/C - \dot{K}/K$.

These expressions will prove useful in deriving the solution paths of k and κ from arbitrage relationships discussed in the next section.

15.A.2 Arbitrage Relationships

In deriving the specification of the budget constraint determining the evolution of net foreign assets per unit of physical capital stock, *z,* we make two modifications: first, in order to capture the effects of trade on the domestic economy, we define a domestic supply constraint, as below:

$$Y = C + MGS^* + RXGS + G + I\left(1 + \frac{a_1 I}{2K}\right) + R\left(1 + \frac{a_2 R}{2HC}\right). \tag{15.A.15}$$

In this identity, MGS^* represents goods produced domestically for foreign sales, and $RXGS$ denotes exports to the rest of the world. Assuming for convenience that government revenues and expenditures are in balance at every point in time, i.e., $T = G$, (15.A.15) can be solved for G and the resulting expression substituted for the lump-sum tax rebate, *T*, in (15.5). Upon rearranging terms, one obtains the following specification of the budget constraint.

$$\begin{aligned}\dot{b} &= \tau_y Y + \tau_c C + \left(1 + \tau_b\right) rb + MGS^* + RXGS \\ &-\left(1 - \tau_c\right)\left(\sigma \cdot MGS + \sigma_r RMGS\right).\end{aligned} \tag{15.A.16}$$

By manipulating this equation to yield the time rate of change in *z*, which by definition is $\dot{z} = \dot{b} / K - b\dot{K} / K^2$, the following equation can be derived, in which MGS^* / K is substituted for by the equivalent expression, $MGS^* / K^* \cdot K^* / K$, and rewritten in terms of the variables of the estimated model as $mgs^* \cdot K^* / K$.

$$\begin{aligned}\dot{z} &= \tau_y y + \tau_c c + \left(1 + \tau_b\right) rz - \left(1 - \tau_c\right)\left(\sigma mgs + \sigma_r rmgs\right) \\ &+ \frac{K^*}{K} mgs^* + rxgs - zk.\end{aligned} \tag{15.A.17}$$

Critical intertemporal arbitrage (or trade-off) relationships in the model follow directly from the co-state equations in the first-order necessary conditions of the intertemporal optimization decisions of the representative agents. These conditions are as follows:

$$\frac{\partial H}{\partial b} = -\dot{\lambda} e^{-\beta t} + \lambda \beta e^{-\beta t} \;\Rightarrow\; \lambda e^{-\beta t}\left(1 - \tau_b\right) r = -\dot{\lambda} e^{-\beta t} + \lambda \beta e^{-\beta t}, \tag{15.A.18}$$

$$\frac{\partial H}{\partial K} = -\dot{q}_1' e^{-\beta t} + q_1' \beta e^{-\beta t} \;\Rightarrow\; \lambda e^{-\beta t}\left[(1 - \tau_y) Y_K + \frac{a_1 I^2}{2K^2}\right] = -\dot{q}_1' e^{-\beta t} + q_1' \beta e^{-\beta t}, \tag{15.A.19}$$

$$\begin{aligned}\frac{\partial H}{\partial HC} = -\dot{q}_2' e^{-\beta t} + q_2' \beta e^{-\beta t} \;\Rightarrow\; & \lambda e^{-\beta t}\left[\left(1 - \tau_y\right) Y_{HC} + \frac{a_2 R^2}{2HC^2}\right] \\ & = -\dot{q}_2' e^{-\beta t} + q_2' \beta e^{-\beta t}.\end{aligned} \tag{15.A.20}$$

Dividing both sides of equation (15.A.18) by $\lambda e^{-\beta t}$ and rearranging terms, yields the Keynes-Ramsey consumption rule[15]:

$$\frac{\dot{\lambda}}{\lambda} = \beta - \left(1 - \tau_b\right) r . \tag{15.A.21}$$

The trend term $e^{-\beta t}$ can be eliminated from the co-state equation corresponding to K, (15.A.19) as can the market value of capital, q_1, by making use of the definition (15.A.13). Differentiating the zero-order optimality condition for *I* given in equation (15.A.13) fully with respect to time and making appropriate substitutions into (15.A.19) one obtains a restatement of the co-state equations strictly in terms of observable quantities, which also defines the evolution of the rate of change in the physical capital stock.

$$\dot{k} = \left(\frac{1}{a_1} + k\right)\left(1 - \tau_b\right) r - \frac{\left(1 - \tau_y\right)}{a_1} \mu_2 y - \frac{1}{2} k^2 . \tag{15.A.22}$$

Proceeding similarly, the co-state equation corresponding to *HC*, (15.A.20), can be rewritten as below.

$$\dot{\kappa} = \left(\frac{1}{a_2} + \kappa\right)\left(1 - \tau_b\right) r - \frac{\left(1 - \tau_y\right)}{a_2} \frac{y\varepsilon}{h} - \frac{1}{2} \kappa^2 . \tag{15.A.23}$$

15.A.3 Transversality Conditions

Intertemporal optimality requires the following transversality conditions to hold at each point of time:

$$\lim_{t \to \infty} \lambda b e^{-\beta t} = 0, \quad \lim_{t \to \infty} q_1' K e^{-\beta t} = 0, \quad \lim_{t \to \infty} q_2' HC e^{-\beta t} = 0. \tag{15.A.24}$$

As all three of these conditions are equal to zero and must be met simultaneously, they can be incorporated in one condition to be imposed on the model in estimation.. Making substitutions to accommodate newly defined variables and zero-order optimality conditions, (15.A.24) can be rewritten as

$$\lim_{t \to \infty} \lambda e^{-\beta t} \left(b + \left(1 + a_1 k\right) K + \left(1 + a_2 \kappa\right) HC\right) = 0 . \tag{15.A.25}$$

Dividing all the terms by *K* will allow rewriting this equation in terms of the variables of the model:

$$\lim_{t \to \infty} \lambda e^{-\beta t} \left(z + 1 + a_1 k + h + a_2 \kappa h\right) = 0 . \tag{15.A.26}$$

---

[15] This equation is used to substitute for $\dot{\lambda}/\lambda$ in equation (15.A.9), determining the time path of c.

15.A.4 Other Optimality Conditions

Recalling the static optimality condition for employment of energy resources, we have,

$$P^e = Y_E = \alpha e^{\rho t} HC^{\varepsilon} L^{\mu_1} K^{\mu_2} \mu_3 E^{\mu_3 - 1} . \qquad (15.A.27)$$

where the real unit price of energy, P^e, is equated to the marginal product of energy. Isolating *E*, dividing both sides of the resulting expression by *K*, and differentiating logarithmically with respect to time yields the following equation.

$$(1-\mu_3)\left(\frac{\dot{E}}{E} - \frac{\dot{K}}{K}\right) - \varepsilon \frac{\dot{K}}{K} = \rho + \varepsilon\left(\frac{\dot{HC}}{HC} - \frac{\dot{K}}{K}\right) + \mu_1\left(\frac{\dot{L}}{L} - \frac{\dot{K}}{K}\right) - \frac{\dot{P}^e}{P^e} . \qquad (15.A.28)$$

upon rearranging and substituting the notation employed above—n, h, l and p^e —for energy, human capital, labor per unit of physical capital, and the proportionate rate of change in the price of energy, the equation defining the time path of energy per unit of capital can be written as

$$\dot{n} = \frac{n}{(1-\mu_3)}\left(\rho + \varepsilon\kappa + \mu_1 \frac{\dot{l}}{l} - p^e\right). \qquad (15.A.29)$$

The time path of GDP per unit of capital can also be stated similarly. Dividing equation (15.1) by *K*, differentiating logarithmically with respect to time, and isolating time rate of change in *y*, yields

$$\dot{y} = y\left(\rho + \varepsilon\kappa + \mu_1 \frac{\dot{l}}{l} + \mu_3 \frac{\dot{n}}{n}\right). \qquad (15.A.30)$$

Appendix 15.B Data Aggregation

The published sources of data used in estimation of the model are listed below. The procedure followed in aggregating to bloc level data collected at the national level is then presented in detail for each variable.

15.B.1 Published Sources of Data

WDI, *World Development Indicators*, annually, World Bank.
DOT, *Direction of Trade*, quarterly, International Monetary Fund.
NA, *National Accounts*, quarterly or annually, Detailed Tables, Vol:2, OECD.
IFS, *International Financial Statistics*, quarterly or annually, International Monetary Fund.
IEA, *International Energy Annual*, annually, Energy Information Administration of the US Government.
EPT, *Energy prices and taxes*, quarterly, OECD.

15.B.2 Variable Definitions:

The blocs of national economies aggregated in this study are as follows:

EU15: *European Union Countries* comprising Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, United Kingdom.
NAC13: *New Accession Countries* comprising Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovak Republic, Slovenia, Turkey.

General Notation The following general notation is used in defining the variables of the model:[i]

$\varepsilon_t^{i/US\$}$: Exchange rate, local currency per US dollar (IFS).

P_t^{US} : Price deflator for US where 1995=1 (WDI).

P_t^i : Price deflator for country i where 1995=1 (WDI).

Y_t^i : Gross Domestic Product (GDP) of country i[1,ii] (IFS).

P : EU15 price deflator

$$P_t = \sum_{i=1}^{15} \left[\frac{\dfrac{Y_t^i / \varepsilon_t^{i/US\$}}{P_t^{US}}}{\sum_{j=1}^{15} \dfrac{Y_t^j / \varepsilon_t^{j/US\$}}{P_t^{US}}} \right] * P_t^i .$$

P^* : NAC13 price deflator

$$P_t^* = \sum_{i=1}^{13} \left[\frac{\dfrac{Y_t^i / \varepsilon_t^{i/US\$}}{P_t^{US}}}{\sum_{j=1}^{13} \dfrac{Y_t^j / \varepsilon_t^{j/US\$}}{P_t^{US}}} \right] * P_t^i .$$

Macroeconomic Variables

Net foreign assets:

b: EU15 net foreign assets in constant 1995 EUROs.

$$b_t = \sum_{i=1}^{15} \left[NFA_t^i / \varepsilon_t^{i/euro} / P_t \right],$$

where NFA_i= Net foreign assets in current terms in local currency (IFS).

b^*: NAC13 net foreign assets in constant 1995 NACCUs.

$$b_t^* = \sum_{i=1}^{13}\left[NFA_t^i / \varepsilon_t^{i/US\$} * \left(\varepsilon_t^{NACCU/US\$} / P_t^*\right)\right].$$

Private consumption:

C: EU15 private consumption in constant 1995 EUROs.

$$C_t = \sum_{i=1}^{15}\left[C_t^i / \varepsilon_t^{i/euro} / P_t\right],$$

where C_i= Private consumption in current terms in local currency (IFS[iii]).

C^*: NAC13 private consumption in constant 1995 NACCUs.

$$C_t^* = \sum_{i=1}^{13}\left[C_t^i / \varepsilon_t^{i/US\$} * \left(\varepsilon_t^{NACCU/US\$} / P_t^*\right)\right].$$

Net capital formation:

I: EU15 net capital formation in constant 1995 EUROs.

$$I_t = CF_t - CC_t = \sum_{i=1}^{15}\left[CF_t^i / \varepsilon_t^{i/euro} / P_t\right] - \left[CC_t^i / \varepsilon_t^{i/euro} / P_t\right],$$

where CF_i : Capital formation in current terms in local currency (IFS[iv],[v]),
CC_i : Capital consumption in local currency (NA[vi],[vii]).

I^*: NAC13 net capital formation in constant 1995 NACCUs.

$$I_t^* = \frac{\sum_{i=1}^{13} CF_t^i / \varepsilon_t^{i/US\$}}{\sum_{j=1}^{3} CF_t^j / \varepsilon_t^{j/US\$}}\left[\sum_{j=1}^{3}\frac{CF_t^j}{\left(CF_t^1 + CF_t^2 + CF_t^3\right)}\left(\frac{CF_t^j - CC_t^j}{\varepsilon_t^{j/US\$}}\right)\right] * \frac{\varepsilon_t^{NACCU/US\$}}{P_t^*}$$

Trade Variables

Imports from the other bloc:

MGS: EU15 total imports from the NAC13 in 1995 NACCUs.

$$MGS_t = \sum_{i=1}^{15}\left(MGS_t^i \cdot \varepsilon_t^{NACCU/US\$} / P_t^*\right),$$

where MGS_i= Imports by the EU15 from the NAC13 in 1995 US\$ (DOTS).

MGS*: NAC13 total imports from the EU15 in 1995 EUROs (DOTS).

$$MGS_t^* = \sum_{i=1}^{13}\left(MGS_t^i \cdot \sigma_{r,t}\right)\big/P_t \, .$$

Imports from the rest of the world:

RMGS: EU15 total imports from the rest of the world in 1995 US$

$$RMGS_t = \sum_{i=1}^{15}\left(RMGS_t^i / P_t^{US}\right),$$

where $RMGS_i$= Imports from the rest of the world in 1995 US$ (DOTS).

RMGS*: NAC13 total imports from the rest of the world in 1995 US$ (DOTS).

$$RMGS_t^* = \sum_{i=1}^{13}\left(RMGS_t^i / P_t^{US}\right).$$

Other Aggregate Variables

Labor:

L: EU15 labor force.

$$L_t = \sum_{i=1}^{15} L_t^i \, ,$$

where L_i= Labor force for country i (WDI).

L^*: NAC13 labor force.

$$L_t^* = \sum_{i=1}^{13} L_t^i \, .$$

CO_2 emissions:

M: EU15 total CO_2 emissions.

$$M_t = \sum_{i=1}^{15} M_t^i \, ,$$

where M_i= CO_2 emissions in million (10^6) metric tons of Carbon equivalent (IEA).

M^*: NAC13 total CO_2 emissions.

$$M_t^* = \sum_{i=1}^{13} M_t^i .$$

Energy use:

E: EU15 total commercial energy use.

$$E_t = \sum_{i=1}^{15} E_t^i ,$$

where E_i= Commercial energy use in kilo tones of oil equivalent (WDI).

E^*: NAC13 total commercial energy use.

$$E_t^* = \sum_{i=1}^{13} E_t^i .$$

Research and Development (R&D) Expenditures:

R : EU15 R&D expenditures in constant EUROs.

$$R_t = \left(\sum_{i=1}^{14} R_t^i * \left(GDP_t^i / \varepsilon_t^{i/euro} \right) \right) / P_t ,$$

where R_i= R&D expenditures as percentage of GDP (WDI)[viii].

R: NAC13 R&D expenditures in constant NACCUs.

$$R_t = \sum_{j=1}^{11} R_t^j * GDP_t^j / \varepsilon_t^{j/US\$} * \left(\varepsilon_t^{NACCU/US\$} / P_t^* \right) .$$

Energy price:

P^e: EU15 retail oil products energy price index where 1995=1.

$$P_t^e = \left(\sum_{i=1}^{12} P^{e\,i}_{\;t} * \left(Y_t^i / \sum_{j=1}^{12} Y_t^j \right) \right) \Big/ P_t ,$$

where P^i= Retail oil products energy price index for country i (EPT)[ix], [x].

P^{e*}: NAC13 retail oil products energy price index where 1995=1.

$$P^{e*}_t = \left(\sum_{i=1}^{3} P^{e\,i}_{\;t} \left(Y_t^j / \sum_{j=1}^{3} Y_t^j \right) \right) / P_t^* .$$

Exchange Rates

Normalized real exchange rate:

σ_r: Normalized real exchange rate between EU15 and the rest of the world (DOT and IFS).

$$\sigma_{r,t} = \left(\varepsilon_t^{euro/US\$} \cdot P_t^{US} / P_t\right) \Big/ \varepsilon_{1995/4}^{euro/us\$} .$$

Trade-weighted normalized real exchange rate:

σ: Trade-weighted normalized real exchange rate between EU15 and the NAC13 (DOT and IFS).

$$\sigma_t = \sum\nolimits_{j=1}^{13} \left(\frac{XGS_t^j + MGS_t^j}{\sum\nolimits_{i=1}^{13}\left(XGS_t^i + MGS_t^i\right)} \varepsilon_t^{j/euro} \right) \cdot \frac{1}{\varepsilon_{1995/4}^{j/euro}} \cdot \frac{P_t}{P_t^*} .$$

σ_r^*: Trade-weighted normalized real exchange rate between NAC13 and the rest of the world (DOT and IFS).

$$\sigma_r^* = \left(\sum\nolimits_{i=1}^{13} \frac{XGS_t^i + MGS_t^i}{\sum\nolimits_{j=1}^{13}\left(XGS_t^j + MGS_t^j\right)} * \frac{\varepsilon_t^{i/US\$}}{\varepsilon_{1995}^{i/US\$}} \right) * \frac{P_t^{US}}{P_t^*} .$$

---

[i] t subscript denotes time in all the variables.

[ii] Quarterly data are interpolated from annual figures for Cyprus, Hungary, Poland, Romania and Slovenia.

[iii] Capital consumption in local currency is available for Czech Republic, Slovakia and Turkey.

[iv] Quarterly data are interpolated from annual figures for Greece, Ireland and Luxembourg.

[v] Quarterly data are interpolated from annual figures for Cyprus, Hungary, Poland, Romania and Slovenia.

[vi] Quarterly data are interpolated from annual figures for Greece, Ireland, Luxembourg, Portugal, Spain and United Kingdom.

[vii] Capital consumption in local currency is available for Czech Republic, Slovakia and Turkey.

[viii] We have assumed the available data series to sum of GDP of these countries ratio holds for the aggregate bloc as well. In the EU15, data are not available for Luxembourg, neither for Lithuania and Malta in the NAC13.

[ix] In the EU15, data are available for Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Portugal, Sweden, United Kingdom.

[x] In the NAC13, data are available for Czech Republic, Hungary and Poland.

16 Modeling Globalization: A Spatial Econometric Analysis

Bernard Fingleton

Department of Land Economy, University of Cambridge, Cambridge, UK

16.1 Introduction

We live in a state of global interdependency, in which actions, events and decisions made in one country may have ramifications for many others. Although the phenomena associated with globalization are far reaching and the concern of many disciplines, from the perspective of a geographical economist, it is natural to focus on the economic consequences and causes of globalization. I therefore commence with a working definition of globalization as 'the 'increased interdependence of national economies, and the trend towards greater integration of goods and factor markets' (Neary 2003), in an attempt to provide a focus for the discussion.

The consequences of increased economic interdependence and integration are deep and wide, affecting many sectors of society, many institutions and many aspects of economic production and social and economic development. Clearly conditions have changed, so that societies that were in some kind of long-run stable equilibrium, have received shocks from outside which sometimes have had negative as well as beneficial consequences. Tribes deep in the Amazon jungle drinking coca-cola and wearing denim jeans are just one illustration of the way in which even the previously most detached parts of our human family are now consumers and producers responding to, and interacting with, the global economy. However, if the consequences were confined to the already marginalized, it is unlikely that attention would be focused on globalization in the way that it is. What appears to be a distinct possibility is that globalization may have profound negative effects on sectors of the economy and society in wealthy countries also. The fear is that third world growth associated with globalization will come at the expense of jobs and wages in the developed economies. The cause of the upward drift in unemployment rates experienced in many European countries, together with the increased within-country income disparities between skilled and unskilled workers, may be partly attributed to low wage competition from overseas.

The causes of increased international interdependence are various and debatable, but technological change related to communications is at the core of the matter. Communication between remote localities is now much easier, quicker and less expensive than in previous eras, and this has had important consequences for the interlinked processes involved in the creation, production and dissemination of goods and services. It is now possible for companies to search across the globe for locations that are optimal for production, with appropriately skilled and competitively priced labor and with good access to suppliers and markets. These locations will not necessarily be where the company HQ is located, or where research and development is carried out, since with the new communications technology that is now available, the location of these different but closely integrated functions can be individually optimized. Optimizing individual location choices in this way often provides net benefits to the company that are greater than the returns one would receive if they were say, located on the same site, or within the same country. Part of this process of functional separation has been the emergence of global cities (Sassen 1994) which seemingly act as command and control centers for the global economy, and the development of a ‘space of flows’ (Castells 1989, 2000) operating as a virtual information highway, linking key centers of innovation, production and control, allowing localized specialization but at the same time excluding some cities, regions and countries from higher value-added activities. The inference that follows from this technology-enabled division of labor is that it is closely implicated in the mechanism that has produced a large and growing discrepancy between the wage rates and living standards both within and between countries.

However to characterize the outcome of this state of enhanced interdependence as simply persistent or increasing GDP per capita, productivity and wage discrepancies within and between countries would be a gross generalization of the possible dynamics ensuing from globalization. With the aid of a formal model, we are able to obtain clearer and deeper insights and a richer account of the causes of globalization and of the possible paths to be taken by national economies as the global economy evolves. There are costs to be borne however by attempting to formally model globalization, rather than simply describing it. One is that a model entails simplifying assumptions, and these may be difficult to support in an empirical sense. Operationalizing formal theory, testing assumptions and confronting the predictions made by theory with data brings us into the realm of econometrics, and it is this meeting point, and the difficulties it engenders, which is the concern of this paper.

16.2 Theorizing Globalization

While we could look to a variety of theoretical stances to produce global dynamics, this paper takes the model of Krugman and Venables (1995) as the key point of departure and as a leading contribution to our understanding of globalization, since it encapsulates theoretical assumptions, relating essentially to the presence of increasing returns to scale, which are in accord with the main consensus of the geographical economics community. This theory provides a clear prediction of rising then falling inequality, characterized by an inverted U curve. The prediction is based on the combination of increasing returns within countries and inexorably-falling between-country trade or transport costs, with the consequence that the economies of scale inherent in a specific location increase in significance and then recede, ultimately leading to equalized real wage rates as a long-run equilibrium. As Krugman and Venables (1995) explain, 'at high transport costs all countries have some manufacturing, but when transport costs fall below a critical value, a core-periphery spontaneously forms, and nations that find themselves in the periphery suffer a decline in real income. At still lower transport costs there is convergence of real incomes, in which peripheral nations gain and core nations may lose.' We therefore have a formal model predicting what many politicians and critics of globalization have feared, the impoverishment of the peripheral nations followed by the possible loss of low-skilled jobs to industrializing third world countries. It is therefore important that this theory is put under the empirical microscope, since otherwise policy instruments may be put in place that are wrongly conceived. It may be the case that wage disharmony between countries and within them may be consequence of other factors that should be the subject of policy, rather than being an inevitable outcome of globalization.

In order to abstract from the complicating influence of reality, the Krugman and Venables (1995) model, which is elaborated in Fujita *et al.* (1999), starts from a position of no inherent differences between countries apart from location. The theoretical basis is a market structure derived from Dixit-Stiglitz monopolistic competition theory, with manufactures under monopolistic competition, and the manufacturing sector also produces intermediates, which are also under increasing returns with transport costs, and these provide the country-specific externalities leading to agglomeration of industrial activity. The mechanism is as follows, we start out with (for some reason) a marginally greater level of industrial development in one region. Two is the usual number for the purposes of exposition, often labeled North and South and pertaining to groups of national economies, not regions within a single national economy. This marginal advantage creates a bigger market for intermediates, and this bigger (more varied) intermediate sector in turn reduces costs for final goods, which creates an even bigger market for intermediates, and it is this market interdependence via input-output linkages that provides the externalities. Unlike some NEG models, in the long-run real wage differences are not eliminated by migration, but by capital being ultimately attracted to no-longer remote peripheral locations. Initially, as trade costs fall there is divergence in real wages as the core, with its marginal advantage, becomes dominant and labor demand drives up real wages. In contrast labor demand falls in the periphery. But ever-falling transport costs ensure that a point is reached at which the periphery becomes competitive and starts to reindustrialize. As Krugman and Venables (1995) state, 'At some point the decline in transport costs will be sufficient that the lower wage rate in the periphery more than offsets the disadvantage of being remote from markets and suppliers. At this point manufacturing will have an incentive to move out from the core to the periphery once again, forcing a convergence of wages rates.' Close access to markets and suppliers becomes less relevant and therefore the core loses its rationale.

There are some obvious limitations to the above theory which weaken its attraction. One is that the theory assumes that monopolistic competition is an appropriate market structure for manufactures. Neary (2003) considers globalization under a model of oligopoly, arguing that the presence of strategic interaction, with firms relatively large within their sectors, is much more realistic than assuming that they are incredibly small, equal in size at equilibrium, and with no influence on the market presence or behavior of others. Another limitation of the story told by Krugman and Venables (1995) is the assumption that international labor migration is relatively unimportant compared with international trade and the mobility of capital, whereas the empirical facts point also to extensive international migration. Peeters and Garretsen (2000) develop a model that introduces international migration in skilled labor, which suggests that the outcome of globalization is increased regional economic integration rather than global economic integration. However, even this move towards greater realism is insufficient, because there is also extensive international migration of unskilled labor, as is evident to anyone analyzing the labor force of central London for example, which has numerous nationalities occupying unskilled jobs support the productive activity of the city. More precisely, while the migrant labor may be skilled, it is often employed in unskilled jobs, or is unemployed, in the host country. There are in the British economy, for example, people such as qualified engineers and teachers from countries such as Iran employed as hospital cleaners, part of the reason being inadequate language skills. An additional criticism of the Krugman and Venables (1995) model is that the labor market always clears, but in reality there are wide discrepancies in both unemployment and wage rates. What the data show are different unskilled worker wage and unemployment rates, with wage inflexibility causing globalization to affect not so much wages rates as unemployment rates.

The globalization process seems to involve a step-like progression, with rapid development switching from country to country, so that if we take a snapshot at any particular time we see certain countries (e.g., Taiwan, S Korea) moving up the GDP per capita ladder quite rapidly, following in the wake of earlier rapid-developers such as Japan, which now are at a high level of GDP per capita, although now at a lower rate of growth. Fujita et al. (1999) provide a formal mechanism for the step-like progression, that is based on an ever-increasing rise in the demand for manufactures, which they attribute to an inexorable growth in the level of technology. Therefore in their ‘history of the world part II’, rather than falling transport costs causing the core region to initially benefit from external economies derived from linkages at the expense of relative real wage rates in the periphery, in this formalization rising real wages in the core are a response to a rise in manufacturing demand. Again we see this discrepancy reduce in the longer run, as the wage gap between the high and low GDP per capita countries becomes unsustainable, with firms migrating to a low cost country as wages become too high and the benefits of a low cost production site become apparent. Once the process starts, the low cost country becomes increasingly attractive to inward investment, since it too begins to benefit from a build-up of upstream suppliers and downstream customers and the self-reinforcing effects of a greater variety of linked industries. The consequence of this is that wages in the second country also start to rise, and this lead in turn to firms migrating to a third, lower-cost, location.

Thus, in any one instant, we should according to this model see wide discrepancies between wage levels in different countries, with slower wages growth in countries at the top of the wages ladder due to a loss of manufacturing jobs reflecting a slowdown in labor demand as jobs migrate to low wage cost locations. Some lower wage countries will experience rapidly rising wage levels as the demand for labor increases and as they benefit from self-reinforcing externality effects. On the other hand other low wage countries that have not benefited from the inflow of manufacturing of manufacturing capital and jobs will tend to have a low rate of wage growth.

16.3 An Empirical Model

We can model this relationship quite succinctly as a non-linear relationship between wage level and wage growth, with the non-linearity reflecting the changing importance of positive externalities as the wage level changes. At low wage levels, growth is slow because there is a lack of interlinked industries upstream and downstream. As these develop, we see growth accelerate and wage levels rise, until the change in wage levels reaches a peak beyond which the rate of change falls. This downward turn reflects the loss of demand for manufacturing jobs as firms migrate to lower cost locations. In short, we model the evolution of the economy by a quadratic function (see fig. 16.1).

Of course, wage levels in any one country will change due to a multitude of factors, many of which we cannot know. In the interest of realism, we nevertheless need to represent these additional effects on the change in wages, and one obvious and simple solution is to add a stochastic error term ε to our model, in which case the functional relationship between wage change Δw and wage level w is given by the quadratic function:

$$\begin{aligned} &\Delta w = aw^2 + bw + c + \varepsilon = Xf + \varepsilon \\ &\varepsilon \sim N(0, \sigma^2) \end{aligned} \tag{16.1}$$

in which a, b and c are regression coefficients, while in the matrix form X is an n by 3 matrix with columns equal to w^2 , w, and a vector of ones, and f is the 3 by 1 vector containing a, b and c. In order to be consistent with the globalization story of rising than diminishing externality effects, this model should have significant regression coefficients for both w and w^2 with a positive coefficient on the former and a negative coefficient on the latter.

Up to this point, we have considered globalization exclusively, with no discussion of how this might relate to spatial econometric models. However, it is apparent that equation (16.1) is very incomplete as a model of globalization. We have assumed that wage change will partly be a response to the random shocks ε, which for any single country will cause the actual wage change to differ from what one would expect on the basis purely of the wage level in that country. But, in considering globalization, it is impossible to ignore the fact that shocks are transmitted worldwide, in other words a shock to one economy is invariably a shock to others also. It turns out that this interdependence of economies can be explicitly modeled via the so-called spatial error model (see Anselin, 1988, Upton and Fingleton, 1985), so that :

$$\begin{aligned} &\Delta w = Xf + \xi \\ &\xi = \rho W \xi + \varepsilon \\ &\varepsilon \sim N(0, \sigma^2 I) \end{aligned} \tag{16.2}$$

in which W is the n by n matrix defining the interconnectivity between countries. The simplest possible structure for W is as a set of ones and zeros, with ones defining contiguous countries and zeros defining other non-contiguous countries. This seems however to be unnecessarily restrictive, since one would expect the direct effects of shocks to be conditioned by trade costs which are proxied by the distance between trading economies, in which case we would prefer to assume that:

$$\begin{aligned} &W_{ij} = \frac{d_{ij}^{-\beta}}{\sum_j d_{ij}^{-\beta}} \\ &W_{ij} = 0, d_{ij} > 10{,}000 miles \end{aligned} \tag{16.3}$$

in which d_{ij} is the great circle distance between the capital cities of countries i and j and β defines the rate of decay with distance. For example, with β relatively large, then the impact of a shock will be confined more to nearby countries than if β was relatively small. Note also that the W matrix has been standardized by dividing by the total of each row. This is not necessary, but row totals summing to one make the estimated coefficient ρ easier to interpret, since it has an upper bound equal to one. The effect of standardizing is to make distances relative, so that if for example there are three capital cities in row j at distances 100, 200 and 300 miles, the standardized W values will be the same as if the distances are 1000, 2000 and 3000.

One alternative to the use of distances would be to define the cells of the W matrix directly by international trade data, assuming that shock-effects would be proportional to the trade links between countries. While this in principal would seem to be a good idea, and trade data has been used in the past, for instance as an indicator of the intensity of R&D spillovers between OECD countries (Coe and Helpman, 1995, Verspagen 1997), there are also some difficulties with this approach. In the case of a sample of countries that includes underdeveloped countries, obtaining comprehensive and accurate trade data may not be easy. Also, trade volumes and directions vary considerably over time, and therefore to convert these into a viable W matrix format would require some considerable simplification and numerous assumptions, which may be hard to justify.

An alternative would be to examine in more detail the structure of the residuals from fitting a model without any spatial interaction effects, for example the residual correlogram may suggest the range of distances over which the spatial effects may extend. Also, the Moran scatterplot and related techniques may suggest countries that are particularly closely linked, thus refining our modeling. However, these more sophisticated techniques are for the future. In the empirical modeling in this paper, we base the W matrix simply on great circle distances.

16.4 Model Estimates

The estimation that follows is based on a sample of 98 countries, with data taken from the Penn World Table Version 6.1 (October 2002). The variable Δw is approximated by the change over the period 1970 to 2000 in real GDP per worker, expressed in 1996 constant prices. Table 16.1 gives the initial model fitted as equation (16.1), with the resulting estimates of a, b, c and σ together with associated t-ratios and goodness of fit statistics. Interestingly, for these 98 countries, a quadratic function does not seem to describe the relationship between Δw and the initial level of GDP per worker (w) in 1970, a date which many would consider to be the time at which the globalization process commenced. Of course this is an exceedingly simple model, only accounting for less than 30% of the variance in Δw.

When the residuals from the table 16.1 model are examined, its limitations are unmistakable. There is evidence of very significant residual spatial autocorrelation. For example, using the standardized W matrix as defined in equation (16.3) with $\beta = 1.0$, Moran's I has a value equal to 0.1179 compared with an expected value equal to -0.01447 and a variance of 0.0004053 under the equations defining the randomization moments. This gives a standardized value equal to 6.574, which is well outside the range of plus or minus 1.96 that one would conventionally use in conjunction with the $N(0,1)$ distribution to define the significance of the test statistic. This result is reaffirmed by the empirical randomization distribution, which for 100 replications has a mean equal to -0.01332 and variance equal to 0.0004029, giving a standardized value of 6.536.

Table 16.1. OLS estimates of equation (16.1)

Parameter	estimate
c	-997.0
t-ratio	-0.48
b	1.022
t-ratio	3.17
a	-0.00001329
t-ratio	-1.55
σ	10011.0
loglikelihood	-1040.2758
r-squared	0.2912

The residual autocorrelation could be due to omitted spatially autocorrelated variables, or it might simply be a spatial error process reflect the transmission of shocks between 'neighboring' (as defined by *W*) countries. In other words in may be substantive, or simply a so-called 'nuisance' effect. Of course there are many variables that one would wish to introduce in order to explain in a substantive way the causes of the change in wage levels over the study period, but in this chapter, we opt for the simpler approach, which does not demand additional data and hypotheses beyond those suggested from the foregoing globalization theory. Consequently, the preferred model is given by equation (16.2).

Table 16.2. ML estimates of equation (2) (non-standardized *W* matrix)

	$\beta = 0.1$	$\beta = 0.5$	$\beta = 1$	$\beta = 1.5$	$\beta = 2$	$\beta = 2.5$
c	-1841.726	-3822.156	-1170.822	-801.004	-934.415	-979.607
	-0.56	-0.82	-0.38	-0.36	-0.45	-0.47
b	1.084	1.154	1.133	1.086	1.039	1.024
	3.45	3.61	3.44	3.33	3.23	3.21
a	-1.4790E-05	-1.8236E-05	-2.1123E-05	-1.7905E-05	-1.4603E-05	-1.3574E-05
	-1.79	-2.20	-2.48	-2.08	-1.71	-1.60
ρ	0.01000	0.05000	0.1790	0.1660	0.07500	0.02800
	2.615	7.730	6.002	3.557	1.410	0.6276
σ	9724	9430	9228	9579	9787	9838
loglikelihood	-1039.3396	-1037.0447	-1035.4702	-1038.4578	-1039.8569	-1040.1528

Given the lack of precise knowledge f the structure of the *W* matrix, I have opted to search through a range of specifications, initially working with the non-standardized *W* matrix (equation 16.3 without the denominator). Table 16.2 gives the ML estimates of equation (16.2) based on the same data as was used for table 16.1, with different estimates obtained for the different *W* matrices defined by alternative β values. The most notable feature of these estimates is that the existence of the spatial error significantly improves the level of fit, and for the maximum of the maximized loglikelihoods ($\beta = 1$) the quadratic function is a significant feature of these data, with parameter b always positive and significant, and c always negative. The non-standardized *W* has implications for the estimate of ρ, which is automatically constrained within upper or lower bounds given by the inverse maximum and minimum eigenvalues of the *W* matrix. In order to satisfy the constraint, which ensures a stable autoregressive error process (see Fingleton, 1999), the likelihood function includes a term that acts as a penalty or weighting function. This has the effect that the likelihood, which is based on a normality assumption, diminishes sharply as ρ approaches its upper or lower bound. Therefore the different ρ estimates are not directly comparable, since each different non-standardized *W* has different eigenvalues. These eigenvalues also facilitate the estimation of the asymptotic variance-covariance matrix, which provides the standard errors used to give the *t*-ratios in table 16.2.

Table 16.3 gives the corresponding estimates based on the standardized *W* matrix. In this case what matters is relative not absolute distance, and because *W* has been standardized the maximum eigenvalue is equal to 1 regardless of the assumed value of β, and this means that the ρ estimates can be directly compared.

The indication is that there is a very significant positive spatial interaction over most of the β range considered, and that the optimal β, on the basis of the maximized log likelihoods, is close to 1.0. Table 16.3 on the whole supports the hypothesis of a significant quadratic function, with the slightly higher log likelihood using the standardized *W* suggesting that the data indicate a slight preference for relative rather than absolute distances. With $\beta = 1.0$, the (unrestricted) log likelihood (L_U) equal to -1033.5740 is significantly above the restricted ($\rho = 0$) log likelihood (L_R) of the OLS model (-1040.2758). Under the null hypothesis that ρ is equal to 0, the quantity $2(L_U - L_R)$ is distributed as χ_1^2, giving a critical value of 3.84 with a 0.05 Type I error rate. However 2(-1033.5740+1040.2758) = 13.40, leading to rejection of the null.

Table 16.3 ML estimates of equation (16.2) (standardized *W* matrix)

	$\beta = 0.1$	$\beta = 0.5$	$\beta = 1$	$\beta = 1.5$	$\beta = 2$	$\beta = 2.5$
c	-1554.897	-2107.976	-1708.666	-890.890	-902.473	-1048.610
	-0.43	-0.35	-0.28	-0.27	-0.33	-0.42
b	1.092	1.171	1.235	1.239	1.231	1.215
	3.49	3.70	3.77	3.66	3.61	3.57
a	-1.4954E-05	-1.8258E-05	-2.2538E-05	-2.3690E-05	-2.2989E-05	-2.1751E-05
	-1.81	-2.22	-2.70	-2.76	-2.63	-2.47
ρ	0.6870	0.8340	0.8430	0.6490	0.4750	0.3640
	3.112	7.166	8.137	4.710	3.791	3.150
σ	9697	9386	9027	9023	9173	9327
loglikelihood	-1039.1475	-1036.5272	-1033.5740	-1033.9498	-1035.3288	-1036.5437

Although the initial model estimates of table 16.1 did not provide evidence of a quadratic relationship, we have shown that this was due to the model being badly specified, with significant unmodeled residual spatial autocorrelation. Modeling the error dependence produces strong empirical support for the quadratic function.

Moreover, the quadratic relationship appears to be a more pronounced feature of the data on eliminating one aberrant case. According to the diagnostics for the OLS model, the most prominent outlier is Luxembourg, which is a highly influential case as indicated by the value of Cook's statistic (equal to 11.9591), due to a very large standardized residual (4.8372) combined with a high leverage. Normally with OLS regression we can either remove the case in question from the data set, or equivalently eliminate it by fitting a case-specific dummy *D* (with 1 in the cell corresponding to the case to be eliminated, and zeros elsewhere). The effect of doing this is to account completely for the outlier, since the residual for the case always turns out to be zero. This in effect means that the case has been removed from the data set, since it is completely explained by the estimated coefficient d for the dummy *D*. For the outlier or aberrant case, it does not matter what the value of the dependent variable is, since it is always completely accounted for by the estimated case-specific dummy, so that if one were to substitute an alternative value of the dependent variable it would make no difference to the parameter estimates, standard errors or fitted values of the overall model.

The presence or absence of Luxembourg in the data does make a difference to the estimates obtained for the model, highlighting its influence in the overall fit. The OLS estimates of table 16.4 indicate that the presence of the dummy *D* (with coefficient d) produces a much better fitting model, since it completely accounts for the very significant positive residual, which goes to zero. Nevertheless the OLS model remains badly specified, on account of the residual autocorrelation that is a consistent feature of the data and which is not eliminated by modeling the outlier. Using the standardized *W* matrix as defined in equation (3) with $\beta = 1.0$, the expected value of Moran's I is -0.01393 and the variance is 0.0004085, which gives a standardized value equal to 6.521. This results is reproduced by the empirical randomization (100 replications), which have mean equal to -0.01294, variance equal to 0.0004201, and standardized value equal to 6.383. In addition, the LM test of error autocorrelation gives a test statistic equal to 41.10, which has a p-value of 1.4473E-10 in the χ_1^2 distribution. The LM test for an omitted spatial lag gives a test statistic of 35.20 with a p-value equal to 2.9683E-09 in χ_1^2. These LM tests therefore provide some empirical support for the spatial error rather than the spatial lag specification.

In table 16.4, we give the parameter estimates of the spatial error models including the dummy variable *D*, showing the best-fitting models across the range of β values. For the *W* standardized model, this results in a switch from $\beta = 1.0$ to $\beta = 1.5$, since this produces a higher likelihood and lower $\hat{\sigma}$ than obtained using other β values. The presence of the dummy variable *D* in the spatial error models does not have precisely the same effect as with the OLS specification, although its presence does reinforce the significance of the quadratic function, as is apparent from the table 16.4 estimates. The presence of D in the spatial error model serves to pull the observed and fitted values closer together, but does not make them equal as in the OLS specification. The reason for this is the across country spillover effect embodied within this model.

The spillover across countries becomes apparent if we examine the Durbin representation of the spatial errors model, which in matrix form is

$$\begin{aligned} &\Delta w = \rho W \Delta w + Xf - WX\gamma + \varepsilon \\ &\varepsilon \sim N(0, \sigma^2 I) \end{aligned} \tag{16.4}$$

in which $W\Delta w$ is the *n* by 1 vector from the matrix product of *W* and Δw, *X* is the *n* by 4 matrix with columns equal to the exogenous variables (including the unit vector and D, *f* is the 4 by 1 vector of coefficients (*a*, *b*, *c*, d), *WX* is the *n* by 4 matrix resulting from the matrix product of *W* and *X* and γ is the vector of coefficients. The spatial Durbin form makes explicit the parametric constraints that are involved in the spatial error model, since both are equivalent so long as $\gamma = \rho f$. In contrast to equation (16.4), equation (16.5) highlights the fact that the fitted values from the spatial Durbin with the appropriate parameter constraints simply equal the matrix product of *X* and the estimated *f*.

$$\widehat{\Delta w} = (I - \hat{\rho}W)^{-1}(X\hat{f} - \hat{\rho}WX\hat{f}) = X\hat{f} \qquad (16.5)$$

and only if $\rho = 0$ would we find that $\Delta w - \widehat{\Delta w} = 0$ for the outlier.

Table 16.4 Results of introducing a case-specific dummy

	OLS	W non-standardized β= 1.0	W standardized β= 1.5
c	-1986.0	-1757.588	-1458.585
	-1.09	-0.90	-0.40
b	1.308	1.372	1.550
	4.57	4.83	5.45
a	-0.00002423	-2.7855E-05	-3.4296E-05
	-3.14	-3.69	-4.76
d	51862.0	49348.444	49698.411
	5.54	5.98	6.90
ρ		0.1880	0.7680
		5.541	7.289
σ	8737.0	8171	7337
loglikelihood	-1026.4391	-1023.2957	-1014.9346

To summaries, from table 16.4, on eliminating the influential case, the OLS estimates support the hypothesized quadratic function, since parameter a is significantly below zero. However the OLS model, with R^2 equal to 0.4658, is clearly inferior to the spatial econometric models, which both have significantly higher log likelihoods reflecting the fact that the ρ estimate is significantly different from zero, and for these better specified models, the presence of the case-specific dummy is not necessary to maintain the quadratic hypothesis, although treating Luxembourg as a special case does strengthen the models for the remaining countries.

16.5 Implications of the Model

The assumption of a quadratic function has long run implications for globalization, but to obtain a clear picture of the quadratic relationship between $\widehat{\Delta w}$ and *w,* one might first consider whether it is necessary to eliminate other effects on $\widehat{\Delta w}$ due to spatial interaction. However unlike the analysis in Fingleton, *et al.* 2005),[1]

---

[1] This analysis in Fingleton, Igliori and Moore(2005) focuses on economic clusters within Great Britain and involves a space lag specification, which we denote by $Y = \rho WY + Xb + e$, in which WY is the endogenous space lag. It follows that $\hat{Y} = (I - \hat{\rho}W)^{-1}X\hat{b}$, and assuming $|\rho| < 1$ this means that

$\hat{Y} = (\sum_i \hat{\rho}^i W^i)X\hat{b} = X\hat{b} + \hat{\rho}WX\hat{b} + \hat{\rho}^2W^2X\hat{b} + \hat{\rho}^3W^3X\hat{b} + ...$ with i going from 0 to infinity and in which $W^0 = I$, W^2 is the matrix product of W and W, and W^i is the matrix product of involving the space lag specification, in this paper the fitted values from the space error model require no additional adjustment. As equation (16.5) implies, the presence of the spatial effects does not distort the quadratic relationship between the unadjusted values $\widehat{\Delta w}$ and w, and this is also evident from plotting $\widehat{\Delta w}$ against w as in fig. 16.1.

v

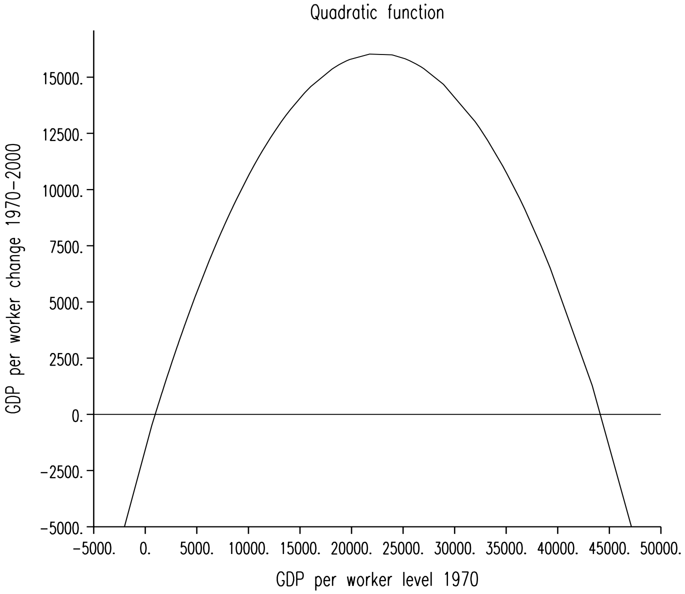


Fig. 16.1. Relationship between GDP per Worker in 1970 and Change 1970-2000

Fig. 16.1 illustrates graphically what we know mathematically, that there is a solution to the quadratic with two roots (which would be coincident if $\hat{b}^2 = 4\hat{c}\hat{a}$), since $a \neq 0$ and $0 \leq \hat{b}^2 - 4\hat{a}\hat{c}$. Using the standardized W matrix with $\beta = 1.5$ and the ML estimates given in Table 4, solving for the roots using:

$$x = \frac{-\hat{b} \pm \sqrt{\hat{b}^2 - 4\hat{a}\hat{c}}}{2\hat{a}}$$

yields the points, (961.5,0) and (44233,0).

The long-run dynamics implied by the model are that each country will gravitate to either the stable upper root or to zero GDP per worker (!), unless it is exactly at the unstable lower root. In other words, countries with GDP per worker higher than 44233 will experience falling GDP per worker until they reach the

---

W^{-1} and W. This shows that the fitted value for the change in Y in any one area depends on the Xs in the other areas also, so to obtain the quadratic function the fitted values have to be adjusted by subtracting these spillover terms.

stable equilibrium point at zero GDP per worker change. GDP per worker in countries below the unstable equilibrium of 961.5 will see GDP per worker fall towards zero. Countries between the two roots experience positive GDP per worker change up to the stable equilibrium point. Therefore, in principle, we have a bipolar process with convergence of some countries to the same upper root, while some poorer countries are inexorably left behind. In practice all but two (Malawi and Guinea-Bissau) of the 98 countries lies within between the two roots, and therefore the conjectured paths for each country, assuming no extraneous factors as for example in the case of Luxembourg, are as in fig. 16.3. The country with the highest level of GDP per worker in 1970 (equal to 43346) is Switzerland. This lies marginal below the stable upper root, and therefore experiences practically no change in GDP per capita level. Unfortunately Guinea-Bissau's and Malawi's GDP per worker falls inexorably to zero.

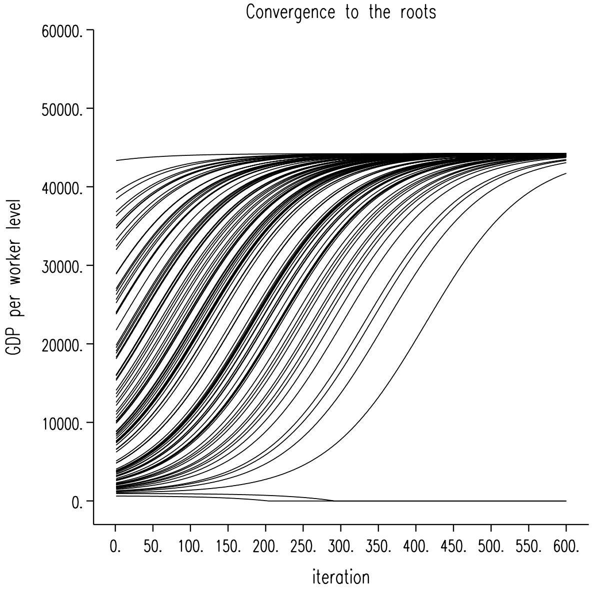


Fig. 16.2. Convergence to the roots.

The very simple spatial econometric model that has been estimated has some dynamic implications that are not inconsistent with the theoretical analysis of the globalization process. For instance, the functional form and estimated parameters indicate that the highest wage countries experience a slow and minimal upward drift in wage levels, while most low wage economies will see quite sharply rising wage rates at some stage of their development, leading ultimately to a long-run stable equilibrium at which wage rates (or GDP per worker) are constant across countries. In common with the theoretical analysis, it is apparent from fig. 16.2 that some countries will reach the equilibrium level earlier than others, while some very low wage level countries will be very slow to respond and only reach the equilibrium wage level at some distant point in the future. This is suggestive of a globalization process that takes time to impact upon those economies that lack the forward or backward linkages that are characteristic of more mature and developed economies, and which provide the external effects that counteract rising wage levels. However, as time passes, even the initially very un-industrialized economies do attract some mobile capital, and demand for labor starts to rise, along with the number and variety of upstream suppliers and downstream markets that sustain growth even with rising wage levels.

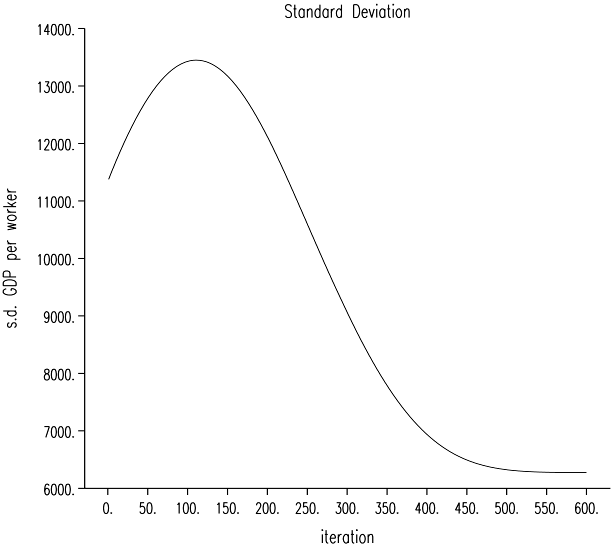


Fig. 16.3 Standard deviation.

One feature of the empirical model that is not evident from the theoretical globalization story is the downward shift in GDP per worker levels in countries (in this data set Guinea-Bissau and Malawi) that are initially below the lower equilibrium. In their case, there is no wage level that will prove attractive to mobile capital and they go into free fall! This is a rather depressing scenario, but in reality, notwithstanding the experience of many countries on the continent of Africa, it is hard believe that this will be the case. In reality it is hoped that the roots to which each country converges will also be influenced by transfer payments and various form of aid from international agencies, so that the roots become:

$$x = \frac{-\hat{b} \pm \sqrt{\hat{b}^2 - 4\hat{a}\hat{c}}}{2\hat{a}}$$

$$\hat{c} = g_0 + g_1 X_1 + g_2 X_2 + \dots g_v X_v$$

in which x and $\hat{c}$ are non-constant vectors which are determined by the set of variables $X_1 \ldots X_v$ (see Fingleton, *et al.* 2005) that reflect the quantities transferred from the developed economies.

16.6 The Impact of Shocks

One feature of the global economy is the way in which shocks to one economy have consequences in other economies. This spatial interaction is directly embodied within the very simple spatial error model that has been estimated here, since from equation (16,2) we have

$$\begin{aligned} \Delta w &= Xf + (I - \rho W)^{-1} \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I) \end{aligned} \tag{16.6}$$

This implies that a shock emanating from one economy will, via the inverse spatial transformation, impact all other countries, with the magnitude of the impact being determined by the size of the initial disturbance, value of the parameter ρ, by the assumed structure of the W matrix.

Note that for $\rho < 1$:

$$(I - \rho W)^{-1} \varepsilon = (\sum_{i=0}^{\infty} \rho^i W^i) \varepsilon = \varepsilon + \rho W \varepsilon + \rho^2 W^2 \varepsilon + \rho^3 W^3 \varepsilon + \ldots$$

in which $W^0 = I$, W^2 is the matrix product of W and W, and W^i is the matrix product of W^{i-1} and W. The effect of shock are therefore felt directly within each country receiving a shock, and there is an indirect effect due to $\rho W \varepsilon$ which affects only those countries linked via the W matrix (i.e. for which there is a non-zero element on the W matrix). If W was a contiguity matrix we might think of these as local effects. However the effect of a shock is global, since it is transmitted also to third party countries that are 'neighbors of neighbors.' The effect via the higher powers of W is also felt in countries that are not linked via non-zero elements of the W matrix. Note that the effect is not one-way. A shock to a country affects the neighbors, and the non-neighbors, but these also affect the country from which the shock emanates. In other words, the full effect of a shock for country k is not simply the shock itself, but the initial shock plus the feedback from the other countries.

This can be illustrated in a simple 9 country world, arranged on a 3 by 3 lattice, with a standardized contiguity matrix based on table 16.5.

Assume a random shock for country k = 1 so that $\varepsilon_1 = -1.2423$, $\varepsilon_{2 \ldots 9} = 0$. Assume $\rho = 0.5$, in which case the vector $(I - \rho W)^{-1} \varepsilon$ = -1.3678, -0.2510, -0.0690, -0.2510, -0.0690, -0.0251, -0.0690, -0.0251, -0.0125. The initial shock ε_1 is not the same as the final outcome for country k (-1.3678) as a result of the feedback effects to *k* with global interaction so that all countries see an impact, regardless of whether or not they are neighbors.

Table 16.5 A W matrix with Weights Given by Rooks contiguity on a 3 by 3 Lattice

	1	2	3	4	5	6	7	8	9
1	0	1	0	1	0	0	0	0	0
2	1	0	1	0	1	0	0	0	0
3	0	1	0	0	0	1	0	0	0
4	1	0	0	0	1	0	1	0	0
5	0	1	0	1	0	1	0	1	0
6	0	0	1	0	1	0	0	0	1
7	0	0	0	1	0	0	0	1	0
8	0	0	0	0	1	0	1	0	1
9	0	0	0	0	0	1	0	1	0

The vector Δw is completely explained by the model plus shocks, but shocks are unexplained and unpredictable so that shocks other than the ones actually experienced could have occurred, in which case we would have observed the vector $\Delta w'$ rather than Δw. We can generate shocks and simulate vector $\Delta w'$ on the assumption that the shocks were as generated and this ability to simulate outcomes provides a useful analytical tool. If we were to follow the same lines as the within-country or within-EU regional analysis of Rey and Montouri (1999) and Le Gallo *et al.* (2003) ideally we would examine the effect of an extraordinary shock to the economy of country *k*, keeping all other shocks at their estimated values, in order to show how the extraordinary country-specific shock affects other countries. Define the vector of estimated shocks as:

$$\hat{\varepsilon} = (I - \hat{\rho}W)(\Delta w - X\hat{f}) \tag{16.7}$$

and then make the country k shock extraordinarily large by adding 3 times the estimated standard error (from Table 4), so that:

$$\hat{\varepsilon}'_k = \hat{\varepsilon}_k + 3\hat{\sigma}$$
$$\hat{\varepsilon}'_j = \hat{\varepsilon}_j ; j \neq k$$

The simulated values are therefore:

$$\Delta w' = X\hat{f} + (I - \hat{\rho}W)^{-1}\hat{\varepsilon}' \tag{16.8}$$

and we look at the effect of the shock to country *k* by comparing the vectors Δw, which simply contains the observed values of the dependent variable, and $\Delta w'$. With this in mind, but with a difference in methodology, the analysis below focuses on the impact of shocks to three leading economies, the USA, UK and Japan. We look at the absolute and relative magnitudes of the shock-effects as they impact both the country of origin and other countries.

The reason why the method actually employed differs from that already described is primarily because of data limitations, which make a straightforward 98-country simulation unacceptable. The estimates in table 16.4 are based on 98 countries for which the Penn World Tables provide data, and it is true that these 98 are, on the whole, the leading economies, but nevertheless some large countries such as Russia and Germany are omitted from the analysis since there is a lack of data for both the start and end years of the period 1970-2000. For a convincing global simulation analysis, one would envisage shocks being transmitted to and from all countries in the global economy, whereas working with only 98 countries we would pick up only the partial effects of any transmission that would in reality occur.

Looking at the snapshot of data for the year 2000 in the Penn World Tables, we find data for 130 countries, and using this cross section of GDP per worker levels we carry out a simulation of the impact of a shock to economy *k*. Although 130 countries fall somewhat short of a complete global coverage, nevertheless we feel justified in treating this larger set as the global economy (more or less!).

One strategy for illustrating the 'spatial reach' and magnitude of shocks to the USA, UK and Japan would be to take the estimated coefficients $\hat{\rho}, \hat{f}$ and $\hat{\sigma}$ given in table 16.4, and apply them to the new data based on w for the year 2000, thus enabling a calculation such as equation (16.8) above. The exogenous regressors w, w^2 and D, together with the unit vector, would then form the columns of the 130 by 4 matrix X_c, which is equivalent to X. A new (130 by 130) matrix, which is denoted by W_c , is constructed on the same basis as W (equation 16.3), using great circle distances for all 130 countries.

However, we cannot obtain an estimated residual vector using the equivalent of equation (16.7) because there are no data on Δw , which is of course why the additional 32 countries were not included in the estimation at the outset. In order to proceed, we might consider constructing a vector of disturbances ε_c by selecting at random from the assumed N(0, $\hat{\sigma}^2 I$) distribution, so that

$$\begin{aligned} \Delta w_c &= X_c \hat{f} + (I - \hat{\rho} W_c)^{-1} \varepsilon_c \\ \varepsilon_c &\sim N(0, \hat{\sigma}^2 I) \end{aligned} \tag{16.9}$$

is a vector of virtual observations Δw_c that (possibly) could occur. This vector might then be compared to the vector $\Delta w_c'$ that could occur if country *k* received an extraordinary shock as a result of adding $3\hat{\sigma}$ to the country k disturbance. This choice of magnitude means that we consider the shock not to be a typical realization of the random variable ε_c, but to be an extraneous event with a size that is outside the normal range one would associate with random variation from $N(0, \hat{\sigma}^2 I)$. The assumption is that the extraneous shock is transmitted between countries in the same way as stochastic disturbances, so that we can consider the combined effect of the country k disturbance plus the extraneous shock to be $\varepsilon'_{ck} = \varepsilon_{ck} + 3\hat{\sigma}$ while the remainder of the vector ε'_{cj} is equal to the initial vector of stochastic disturbances ε_{cj}, for $j \neq k$. On this basis the simulated values are given by

$$\Delta w'_c = X_c \hat{f} + (I - \hat{\rho} W_c)^{-1} \varepsilon'_c \tag{16.10}$$

and the difference between Δw_c and $\Delta w'_c$ gives the 'global' impact of the extraordinary shock to economy *k*. However, it is easy to show that the vector of differences $\Delta w'_c - \Delta w_c$ is the same regardless of the random numbers actually chosen for ε_c, and is also independent of X_c, being equal to $(I - \hat{\rho} W_c)^{-1} s$ where $s_j = 0 \ (j \neq k)$ and $s_k = 3\hat{\sigma}$. Therefore the impact on country j of a shock to country *k* depends on the size of the shock and on *j*'s relative position as embodied within W_c. It also depends on the estimate $\hat{\rho}$ and therefore indirectly on the specification given as equation (16.2).

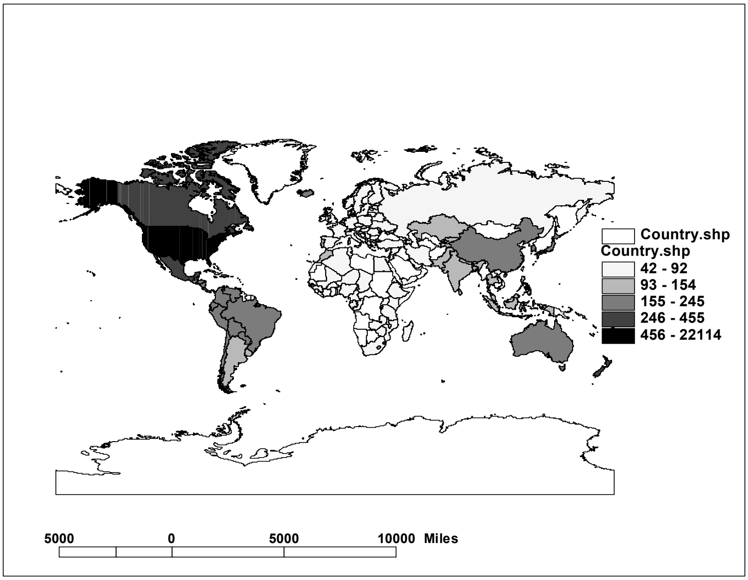


Fig. 16.4 Absolute Impact of a Shock to USA

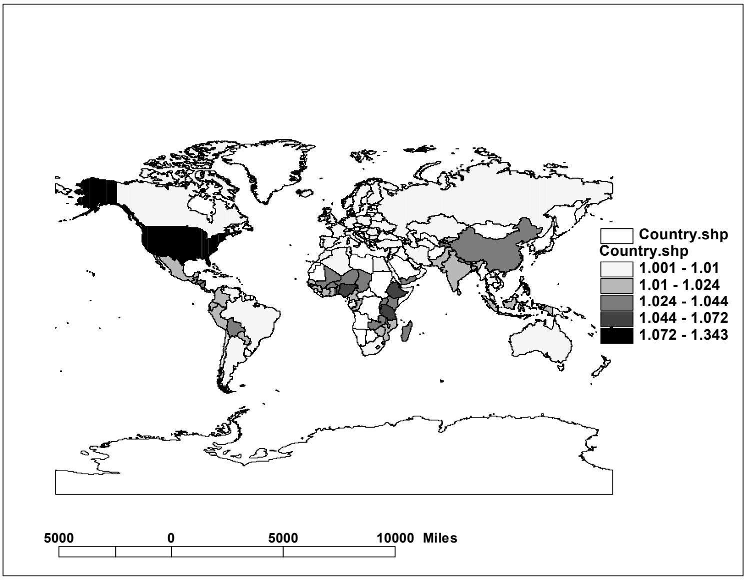


Fig. 16.5. Local impact of a US shock

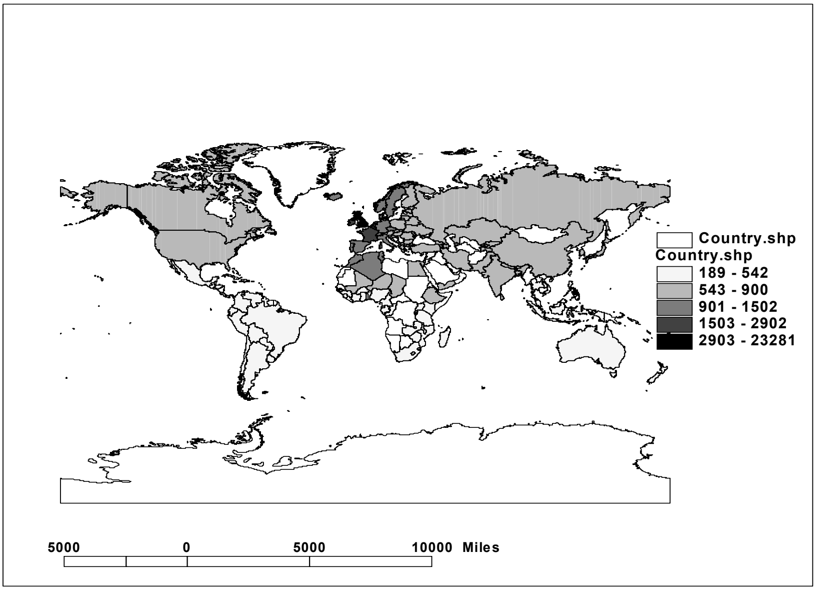


Fig. 16.6. Absolute Impact of a shock to the UK

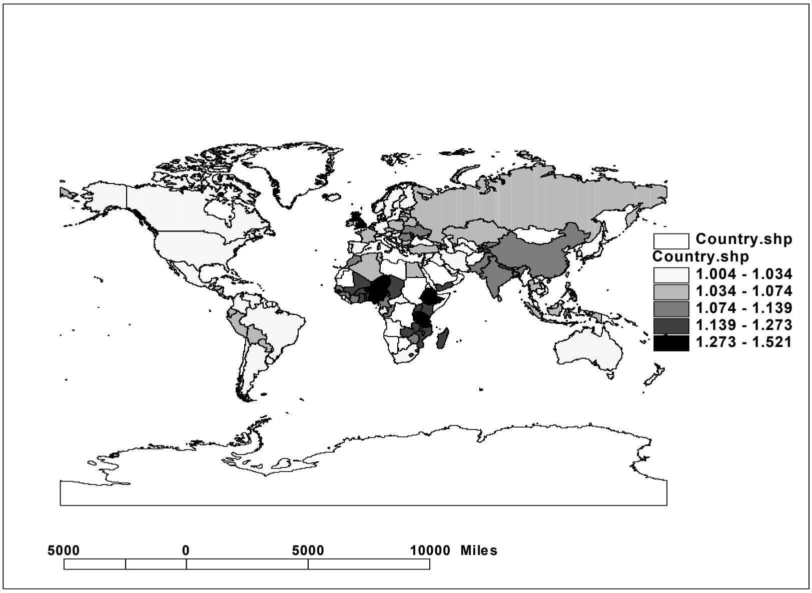


Fig. 16.7 Local impact of a UK Shock

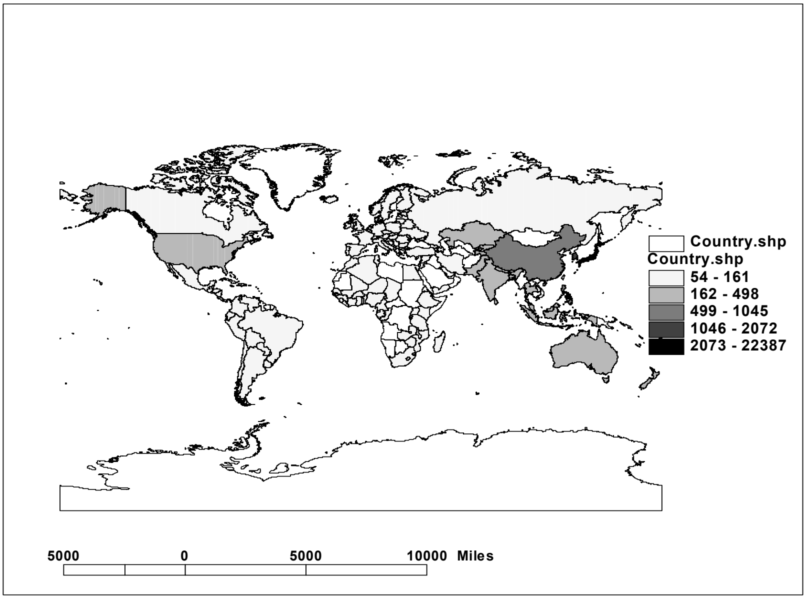


Fig. 16.8. Absolute impact of a shock to Japan.

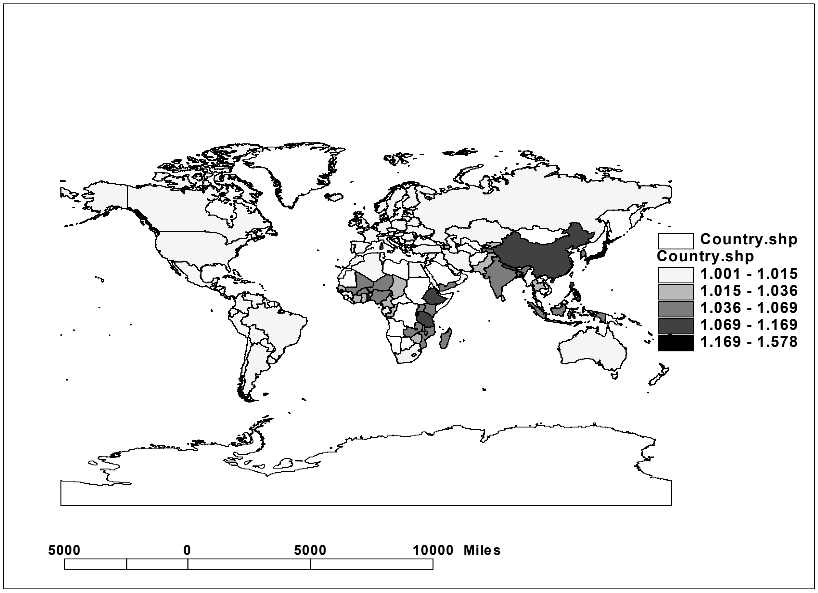


Fig. 16.9 Local impact of a Japan shock

A shock to the economy of the USA equal to $3\hat{\sigma}$ gives a value $\Delta w_c' - \Delta w_c =$ 22113.6 in the USA, which amounts to 61.86% of the total effect of 35748 obtained by summing across all countries. By contrast, Canada receives only 1.27% of the total effect, and Mexico 1.08%, with all other countries receiving lesser amounts. For example, Britain receives only 0.21% of the total global effect of a shock to the US economy. Fig. 16.4 shows that most of the impact of the shock is felt in the Americas and in the Far East and Australasia, with minimal effects in Europe and Africa. However, when we examine the local impacts, obtained by calculating the percentage increase in GDP per worker in each country as a result of a shock to the US economy, a different picture emerges, as shown by fig. 16.5. For the USA, the (arbitrarily sized) shock raises the year 2000 level of GDP per worker by 34.26%, but also has quite sizeable local effects in poor African, Asian and Latin American countries also. For example, the shock to the US economy increases GDP per worker by 2.69% in China, by 5.18% in Nigeria, and by 7.20% in Tanzania.

The global effects of a shock of equal magnitude emanating from the UK economy turns out to be much larger and far reaching than one located in the US, with $\Delta w_c' - \Delta w_c =$ 23280.6 in the UK only amounting to 21.57% of the total effect summing across all countries, which is equal to 107884. The difference between the impacts of the US and the UK is entirely attributable to the differences in their locations, and we see from fig. 16.6 a much more global spread of large effects from the UK shock, but with nearby countries such as Ireland (2.69%), France (2.47%), Belgium (2.38%), Netherlands (2.35%), Switzerland (1.39%), Spain (1.17%) and Germany (1.09%) receiving relatively large shares of the total effect, and of course these big effects will also have knock-on effects to other countries. Figure 16.7 shows the local impacts, again highlighting the sensitivity of low-income African and Asian countries. While the net effect of the initial shock amounts to 52.14% of year 2000 GDP per worker in the UK, it is equal to 44.02% in Tanzania, 34.63% in Nigeria and 9.40% in China. However it should be recognized that a shock of a fixed magnitude is much less likely for the UK economy that it is for the US economy. For the US, the initial shock of $3\hat{\sigma}$ is equal to 34% of the year 2000 GDP per worker level, whereas for the UK it equals 49%. Rather than being a drawing from the homoscedastic error distribution for ε_c, we can make this assumption because we take the shock to be an extraordinary one relative to that distribution.

A shock to Japan equal to $3\hat{\sigma}$ (57% of year 2000 GDP per worker) gives $\Delta w_c' - \Delta w_c = 22386.8$ as the effect on Japan itself, again taking into account simultaneous feedback effects from other economies, and this amounts to 55.44% of the total overall effect across all countries of 40382. Most of the impact is concentrated in the Far East, as shown by fig. 16.8, although there are noticeable effects in Australasia, the Indian sub-continent and in the US also. South Korea is the country most affected in absolute terms, accounting for 5.13% of the total effect, followed by China with 2.59%, and the Philippines, New Zealand, Australia, Thailand and India all account for more than 1% of the total. In terms of the effect relative to year 2000 GDP per worker in each country, the most noticeable local impacts (fig. 16.9) are on Japan itself (57.79% of year 2000 GDP per worker), on China (16.92%), and on a range of low income African and South East Asian countries.

16.7 Conclusions

This Chapter has attempted to link globalization to spatial econometric analysis by estimating a very simple empirical model that has features in common with the main predictions of globalization theory, as expounded by Krugman and Venables (1995) and Fujita *et al.* (1999). The empirical model indicates the presence of a quadratic relationship between the change in GDP per worker over the period 1970 to 2000 and the level of GDP per worker in the year 1970. The model has features not atypical of the principal predictions of globalization theory, particularly increasing GDP per worker or wage level disparity as globalization takes effect and a core-periphery structure emerges, followed by convergence as capital migrates from the leading economies to lower cost countries. However the model also predicts that not all countries will converge to the same stable equilibrium at which wage levels will be equalized. The existence of a quadratic relationship implies that there are roots and that, depending on a country's initial position with respect to the roots, it will converge on a stable upper root or indeed go into free fall, with no wage level low enough to attract capital and start a process of upward movement in wage levels. It appears that there are countries in this position, and may need rescuing.

The chapter also focuses on spatial effects, which are evidently an important aspect of the data. We have chosen to model these as stochastic disturbances, which are transmitted simultaneously between economies on a global basis. There are important implications of simultaneous spatial interaction for the globalization process, and these have been illustrated by an analysis of the impact of extraordinary shocks to three leading economies, the US, UK and Japan. This shows that the more central and connected location of the UK within Europe results in larger and more widespread shock effects across the globe, while at the other extreme the impact of a shock to the US economy of the same magnitude has a lesser effect overall, focused mainly but not exclusively on nearby countries and the Americas. Of course this does not mean that a shock to the US economy is less likely to have global repercussions, particularly when one examines the impacts relative to GDP per worker levels, and also it seems reasonable to assume that an extreme shock of a given magnitude would be more likely to emanate from the US economy that from smaller and less wealthy economies.

References

Anselin L. (1988) Spatial econometrics : Methods and models. Kluwer, Dordrecht

Castells M (1989) The informational city: Information technology, economic restructuring, and the urban regional process. Basil Blackwell, Oxford

Castells M (2000) The rise of the network society (2nd ed) Blackwell, Oxford

Coe DT, Helpman E (1995) International R&D spillovers. European Economic Review 39:859-887

Fingleton B (1999) Spurious spatial regression: Some Monte-Carlo results with a spatial unit root and spatial cointegration. Journal of Regional Science 39:1-19

Fingleton B, Igliori DC, Moore B (2005) Cluster dynamics: New evidence and projections for computing services in Great Britain. Journal of Regional Science 45:283-311

Fujita M, Krugman P R, Venables A (1999) The spatial economy : Cities, regions, and international trade. MIT Press, Cambridge, Massachusetts

Krugman PR, Venables A (1995) Globalization and the inequality of nations. Quarterly Journal of Economics 60:857-880

Le Gallo J, Ertur C, Baumont C (2003) A spatial econometric analysis of convergence across European regions, 1980-1995. In: Fingleton B (ed) European regional growth. Springer-Verlag, Berlin

Neary JP (2003) Globalization and market structure. Journal of the European Economic Association 1:245-271

Peeters J, Garretsen H (2000) Globalisation, wages and unemployment: An economic geography perspective, CESifo working paper no. 256. CESifo GmbH Center for Economic Studies, Munich

Rey SJ, Montouri BD (1999) US regional income convergence: A spatial econometric perspective. Regional Studies 33:143-156

Sassen S (1994) Cities in a world economy. Pine Forge Press, Thousand Oaks, CA

Upton GJG, Fingleton B (1985) Spatial data analysis by example, Vol. 1. Wiley, Chichester

Verspagen B (1997) Estimating international technology spillovers using technology flow matrices. Weltwirtschaftliches Archiv 133:226-248

17 Risk and Growth: Theoretical Relationships and Preliminary Estimates for South Africa*

Russel J. Cooper[1] and Kieran P. Donaghy[2]

[1]Department of Economics, Macquarie University, Sydney, NSW Australia
[2]Department of City and Regional Planning, Cornell University, Ithaca, New York, USA

17.1 Introduction

The issue of the costs of economic instability relative to the costs of reduced growth is one which has created considerable controversy both politically and amongst the ranks of economists. The publication of Lucas's (1987) calculations, which claimed to show an enormous relative welfare benefit in favor of promoting growth rather than stability, has been accepted in some quarters as a strong justification for small government. On the other hand, a significant but much less influential rearguard action continues to be fought. At the level of macroeconomic research, Ramey and Ramey (1995) present evidence on a negative link between volatility and growth. They conclude that:

> "..... by assuming no interaction between volatility and growth, the theoretical business cycle and growth literatures omit important elements. These omissions can lead to questionable conclusions, such as Lucas's (1987) calculation of the potential benefits of eliminating business-cycle volatility." (Ramey and Ramey (1995) p. 1148).

The Lucas calculation was based upon a "microfoundations" or "optimizing" approach to macroeconomics, however, and it would be rather easy in this context to consign the Ramey and Ramey results to the fate of those aggregative studies which do not have the status of explicit microeconomic underpinnings. On the

---

* This chapter is a revised version of a paper presented to Econometric Society World Congress, Seattle, Washington, USA, August 11-16, 2000.

other hand, work of Obstfeld (1994) set in a similar "representative consumer" microfoundations context to that of Lucas, lends weight to the view that the issue should be regarded as far from settled. Obstfeld concludes that:

> ".....the cost of U.S. consumption variability is substantially higher than in Lucas (1987), although still quite small compared with the benefit of an extra annual percentage point of trend consumption growth". (Obstfeld (1994), p.1472),
>
> and
>
> ".....application of this paper's framework to other countries yields many instances, especially in the developing world, of much higher variability costs than those found for the United States". (Obstfeld (1994), p.1473).

Interestingly, Obstfeld's results arise, at least in part, not from conjecturing a link between volatility and growth but from actually breaking a rather rigid link between the microeconomic influences of volatility and growth which is present in the great majority of the microfoundations approaches - namely the link between a measure of the representative consumer's willingness to forgo current consumption for the sake of growth (typically constructed as the intertemporal elasticity of substitution) and a measure of risk aversion (usually defined in terms of the Arrow-Pratt measure of relative risk aversion and often applied to what is termed "consumption risk"). By breaking the rigid link between aspects of consumer preferences which dictate willingness to substitute current for future consumption and other aspects which dictate willingness to accept variability in circumstances, the data may be allowed to play a greater role in deciding the issue of the relative importance of volatility and growth. The difficulty, however, is to break the link realistically - from an empirical viewpoint - while maintaining the microfoundations approach.

Obstfeld's choice of the non-expected utility maximizing framework, based on the specification of Weil (1990) and the theoretical developments of Epstein and Zin (1989, 1991), is an explicit recognition of the need to break the nexus between the intertemporal elasticity of substitution and the coefficient of risk aversion. In this paper an alternative approach to breaking this link is pursued, an approach which generalizes the functional forms underlying the specification of preferences and technology. It is shown that, if the focus for measurement of risk aversion is made consistent with the microeconomic optimization problem of an intertemporal utility maximizing consumer and if a natural interpretation of the concept of intertemporal substitutability is adopted which is not dependent on a particular constant parametric specification, then it is not necessary to depart from the intertemporally additive expected utility paradigm to break the rigid link between risk aversion and intertemporal substitutability. However, if one is to work within this paradigm it is necessary either to generalize the functional form of the period utility or felicity function or to generalize the functional representation of technology. This suggests that explicit emphasis be placed on the issue of functional form specification, an approach which seems warranted in view of the potential sensitivity of volatility versus growth welfare calculations to the specification of the functional form of utility and technology.

In the introduction to his book *Methods of Macroeconomic Dynamics*, Turnovsky (1995) describes the process of increasing sophistication in the development of microeconomic foundations for macroeconomic models. This has become apparent in the increasing attention to logical detail which is given in the specification of the objectives, choices and constraints facing consumers, firms and government. The environment of analysis has moved from an earlier paradigm in which agents operated in an atemporal closed economy context, with intertemporal aspects tacked on in an *ad hoc* manner, to a fully intertemporal optimizing framework in a deterministic context, with increasing attention being paid to open economy choices and constraints, and more recently into an explicitly stochastic environment which introduces options for studying a range of real world phenomena previously treated in a cursory manner.

Naturally, the developments in the attention of macroeconomics to microeconomic foundations have not been achieved without cost. In those cases where explicit closed form solutions for agents' optimal behavior have been derived, these have typically been with the aid of simplifying assumptions both on preferences and technology. However, for analyses supported by empirically based parameter values, general ("flexible") functional form specifications are desirable. The point of departure of the current paper is therefore to bring together theoretical developments associated with the modeling of representative agents in a continuous time stochastic environment, along the lines expounded by Turnovsky, with specifications of preferences and technology more general than the typical linear or linearly homogeneous technology and isoelastic preferences. The objective is to provide a theoretical basis for specification of empirically robust macroeconomic models based upon microeconomic foundations in which there is no *apriori* rigid link between concepts of critical relevance to distinguishing the relative importance of volatility and growth.

The plan of this chapter is as follows. In Section 17.2 the relationship between intertemporal substitutability and risk aversion is considered in more detail in a stochastic intertemporal expected utility maximizing framework. It is shown that the key to maintaining flexibility in this relationship is the modeling of either sufficiently complex preferences or sufficiently complex technology, or preferably both, so as to imply a non-unitary wealth elasticity of consumption. Since specifications of this level of generality are rarely employed, Section 17.3 is devoted to summarizing a method which allows the derivation of consumption functions of this type for which the corresponding utility functions can be explicitly evaluated. The approach is based on intertemporal duality and the specification of consumer profit functions and is an extension to the stochastic case of a technique detailed in a deterministic context in Cooper (2001). Section 17.4 sets out a two country model using generalized preference and technology specifications, while Section 17.5 elaborates a specification of the home-country component of the model for particular specifications of technology and preferences. Section 17.6 begins to employ the model framework in a preliminary application to South African data.

17.2 The Relationship Between Intertemporal Substitutability and Risk Aversion

Consider a representative consumer with an instantaneous von Neumann-Morgenstern utility function $U(z)$ where $z > 0$ is real total consumption expenditure. Following Pratt (1964) and Arrow (1965), for a consumer who faces consumption risk, the expression $-zU_{zz}/U_z$ may be used as a measure of relative risk aversion. However, in the Arrow-Pratt framework it is because the consumer uses the function $U(z)$ as an evaluation of the optimized objective that $-zU_{zz}/U_z$ has a natural interpretation as a relative risk premium which a consumer would be willing to pay to avoid an undesirable risk. That is, the Arrow-Pratt measure applied to $U(z)$ presupposes that $U : \mathbb{R}^+ \to \mathbb{R}$ represents the consumer's optimized, indirect, utility function so that the consumer, constrained by a random initial resource, $z > 0$, uses the function U as a basis for welfare evaluation.

Suppose, however, that the consumer is an intertemporal optimizer with an intertemporally additive preference ordering. The consumer chooses $\{z(t)\}_0^\infty$ in order to maximize the expected discounted present value of the stream of $U(z(t))$ over all future t, conditional on a given initial wealth, k. Let $J(k,\bullet)$ denote this optimized value. Typically J depends upon parameters of the process of evolution of k as well as upon the choice of the path of z. Applying the Arrow-Pratt reasoning to risky initial wealth suggests that $-kJ_{kk}/J_k$ is an appropriate measure of relative risk aversion. If other sources of risk are involved, say in the returns on stored wealth, these could be measured by the dependence of J on the parameters defining the risks.

While risky initial wealth may lead to volatility in consumption, it is nevertheless clear that $-kJ_{kk}/J_k$ is a more appropriate measure of risk aversion than $-zU_{zz}/U_z$ in the intertemporal case. There are at least two reasons for this. Firstly, in intertemporal models z is actually chosen by the consumer so that fluctuations in z are endogenous, not exogenous. Consequently, z is not "risky" in the same sense that k is. That is, while at each point in time the consumer is able to choose z, in the case of k the consumer has no instantaneous choice but must accept the contemporaneous value of the resource as given. Secondly, in the context of an intertemporal optimization it is the optimal value function J which measures the consumer's optimized satisfaction, not the instantaneous utility function U. Therefore, the premium in terms of the initial wealth resource (here denoted by k) foregone to avoid a possible loss in J has meaning in an intertemporal optimizing context, but the premium in terms of consumption foregone to avoid a possible loss in U does not have the same claim to meaning since U is not maximized in this context nor is z outside the instantaneous control of the optimizer.

Having established that the Arrow-Pratt concept of relative risk aversion applies to the (scaled) curvature of the indirect utility function $J(k,\bullet)$, it should be noted that, for a variety of reasons, one would not expect this concept, measured by the construct $-kJ_{kk}/J_k$, to be represented adequately by a constant parameter. In the first instance, $J(k,\bullet)$ depends not only upon the level of wealth but also upon all the intertemporal and risk factors in the model. There is no reason, a` priori, why these factors could be expected to impact to the same extent on consumers drawn from different locations in the wealth distribution or with different demographic characteristics. This is particularly pertinent to modeling the behavior of a so-called representative consumer, since in this case sensitivity of the risk aversion "parameter" to average wealth, its distribution, and to factors driving the evolution of wealth is likely to be important. Secondly, there is even less justification in specifying this concept as a parameter of a consumer's instantaneous utility function. As a concept related to optimized utility it is likely to be extremely unrealistic empirically to link this rigidly to a parameter of the instantaneous function since this implies that all intertemporal factors cancel out in terms of their potential influence on the consumer's extent of risk aversion.

It should be noted that the specifications of Weil (1990) and Epstein and Zin (1991), were not intended to examine the parameter constancy issue. Their specifications use a functional form in which relative risk aversion is represented by a constant parameter. However, this constant relative risk aversion specification has proven to be empirically problematic. The empirical work reported by Epstein and Zin (1991) illustrates the point that the extension to a constant relative risk aversion non-expected utility formulation does not necessarily resolve specification issues. They note that:

> "A troubling pattern that emerges is that the rate of time preference, δ, is often significantly less than zero ... [which] indicates a problem that this model shares with the [constant relative risk aversion] expected utility model ..." (Epstein and Zin (1991), p.282).

Turning to the concept of the intertemporal elasticity of substitution, let us restrict instantaneous utility function specifications $U(z)$ and the process driving k to those for which an optimal expected utility maximizing decision rule exists in autonomous feedback form. The feedback rule is the solution to the first order necessary condition $U_z(z) = J_k(k,\bullet)$. The first order optimization condition may be inverted either to emphasize the dependence of the solution on the evolution of marginal utility $z = Z^{\lambda}(J_k)$ or to emphasize the feedback solution as a synthesized form $z = Z^k(k,\bullet)$[2]. In the former formulation Z^{λ} is the inverse of the U_z

---

[2] Wherever possible upper case letters are reserved for functions which evaluate the variable represented by the corresponding lower case letter. Additionally, a superscript is

function while its argument J_k serves as a conduit for all intertemporal influences on the consumption choice. It follows that $-\partial \ln Z^{\lambda} / \partial \ln J_k$ is a natural choice for the definition of the intertemporal elasticity of substitution. Since $J_k = U_z(z)$ at the optimal choice of z, this implies that $\partial \ln J_k / \partial \ln z = zU_{zz} / U_z$, which in turn suggests that the intertemporal elasticity of substitution may be defined as $-U_z / (zU_{zz})$. This is the usual formula for the intertemporal elasticity of substitution in a continuous time intertemporally additive expected utility maximizing context, it is unambiguous, and it depends only upon the functional form specified for instantaneous utility. Because it relates purely to the consumer's primitive preference ordering, it is natural to parameterize upon the intertemporal elasticity of substitution. However, it may not be empirically realistic to treat this parameter as constant. More plausibly, intertemporal substitutability would increase with the overall consumption level.

In an intertemporally additive expected utility maximizing context subject to a linear stochastic wealth transition equation, it can be shown that a constant intertemporal elasticity of substitution implies a unitary wealth elasticity of consumption and consequently implies a constant (and reciprocal) coefficient of relative risk aversion. However, it can also be shown that this rigid relationship is purely an artifact of the parameter constancy, that is of the functional form. In terms of the above definitions, the relationship between the intertemporal elasticity of substitution (IES) and the coefficient of relative risk aversion (RRA) in an intertemporally additive utility maximizing model may be examined on log differentiation of the first-order condition $U_z(z) = J_k(k,\bullet)$:

$$\left[-zU_{zz} / U_z\right] \left(\partial \ln z / \partial \ln k\right) = -kJ_{kk} / J_k ,$$

which immediately gives:

Wealth Elasticity of Consumption = IES * RRA

This relationship indicates that the link between intertemporal substitutability and risk aversion (in the senses defined here) is not rigid, but neither is it entirely arbitrary. As shown above, a rigid reciprocal link between the IES and the RRA only applies for those instantaneous utility specifications which imply a unitary wealth elasticity of consumption. Thus it is not necessary to move to non-

---

employed where necessary to distinguish functions by their conditioning variables. Thus Z^{λ} denotes the marginal utility conditioned consumption function (the Euler equation) while Z^{k} denotes the wealth conditioned consumption function (the synthesized solution).

intertemporal additivity or to non-expected utility to break this link but merely to move to a more general functional form which does not impose the *à priori* rigid behavioral restriction of a unitary wealth elasticity of consumption.

Moving to a more flexible specification which does not imply a unitary wealth elasticity of consumption, moreover, offers the potential for a more satisfactory resolution of empirical paradoxes such as the so-called "equity premium" puzzle. Those papers which appear to resolve the equity premium puzzle may be interpreted as introducing factors which allow a non-unitary wealth elasticity of consumption.[3] Those papers which appear unable to resolve the puzzle, even if they break the link between alleged measures of risk aversion and intertemporal substitutability, generally retain a unitary wealth elasticity of consumption.[4]

Consideration of the issues raised above suggests that it would be worthwhile to explore the specification of functional forms which do not imply a unitary wealth elasticity. These forms allow recognition of some of the issues which in the quoted literature are raised as important, such as the requirement for a more flexible but intuitively reasonable relationship between intertemporal substitutability and risk aversion, without abandoning either the intertemporal additivity assumption or the expected utility paradigm.

It has been argued above that the notion of the validity of the concept of consumption risk is questionable in the context of a model in which the consumer is free to choose consumption at any period in time. In its place is the notion of resource risk. In the model to be developed below, resource risk flows directly from investment risk. The model is constructed to allow technology and investment decisions to influence the evolution of the wealth available to the expected intertemporal utility maximizer. However, there remains a fundamental exogenous element of uncertainty associated with the evolution of wealth (or, in the consumer-firm context to be developed below, associated with the evolution of the capital stock). This provides a natural vehicle for examination of the link between risk aversion and intertemporal substitutability and consequently, within a representative consumer-firm context, for consideration of the relationship between volatility and growth.

17.3 Notation, Assumptions and Preliminary Results

17.3.1 The Basic Model

In this section a general methodology is proposed to enable the examination of intertemporal substitutability and risk aversion in a representative consumer/firm context. For the purpose of allowing the theoretical specification to represent the microfoundations of a macro model, the aggregate consumption variable is gener-

---

[3] Constantinides (1990), Grossman and Laroque (1990), Mankiw (1986).

[4] Weil (1989), Attanasio and Weber (1989), Kocherlakota (1990). See also the survey articles by Abel (1991) and Kocherlakota (1996).

alized beyond simply consumer aggregate expenditure to represent the "full" expenditure of a representative consumer/firm, and the transition equation is generalized to allow nonlinear technology and to represent the national income identity. The analysis concentrates on the real open economy. Monetary factors are not considered but the model is developed in an explicit two-country context – a home and a foreign country – where the second country could be interpreted either as a major trading partner with other partners exogenous, or possibly as an aggregate of the rest of the world.

Define $U(z)$ as the home country representative consumer/firm's instantaneous utility function, where z is an aggregate of real consumption, net exports and balancing items, representing total current expenditure in the national income identity. Ignoring minor items such as depreciation and valuation adjustments, essentially $z=c+g+x$ where c is real private consumption, g is real government expenditure and x is net exports. The representative agent derives instantaneous utility not only from current consumption but also from the implications of net exports in terms of foreign debt reduction or foreign exchange accumulation.

$U(z)$ is interpreted as an indirect utility function. Its dependence on the relative prices of c, g and x is suppressed for convenience as attention is concentrated on intertemporal aspects of the optimization problem. The budgeting process is essentially two stage. While optimal c, g and x - conditional on aggregate expenditure z - could be determined by Roy's Identity applied to $U(z)$, the following analysis concentrates on the intertemporal choice between z and investment in the country's capital stock.

It is assumed that $U_z(z) > 0$, $U_{zz}(z) < 0$, where single (double) subscripts denote first (second) derivatives respectively. Let $U^*(z^*)$ denote the foreign country representative agent's utility function, which will be subject to analogous assumptions.

Let k denote the real domestic capital stock and α represent holdings of a risky financial asset, measured in real terms, acquired by sale of ownership rights over a component of the productive capital stock. $F(k)$ denotes a restricted (that is, capital stock conditioned) real net output or real value added function, the average return from which is paid to the home country representative consumer-firm, who owns $k-a$ of productive capital. The foreign real net output function is $F^*(k^*)$. The risky investment for the home country representative agent involves equity investment in the foreign country. This pays an expected average return:

$$\left[q+\tfrac{1}{dt}E(dq)\right]F^*(k^*)/k^*.$$

Here q is the real bilateral exchange rate defined in home currency units per foreign currency unit and E denotes the expectation at the current point in time. The expected change in the exchange rate is modeled as responsive to domestic to foreign interest rate differentials, $r - r^*$, and also to uncertainty, in a manner to be specified more completely below. In general equilibrium, the decisions of the representative economic agents affect the values of the domestic and foreign interest rates, r and r^*, respectively. To characterise the optimization problems it is useful to distinguish between:

(i) a coordinated social optimum, in which a central planner determines the optimal controls for both countries simultaneously and allowing for spillover effects;

(ii) a decentralized social optimum, in which separate central planners determine the optimal controls for each country, allowing for externalities of the type occasioned by the influence of the capital stock on interest rates in their respective countries, but taking the outcome of the other country as given (and consistent with a decentralized social optimum in that country); and

(iii) a competitive optimum, in which separate representative agents make optimization decisions in each country in the assumed absence of an impact of their decisions on interest rates in either country, but taking the outcome of interest rates as given (and consistent with the modeled relationships between interest rates and capital stocks).

Ultimately, the modeled relationships between risk and investment in this paper are to be drawn from an analysis of the competitive optimum (iii). However, the solution method requires initial attention to be paid to specification of the decentralized social optimum (ii).

In the case of optimizations (ii) and (iii), at any given point in time, with k predetermined, the domestic agent, whether a social planner or a competitive agent, chooses optimal current expenditure z and the risky asset portfolio choice a in an environment characterized by knowledge of the current values of k, k^*, r, r^* and q. For both the home and foreign country problems, the optimal value functions and the optimal control functions depend upon the initial values of these variables. An admissible domestic expenditure feedback rule is a function $Z : \mathbb{R}^5_+ \to \mathbb{R}_+$, denoted $z = Z(k, k^*, r, r,^* q)$, chosen from a set of admissible functions $\mathbb{Z}\{k, k^*, r, r,^* q\}$. An admissible portfolio allocation feedback rule for the representative domestic agent is a function $A : \mathbb{R}^5_+ \to \mathbb{R}$, denoted $a = A(k, k^*, r, r,^* q)$, chosen from a set of admissible functions $\mathfrak{A}\{k, k^*, r, r,^* q\}$. Similarly, an admissible expenditure feedback rule for the representative foreign agent is a function $Z^* : \mathbb{R}^5_+ \to \mathbb{R}_+$, denoted by $z^* = Z^*(k^*, k, r^*, r, 1/q)$, and which is chosen from a set of admissible functions $\mathbb{Z}^*\{k^*, k, r^*, r, 1/q\}$. An admissible portfolio allocation feedback rule for the representative foreign agent is a function

$A^*:\mathbb{R}^5_+ \to \mathbb{R}$, denoted by $a^* = A(k^*,k,r^*,r,1/q)$, chosen from a set of admissible functions $\mathfrak{A}^*\{k^*,k,r^*,r,1/q\}$.

The home and foreign countries' representative consumer-firms' optimization problems are linked in the sense that each acts under constraints arising from the value of variables which are determined as the outcome of the optimization problem of the other. It is assumed that in each case the representative agent acts as if it is in a competitive situation in which its decision making does not affect that of the agent in the other country. In fact, this competitiveness assumption will be strengthened to facilitate derivations below.

The optimal value functions implied by the optimizations may be defined by equations (17.1a)-(17.1e) below:[5]

$$J(k,k^*,r,r^*,q) = \max_{\{z\in\mathbb{Z},a\in\mathfrak{A}\}} E_0\int_0^\infty e^{-\delta t}U(z(t))dt\,, \tag{17.1a}$$

$$J^*(k^*,k,r^*,r,1/q) = \max_{\{z^*\in\mathbb{Z}^*,a^*\in\mathfrak{A}^*\}} E_0\int_0^\infty e^{-\delta^* t}U^*(z^*(t))dt\,, \tag{17.1a*}$$

subject to the constraints:

$$\begin{aligned} dk(t) = &\left\{\left[\frac{F(k(t))}{k(t)}\right]\left[k(t)-a(t)\right]-z(t)\right\}dt \\ &+\left[\frac{F^*(k^*(t))}{k^*(t)}\right]a(t)\left[q(t)dt+dq(t)\right], \end{aligned} \tag{17.1b}$$

$$\begin{aligned} dk^*(t) = &\left\{\left[\frac{F^*(k^*(t))}{k^*(t)}\right]\left[k^*(t)-a^*(t)\right]-z^*(t)\right\}dt \\ &+\left[\frac{F(k(t))}{k(t)}\right]a^*(t)\left[\left(\frac{1}{q(t)}\right)dt+d\left(\frac{1}{q(t)}\right)\right], \end{aligned} \tag{17.1b*}$$

---

[5] In the approach to be adopted below, the dependence of the optimal value function and optimal control functions upon both initial values of variables and upon parameters is to be investigated. However, the explicit dependence of these functions on parameters is suppressed for notational clarity. The key parameters which are explicit in the optimizations at this point are σ, δ and δ^*, where σ represents risk (the diffusion term in the exchange rate equation (17.1d)), and δ and δ^* represent the time preference rates of the domestic and foreign representative agents respectively. In initially specifying the structure of dynamic optimizations, the time dependence of variables is explicitly indicated. As far as possible, variables are represented by lower case Roman letters and parameters by lower case Greek letters. Upper case letters are used to indicate functions whose value is the associated lower case letter.

$$dr(t) = R(k(t),k^*(t),r(t),r^*(t),q(t))dt + S(k(t),k^*(t),r(t),r^*(t),q(t))dw(t), \tag{17.1c}$$

$$\begin{aligned} dr^*(t) &= R^*(k^*(t),k(t),r^*(t),r(t),1/q(t))dt \\ &\quad + S^*(k^*(t),k(t),r^*(t),r(t),1/q(t))dw^*(t), \end{aligned} \tag{17.1c*}$$

$$dq(t) = \left[r(t) - r^*(t) + \tfrac{1}{2}\sigma^2\right]q(t)dt + \sigma q(t)d\xi(t) \tag{17.1d}$$

$$k(0) = k\ ,\ k^*(0) = k^*,\ r(0) = r\ ,\ r^*(0) = r^*\ ,\ q(0) = q\ . \tag{17.1e}$$

Let $\tilde{z} = \tilde{Z}(k,k^*,r,r,^* q)$, $a = \tilde{A}(k,k^*,r,r,^* q)$, $\tilde{z}^* = \tilde{Z}^*(k^*,k,r^*,r,1/q)$, and $\tilde{a}^* = \tilde{A}^*(k^*,k,r^*,r,1/q)$ denote the optimal feedback solutions.

In this model, the primary source of uncertainty is associated with foreign exchange risk, emanating from the standard Weiner process $d\xi$. The specification of the exchange rate stochastic transition equation (17.1d) has implications for exchange rate expectations in a rational expectations setting. The particular specification chosen in (17.1d) allows a symmetric expectations process to be assumed for the foreign country representative agent, as will be elaborated below. In (17.1c) and (17.1c*) additional apparent uncertainty is associated with (domestic and foreign) interest rate risk, emanating from the specification of standard Brownian motions dw and dw^*. Later, these sources of interest rate risk will be linked to the one fundamental source of (foreign exchange) risk, through additional assumptions related to the absence of unexploited instantaneous arbitrage opportunities.

To allow a simplification of notation in what follows, a time varying risk premium is defined by:[6]

$$\begin{aligned} P(k,k^*,r,r^*,q) &\equiv \left[q + \tfrac{1}{dt}E(dq)\right]F^*(k^*)/k^* - F(k)/k \\ &= q\left[1 + r - r^* + \tfrac{1}{2}\sigma^2\right]F^*(k^*)/k^* - F(k)/k \end{aligned} . \tag{17.2}$$

In order for the home country to undertake a risky investment in the foreign country when it has a risk free alternative at home, the overall risk premium p must be positive for a risk averse representative agent; otherwise no foreign in-

---

[6] At this point, the explicit reference to time is dropped for notational convenience. In addition to its dependence on variables, the risk premium clearly depends upon the parameter σ. This parameter dependence is suppressed for notational convenience in the notation representing the risk premium as a function $P : \mathbb{R}^5_+ \to \mathbb{R}$, denoted by $p = P(k,k^*,r,r^*,q)$.

vestment will occur, since an equal or greater certain rate of return would be available locally.

From (17.1d), $\frac{1}{dt}E(dq)$, the expected rate of currency depreciation from the point of view of the home country, is given by:

$$\tfrac{1}{dt}E(dq) = \left[r - r^* + \tfrac{1}{2}\sigma^2\right]q\,.$$

The modeling of exchange rate expectations in this way has the advantage that the mirror image expectation formation process may be compatibly assumed for the foreign country. For the foreign country, the exchange rate is defined as $1/q$, and applying Ito's Lemma to (1c) implies:

$$d[1/q] = \left[r^* - r + \tfrac{1}{2}\sigma^2\right][1/q]dt - \sigma[1/q]d\xi \qquad (17.1d^*)$$

so that:

$$\tfrac{1}{dt}E(d[1/q]) = \left[r^* - r + \tfrac{1}{2}\sigma^2\right][1/q]\,.$$

The risk premium from the point of view of the foreign country's investments in the home country is:

$$\begin{aligned} P^*(k^*,k,r^*,r,1/q) &\equiv \left[1/q + \tfrac{1}{dt}E(d[1/q])\right]F(k)/k - F^*(k^*)/k^* \\ &= [1/q]\left[1 + r^* - r + \tfrac{1}{2}\sigma^2\right]F(k)/k - F^*(k^*)/k^*\,. \end{aligned} \qquad (17.2^*)$$

It is possible, due to the effect of uncertainty on exchange rate expectations, that each country can perceive advantages in investing in the other simultaneously. Such cross investment will occur if p and p^* are simultaneously positive. If only one of these is positive, investment at that point in time will be one-way.[7]

Based on the above, the drift in the home country's capital stock may be rewritten by rearrangement of (17.1b), (17.1d) and (17.2) as:

$$\begin{aligned} \tfrac{1}{dt}E(dk) &= \frac{F(k)}{k}\left[k - a\right] + \left[q + \tfrac{1}{dt}E(dq)\right]\frac{F^*(k^*)}{k^*}a - z \\ &= F(k) + P(k,k^*,r,r^*,q)a - z \end{aligned}$$

---

[7] A necessary and sufficient condition for the simultaneous positivity of both risk premia, and hence for simultaneous cross-investment, is that $\left[1 + r - r^* + \tfrac{1}{2}\sigma^2\right]\left[1 + r^* - r + \tfrac{1}{2}\sigma^2\right] > 1$.

while similar rearrangement for the foreign country gives:

$$\begin{aligned}\tfrac{1}{dt}E(dk^*) &= \frac{F^*(k^*)}{k^*}\left[k^* - a^*\right] + \left[1/q + \tfrac{1}{dt}E(d[1/q])\right]\frac{F(k)}{k}a^* - z^* \\ &= F^*(k^*) + P^*(k^*,k,r^*,r,1/q)a^* - z^*\end{aligned}$$

Consequently, the resource accumulation constraints in the two countries may be written as:

$$dk = \left\{F(k) + P(k,k^*,r,r^*,q)a - z\right\}dt + \sigma q\left[F^*(k^*)/k^*\right]ad\xi \ , \tag{17.3}$$

$$dk^* = \left\{F^*(k^*) + P^*(k^*,k,r^*,r,1/q)a^* - z^*\right\}dt - \sigma[1/q]\left[F(k)/k\right]a^*d\xi \, . \tag{17.3*}$$

It is evident that the problems of the two countries are essentially symmetric and attention may be directed to determining the optimal solution for the home country. Some preliminary results and rearrangement of the optimization problem using duality theory are first presented in some detail for the home country and then summarized for the foreign country.

17.3.2 A Dual View of the Optimization Problem Using Concepts of Consumer, Firm and Economy-Wide Profit

Define the instantaneous consumer "profit" maximization problem for the home country representative agent:

$$\Phi(\lambda) = \max_z \left\langle U(z) - \lambda z\right\rangle . \tag{17.4}$$

The parameter λ, used in (17.4) to evaluate the cost of expenditure in utility terms in defining instantaneous consumer "profit", will play an important role as the costate variable in the intertemporal optimization problem.[8] The instantaneous optimization problem (17.4) has first order condition:

$$U_z(z) = \lambda . \tag{17.5}$$

Importantly, optimal expenditure as a function of λ can be recovered from the optimized consumer profit function by Hotelling's (envelope) theorem:

---

[8] Consequent to this definition, in an exception to the previously developed notational convention which generally assigns Roman letters to variables and Greek letters to parameters, the Greek letter λ, which may be thought of as a parameter in the atemporal problem (17.4), nevertheless denotes a variable in the context of the intertemporal problem to be developed below as a variant of (17.1).

$$z = -\Phi_\lambda(\lambda). \tag{17.6}$$

We note for later use that λ-conditioned ("Frischian") utility, defined by $U^F(\lambda) = U(U_z^{-1}(\lambda))$, may be constructed from the profit function on combination of (17.6) with a rearrangement of (17.4) as:

$$U^F(\lambda) = \Phi(\lambda) - \lambda\Phi_\lambda(\lambda). \tag{17.7}$$

Next, define the (optimized, current value, deterministic component of the) Hamiltonian:

$$H(k,\lambda) = \Phi(\lambda) + \lambda F(k). \tag{17.8}$$

In the context of instantaneous profit maximization, the interest rate, r, may be employed to represent the implicit rental rate of capital. In a competitive environment, defined as one in which there are no unexploited instantaneous arbitrage opportunities, an appropriate measure of instantaneous economy-wide economic profit is:

$$\Pi(\lambda,r) = \max_{z,k}\left\langle [U(z) - \lambda z] + \lambda[F(k) - rk] \right\rangle \tag{17.9}$$

This follows from a definition of economy-wide economic profit obtained by combining consumer and producer profit, where producer profit is defined firstly as $\max_k \langle F(k) - rk \rangle$, with the instantaneous opportunity cost of capital in production being valued at the available instantaneous interest rate, and then converted to utility terms by use of the scale factor λ.

However, using (17.4) and 17.(8), it will be convenient to work with the intermediate functions Φ and H in what follows. Thus, (17.9) is rewritten as:

$$\Pi(\lambda,r) = \max_k \left\langle H(k,\lambda) - r\lambda k \right\rangle. \tag{17.10}$$

The concept of instantaneous economy wide profit, the maximization of which ensures the exhaustion of instantaneous arbitrage opportunities, has many useful implications. We highlight six of them.

Firstly, the first order condition for (17.10) implies $H_k(k,\lambda) = r\lambda$, which in view of (17.8) further implies, if *k* is optimal for (17.10), that:

$$r = F_k(k) \tag{17.11}$$

Given (17.3), application of Ito's Lemma to (17.11) implies:

$$dr = \begin{Bmatrix} F_{kk}(k)\left[F(k) + P(k,k^*,r,r^*,q)a - z\right] \\ +\tfrac{1}{2}F_{kkk}(k)\left[\sigma q[F^*(k^*)/k^*]a\right]^2 \end{Bmatrix} dt \qquad (17.12)$$
$$+F_{kk}(k)\sigma q[F^*(k^*)/k^*]ad\xi$$

If the assumption of absence of unexploited arbitrage opportunities is to be maintained in the context of problem (17.1), (17.1c) must be compatible with (17.12). Replacing z and a in (17.12) with their optimal feedback forms $\tilde{Z}(k,k^*,r,r^*,q)$ and $\tilde{A}(k,k^*,r,r^*,q)$, then for compatibility of (17.12) with (17.1c) it is necessary that the following specifications apply in (17.1c):

$$R(k,k^*,r,r^*,q) = F_{kk}(k)\begin{bmatrix} F(k) + P(k,k^*,r,r^*,q)\tilde{A}(k,k^*,r,r^*,q) \\ -\tilde{Z}(k,k^*,r,r^*,q) \end{bmatrix} \qquad (17.13)$$
$$+\tfrac{1}{2}F_{kkk}(k)\left[\sigma q[F^*(k^*)/k^*]\tilde{A}(k,k^*,r,r^*,q)\right]^2$$

$$S(k,k^*,r,r^*,q) = F_{kk}(k)\sigma q[F^*(k^*)/k^*]\tilde{A}(k,k^*,r,r^*,q) \qquad (17.14)$$

$$dw\ (= dw^*) = d\xi\,. \qquad (17.15)$$

Specifications (17.13)-(17.15) ensure model consistency of the interest rate processes with the absence of arbitrage. An implication, in the model form examined here, is that the one source of uncertainty flows through from exchange rate risk, via capital stock risk, to interest rate risk.

Second, let $\hat{k}$ denote the value of the capital stock which is compatible with r under the assumption of absence of unexploited instantaneous arbitrage opportunities. Then (17.10) applies and, in view of the assumption that $F_{kk} < 0$, $\hat{k}$ may be written in terms of r on inversion of (17.11) as:

$$\hat{k} = F_k^{-1}(r)\,. \qquad (17.16)$$

Thus, in cases where it is appropriate to make use of the assumption of the absence of unexploited instantaneous arbitrage opportunities, (17.16) allows the optimal capital stock to be explicitly eliminated in terms of the interest rate *r*.

Third, it is clear by inspection of (17.10) that we can define a "quasi-Hamiltonian":

$$\hat{H}(k,\lambda,r) = \Pi(\lambda,r) + r\lambda k \qquad (17.17)$$

which is, very conveniently, linear affine in *k*, and which has the property that, in the assumed absence of unexploited instantaneous arbitrage opportunities, it reduces to the "true" Hamiltonian. That is, given that $\hat{k}$ solves (17.10), it follows from (17.10), (17.16) and (17.17) that:

$$\hat{H}(F_k^{-1}(r), \lambda, r) \equiv H(F_k^{-1}(r), \lambda) . \tag{17.18}$$

Fourth, under the assumption of the absence of unexploited instantaneous arbitrage opportunities, the average rate of return on holdings of productive assets can be defined in terms of r as:

$$M(r) = \frac{F(F_k^{-1}(r))}{F_k^{-1}(r)} . \tag{17.19}$$

In particular, in the case of optimization within the competitive representative agent framework, this allows the risk premium to be written without direct reference to the capital stock. Thus:

$$\hat{P}(r, r^*, q) = q\left[1 + r - r^* + \tfrac{1}{2}\sigma^2\right] M^*(r^*) - M(r) . \tag{17.20}$$

Similar concepts may be defined for the foreign country. Of particular relevance at this point are:

$$r^* = F^*_{k*}(k^*) , \tag{17.11*}$$

$$\begin{aligned} dr^* = & \left\{ \begin{matrix} F^*_{k*k*}(k^*) \left[\begin{matrix} F^*(k^*) + P^*(k^*, k, r^*, r, 1/q) \tilde{A}^*(k^*, k, r^*, r, 1/q) \\ -\tilde{Z}^*(k^*, k, r^*, r, 1/q) \end{matrix} \right] \\ + \tfrac{1}{2} F^*_{k*k*k*}(k^*) \left[\sigma[1/q][F(k)/k] \tilde{A}^*(k^*, k, r^*, r, 1/q) \right]^2 \end{matrix} \right\} dt \\ & -F^*_{k*k*}(k^*) \sigma[1/q][F(k)/k] \tilde{A}^*(k^*, k, r^*, r, 1/q) d\xi, \end{aligned} \tag{17.12*}$$

$$\hat{k}^* = F^{*\,-1}_{k*}(r^*) \qquad , \tag{17.16*}$$

$$M^*(r^*) = \frac{F^*(F^{*\,-1}_{k*}(r^*))}{F^{*\,-1}_{k*}(r^*)} , \tag{17.19*}$$

$$\hat{P}^*(r^*, r, 1/q) = [1/q]\left[1 + r^* - r + \tfrac{1}{2}\sigma^2\right] M(r) - M^*(r^*) . \tag{17.20*}$$

Fifth, in the competitive agent optimization framework in which the average return on domestic capital and the foreign investment risk premium are assumed to be exogenous to the decision maker, the above results allow specification of the transition equations for the home and foreign capital stocks in forms which are, conveniently, linear in k and k^* respectively as:

$$dk = \left\{M(r)k + \hat{P}(r,r^*,q)a - z\right\}dt + \sigma q M^*(r^*)a\,d\xi\,, \tag{17.21}$$

$$dk^* = \left\{M^*(r^*)k^* + \hat{P}^*(r^*,r,1/q)a^* - z^*\right\}dt - \sigma[1/q]M(r)a^*\,d\xi\,. \tag{17.21*}$$

Sixth, for the particular optimization to be pursued below, corresponding to the competitive optimum in which the representative agents undertake optimal decision making without recognising a link between their actions and interest rates, we may further simplify the transition equations for the domestic and foreign interest rates in two useful ways. The first simplification involves employing (17.16) and (17.16*) to eliminate explicit dependence of these transition equations on the capital stocks. To impose (17.16) and (17.16*) on (17.12) and (17.12*) with a minimum of notational clutter, it is also convenient to define some simplifying expressions:

$$F_0(r) \equiv F(F_k^{-1}(r))\,, \qquad F_2(r) \equiv F_{kk}(F_k^{-1}(r))\,, \qquad F_3(r) \equiv F_{kkk}(F_k^{-1}(r))$$

$$F_0^*(r^*) \equiv F^*(F_{k^*}^{*-1}(r^*))\,, F_2^*(r^*) \equiv F_{k^*k^*}^*(F_{k^*}^{*-1}(r^*))\,, \qquad F_3^*(r^*) \equiv F_{k^*k^*k^*}^*(F_{k^*}^{*-1}(r^*))$$

.

The second simplification recognizes that under the assumption of the absence of unexploited arbitrage opportunities, the feedback solutions may be written with capital stock dependence entirely eliminated in terms of interest rates. Again using (17.16) and (17.16*), this implies that the feedback solutions may be expressed as:

$$\hat{Z}(r,r^*,q) = \tilde{Z}(F_k^{-1}(r), F_{k^*}^{*-1}(r^*), r, r^*, q) \tag{17.22a}$$

$$\hat{A}(r,r^*,q) = \tilde{A}(F_k^{-1}(r), F_{k^*}^{*-1}(r^*), r, r^*, q) \tag{17.22b}$$

$$\hat{Z}^*(r^*,r,1/q) = \tilde{Z}^*(F_{k^*}^{*-1}(r^*), F_k^{-1}(r), r^*, r, 1/q) \tag{17.22a*}$$

$$\hat{A}^*(r^*,r,1/q) = \tilde{A}^*(F_{k^*}^{*-1}(r^*), F_k^{-1}(r), r^*, r, 1/q)\,. \tag{17.22b*}$$

These forms of the feedback solutions may be employed in any expressions in which the optimal values of the decision variables are required but where the competitiveness assumption ensures that the values within those expressions are taken as given by the decision makers.

With these notational conventions the interest rate transition equations in the competitive case become:

$$dr = \left\{ \begin{array}{l} F_2(r)\left[F_0(r) + \hat{P}(r,r^*,q)\hat{A}(r,r^*,q) - \hat{Z}(r,r^*,q)\right] \\ +\frac{1}{2}F_3(r)\left[\sigma q M^*(r^*)\hat{A}(r,r^*,q)\right]^2 \end{array} \right\} dt \tag{17.23}$$
$$+F_2(r)\sigma q M^*(r^*)\hat{A}(r,r^*,q)d\xi,$$

$$dr^* = \left\{ \begin{array}{l} F_2^*(r^*)\left[F_0^*(r^*) + \hat{P}^*(r^*,r,1/q)\hat{A}^*(r^*,r,1/q) - \hat{Z}^*(r^*,r,1/q)\right] \\ +\frac{1}{2}F_3^*(r^*)\left[\sigma[1/q]M(r)\hat{A}^*(r^*,r,1/q)\right]^2 \end{array} \right\} dt$$
$$-F_2^*(r^*)\sigma[1/q]M(r)\hat{A}^*(r^*,r,1/q)d\xi \tag{17.23*}$$

17.3.3 A Reformulation of the Intertemporal Optimization Problem

Problem (17.1) is now reformulated in order to exploit these results. This is done by utilizing the transition equations for k, k^*, r and r^* in the forms given by (17.21), (17.21*), (17.23) and (17.23*) respectively, without at this point imposing any relationship between the initial conditions for the capital stock and the interest rate. This somewhat artificial problem contains the competitive variant of problem (17.1) as a special case, namely when the initial conditions for the capital stocks and the interest rates are linked by explicit imposition of constraints (17.11) and (17.11*).

The optimization problems for the home and foreign countries are now able to be expressed in decoupled form. For this purpose, the home country's domestic expenditure feedback rule is now defined as a function $\bar{Z}:\mathbb{R}_+^4 \to \mathbb{R}_+$, denoted $z = \bar{Z}(k,r,r,^* q)$, chosen from a set of admissible functions $\bar{\mathbb{Z}}\{k,r,r,^* q\}$. An admissible portfolio allocation feedback rule for the representative domestic agent is a function $\bar{A}:\mathbb{R}_+^4 \to \mathbb{R}$, denoted $a = \bar{A}(k,r,r,^* q)$, chosen from a set of admissible functions $\bar{\mathfrak{A}}\{k,r,r,^* q\}$. Similarly, an admissible expenditure feedback rule for the representative foreign agent is a function $\bar{Z}^*:\mathbb{R}_+^4 \to \mathbb{R}_+$, denoted by $z^* = \bar{Z}^*(k^*,r^*,r,1/q)$, and which is chosen from a set of admissible functions $\bar{\mathbb{Z}}^*\{k^*,r^*,r,1/q\}$. An admissible portfolio allocation feedback rule for the representative foreign agent is a function $\bar{A}^*:\mathbb{R}_+^4 \to \mathbb{R}$, denoted by $a^* = \bar{A}(k^*,r^*,r,1/q)$, chosen from a set of admissible functions $\bar{\mathfrak{A}}^*\{k^*,r^*,r,1/q\}$.

For the home country, the problem is:[9]

---

[9] Explicit time dependence of relevant variables is reintroduced temporarily in the specification of the home and foreign countries' dynamic optimization problems.

$$\bar{J}(k,r,r^*,q) = \max_{\{z\in\bar{\mathbb{Z}},a\in\bar{\mathfrak{A}}\}} E_0\int_0^\infty e^{-\delta t}U(z(t))dt \tag{17.24a}$$

subject to:

$$\begin{aligned} dk(t) &= \left\{M(r(t))k(t)+\hat{P}(r(t),r^*(t),q(t))a(t)-z(t)\right\}dt \\ &\quad +\sigma q(t)M^*(r^*(t))a(t)d\xi(t), \end{aligned} \tag{17.24b}$$

$$\begin{aligned} dr(t) &= \left\{\begin{array}{l} F_2(r(t))\left[\begin{array}{l} F_0(r(t))+\hat{P}(r(t),r^*(t),q(t))\hat{A}(r(t),r^*(t),q(t)) \\ -\hat{Z}(r(t),r^*(t),q(t)) \end{array}\right] \\ +\frac{1}{2}F_3(r(t))\left[\sigma q(t)M^*(r^*(t))\hat{A}(r(t),r^*(t),q(t))\right]^2 \end{array}\right\}dt \\ &\quad +F_2(r(t))\sigma q(t)M^*(r^*(t))\hat{A}(r(t),r^*(t),q(t))d\xi(t), \end{aligned} \tag{17.24c}$$

$$\begin{aligned} dr^*(t) &= \left\{\begin{array}{l} F_2^*(r^*(t))\left[\begin{array}{l} F_0^*(r^*(t)) \\ +\hat{P}^*(r^*(t),r(t),1/q(t))\hat{A}^*(r^*(t),r(t),1/q(t)) \\ -\hat{Z}^*(r^*(t),r(t),1/q(t)) \end{array}\right] \\ +\frac{1}{2}F_3^*(r^*(t))\left[\sigma[1/q(t)]M(r(t))\hat{A}^*(r^*(t),r(t),1/q(t))\right]^2 \end{array}\right\}dt \\ &\quad -F_2^*(r^*(t))\sigma[1/q(t)]M(r(t))\hat{A}^*(r^*(t),r(t),1/q(t))d\xi(t), \end{aligned} \tag{17.24c*}$$

$$dq(t) = \left[r(t)-r^*(t)+\tfrac{1}{2}\sigma^2\right]q(t)dt+\sigma q(t)d\xi(t), \text{ and} \tag{17.24d}$$

$$k(0)=k\ ,\ r(0)=r\ ,\ r^*(0)=r^*\ ,\ q(0)=q\ . \tag{17.24e}$$

Let $\hat{z}=\hat{\bar{Z}}(k,r,r,^* q)$ and $\hat{a}=\hat{\bar{A}}(k,r,r,^* q)$ denote the optimal feedback solutions for the home country problem (17.24).

For the foreign country, the problem is:

$$\bar{J}^*(k^*,r^*,r,1/q) = \max_{\{z^*\in\bar{\mathbb{Z}}^*,a^*\in\bar{\mathfrak{A}}^*\}} E_0\int_0^\infty e^{-\delta^* t}U^*(z(t))dt, \tag{17.25a}$$

subject to:

$$\begin{aligned} dk^*(t) &= \left\{M^*(r^*(t))k^*(t)+\hat{P}^*(r^*(t),r(t),1/q(t))a^*(t)-z^*(t)\right\}dt \\ &\quad -\sigma[1/q(t)]M(r(t))a^*(t)d\xi(t), \end{aligned} \tag{17.25b}$$

$$dr(t) = \cdots \text{the same specification as (17.24c)} \cdots \tag{17.25c}$$

$$dr^*(t) = \cdots \text{the same specification as (17.24c*)} \cdots \tag{17.25c*}$$

$$d[1/q(t)] = \left[r^*(t) - r(t) + \tfrac{1}{2}\sigma^2\right][1/q(t)]dt - \sigma[1/q(t)]d\xi(t), \text{ and} \tag{17.25d}$$

$$k^*(0) = k^*,\ r^*(0) = r^*,\ r(0) = r,\ q(0) = q\ . \tag{17.25e}$$

Let $\hat{z}^* = \hat{\bar{Z}}^*(k^*, r^*, r, 1/q)$, $\hat{a}^* = \hat{A}^*(k^*, r^*, r, 1/q)$ denote the optimal feedback solutions for the foreign country problem (24).

In the derivations above, equations (17.22) exhibited a relationship between the solutions to problem (17.1) under the additional assumption of competitive agents and feedback solutions required to describe the dependence of the interest rate transition relationships on variables which were taken as exogenous to the decision maker in the competitive environment. These solutions need to be linked logically and compatibly to the solutions of problems (17.24) and (17.25). The relevant linkage definitions are:

$$\hat{Z}(r, r^*, q) \equiv \hat{\bar{Z}}(F_k^{-1}(r), r, r^*, q), \tag{17.26a}$$

$$\hat{A}(r, r^*, q) \equiv \hat{\bar{A}}(F_k^{-1}(r), r, r^*, q), \tag{17.26b}$$

$$\hat{Z}^*(r^*, r, q) \equiv \hat{\bar{Z}}^*(F_{k*}^{*-1}(r^*), r^*, r, q), \text{ and} \tag{1726a*}$$

$$\hat{A}^*(r^*, r, q) \equiv \hat{\bar{A}}^*(F_{k*}^{*-1}(r^*), r^*, r, q)\ . \tag{17.26b*}$$

With the above model characteristics set out, it is now necessary to demonstrate that problems structured as in (17.24) and (17.25) are capable of solution in an analytical form suitable for econometric estimation of the feedback solutions, for very general utility and technology specifications. For this purpose, attention is concentrated upon the decoupled home country problem (17.24). The solution for (17.25) can be derived in an analogous manner.

17.4 Solution of the Intertemporal Problem

17.4.1 The Hamilton-Jacobi Bellman Equation

This section develops the solution to problem (17.24) and, by analogy, the solution to the mirror image problem (17.25), in a form capable of adaptation to econometric estimation of the model equations for quite general preference and technology specifications. The Hamilton-Jacobi-Bellman (HJB) equation for the home country representative competitive agent problem (17.24), constructed by Bellman's Principle of Optimality, involves the representative agent making a trade-off between the acquisition of utility from current expenditure and the acquisition of potential for additional utility in future as an outcome of capital accumulation. The HJB equation for the competitive optimum for the home country is:

$$\delta\,\bar{J}(k,r,r^*,q) = \max_{z,a}$$

$$\left\langle U(z) + \frac{1}{dt}\left\{\begin{array}{l} \bar{J}_k E(dk) + \bar{J}_r E(dr) + \bar{J}_{r*} E(dr^*) + \bar{J}_q E(dq) \\ +\bar{J}_{kr} E(dk)(dr) + \bar{J}_{kr*} E(dk)(dr^*) + \bar{J}_{kq} E(dk)(dq) \\ +\bar{J}_{rr*} E(dr)(dr^*) + \bar{J}_{rq} E(dr)(dq) + \bar{J}_{r*q} E(dr^*)(dq) \\ +\tfrac{1}{2}\left[\bar{J}_{kk} E(dk)^2 + \bar{J}_{rr} E(dr)^2 + \bar{J}_{r*r*} E(dr^*)^2 + \bar{J}_{qq} E(dq)^2\right] \end{array}\right\}\right\rangle \tag{17.27a}$$

where dependence of the partial derivatives of $\bar{J}$ on (k,r,r^*,q) has been suppressed for notational convenience, and where:[10]

$$\frac{1}{dt}E(dk) = M(r)k + \hat{P}(r,r^*,q)a - z, \tag{17.27b}$$

$$\begin{aligned} \frac{1}{dt}E(dr) &= F_2(r)\left[F_0(r) + \hat{P}(r,r^*,q)\hat{A}(r,r^*,q) - \hat{Z}(r,r^*,q)\right] \\ &\quad + \tfrac{1}{2}F_3(r)\sigma^2 q^2 M^*(r^*)^2 \hat{A}(r,r^*,q)^2 \\ &\equiv E^r(r,r^*,q), \end{aligned} \tag{17.27c}$$

---

[10] The standard implications of the continuous time deterministic process dt and the continuous time stochastic (Wiener) process dv have been applied in simplifying the following presentation. In particular, the second order expectations terms (17.27f)-(17.27o) exploit the results: $E(dt)^2 = 0$, $E(d\xi)^2 = dt$ and $E(d\xi)(dt) = 0$.

$$\begin{aligned}\frac{1}{dt}E(dr^*) &= F_2^*(r^*)\left[F_0^*(r^*) + P^*(r^*,r,1/q)\hat{A}^*(r^*,r,1/q) - \hat{Z}^*(r^*,r,1/q)\right] \\ &\quad + \tfrac{1}{2}F_3^*(r^*)\sigma^2[1/q]^2 M(r)^2 \hat{A}^*(r^*,r,1/q)^2 \\ &\equiv E^{r^*}(r,r^*,q),\end{aligned} \tag{17.27d}$$

$$\frac{1}{dt}E(dq) = \left[r - r^* + \tfrac{1}{2}\sigma^2\right]q, \tag{17.27e}$$

$$\frac{1}{dt}E(dk)(dr) = F_2(r)\sigma^2 q^2 M^*(r^*)^2 \hat{A}(r,r^*,q)a, \tag{17.27f}$$

$$\frac{1}{dt}E(dk)(dr^*) = -F_2^*(r^*)\sigma^2 M(r)M^*(r^*)\hat{A}^*(r^*,r,1/q)a, \tag{17.27g}$$

$$\frac{1}{dt}E(dk)(dq) = \sigma^2 q^2 M^*(r^*)a, \tag{17.27h}$$

$$\begin{aligned}\frac{1}{dt}E(dr)(dr^*) &= -F_2(r)F_2^*(r^*)\sigma^2 M(r)M^*(r^*)\hat{A}(r,r^*,q)\hat{A}^*(r^*,r,1/q) \\ &\equiv E^{rr^*}(r,r^*,q),\end{aligned} \tag{17.27i}$$

$$\frac{1}{dt}E(dr)(dq) = F_2(r)\sigma^2 q^2 M^*(r^*)\hat{A}(r,r^*,q) \equiv E^{rq}(r,r^*,q), \tag{17.27j}$$

$$\frac{1}{dt}E(dr^*)(dq) = -F_2^*(r^*)\sigma^2 M(r)\hat{A}^*(r^*,r,1/q) \equiv E^{r^*q}(r,r^*,q), \tag{17.27k}$$

$$\frac{1}{dt}E(dk)^2 = \sigma^2 q^2 M^*(r^*)^2 a^2, \tag{17.27l}$$

$$\frac{1}{dt}E(dr)^2 = F_2(r)^2\sigma^2 q^2 M^*(r^*)^2 \hat{A}(r,r^*,q)^2 \equiv E^{rr}(r,r^*,q), \tag{17.27m}$$

$$\frac{1}{dt}E(dr^*)^2 = F_2^*(r^*)^2\sigma^2[1/q]^2 M(r)^2 \hat{A}^*(r^*,r,1/q)^2 \equiv E^{r^*r^*}(r,r^*,q), \tag{17.27n}$$

$$\frac{1}{dt}E(dq)^2 = \sigma^2 q^2. \tag{17.27o}$$

To solve the first order conditions for a competitive optimum in problem (27) it is important to note that the effect of the assumption of competitive decision making is that the agent acts as if the decision making affects expectations terms involving dk—that is, the five terms (17.27b), (17.27f), (17.27g), (17.27h), (17.27l)—but not expectations terms exclusively involving dr, dr^*, or dq—that is, the nine terms (17.27c)-(17.27e), (17.27i)-(17.27k), and (17.27m)-(17.27o). These terms are treated as exogenous from the point of view of the optimal control decisions of the representative competitive agent.[11] Under these conditions, problem (17.27) can be written in a simplified form as:

$$\delta\bar{J}(k,r,r^*,q) = \max_{z,a}$$
$$\left\langle U(z) + \frac{1}{dt}\left\{ \begin{array}{l} \bar{J}_k(k,r,r^*,q)\left[M(r)k + \hat{P}(r,r^*,q)a - z\right] \\ +\bar{J}_{kr}(k,r,r^*,q)F_2(r)\sigma^2 q^2 M^*(r^*)^2 \hat{A}(r,r^*,q)a \\ -\bar{J}_{kr*}(k,r,r^*,q)F_2^*(r^*)\sigma^2 M(r)M^*(r^*)\hat{A}^*(r^*,r,1/q)a \\ +\bar{J}_{kq}(k,r,r^*,q)\sigma^2 q^2 M^*(r^*)a \\ +\tfrac{1}{2}\bar{J}_{kk}(k,r,r^*,q)\sigma^2 q^2 M^*(r^*)^2 a^2 \end{array} \right\} \right\rangle \quad (17.28)$$
$$+\Omega(k,r,r^*,q)$$

where the components of the HJB equation which are exogenous to the decision making competitive agent have been extracted from the optimization and are summarized as:

$$\begin{aligned} \Omega(k,r,r^*,q) = {} & \\ & \bar{J}_r(k,r,r^*,q)E^r(r,r^*,q) + \bar{J}_{r*}(k,r,r^*,q)E^{r*}(r^*,r,1/q) \\ & +\bar{J}_q(k,r,r^*,q)\left[r - r^* + \tfrac{1}{2}\sigma^2\right]q + \bar{J}_{rr*}(k,r,r^*,q)E^{rr*}(r,r^*,q) \\ & +\bar{J}_{rq}(k,r,r^*,q)E^{rq}(r,r^*,q) + \bar{J}_{r*q}(k,r,r^*,q)E^{r*q}(r,r^*,q) \\ & +\tfrac{1}{2}\bar{J}_{rr}(k,r,r^*,q)E^{rr}(r,r^*,q)^2 + \tfrac{1}{2}\bar{J}_{r*r*}(k,r,r^*,q)E^{r*r*}(r,r^*,q)^2 \\ & +\tfrac{1}{2}\bar{J}_{qq}(k,r,r^*,q)\sigma^2 q^2 \end{aligned} \quad (17.29)$$

Given the structure implied by (17.28) and (17.29), the first order conditions for (17.27) reduce simply to:

---

[11] Additionally, to simplify expressions which follow, first, cross and higher order terms in expectations involving exogenous (from the point of view of the optimizer) interest rates are represented by the $E^i(r,r^*,q)$ notation introduced along with (17.27c), (17.27d), (17.27i), (17.27j), (17.27k), (17.27m) and (17.27n).

$$U_z(z) = \bar{J}_k(k,r,r^*,q), \tag{17.30}$$

and

$$a = \frac{-B(k,r,r^*,q)\bar{J}_k(k,r,r^*,q)}{\sigma^2 q^2 M^*(r^*)^2 \bar{J}_{kk}(k,r,r^*,q)}, \tag{17.31}$$

where:

$$\begin{aligned} B(k,r,r^*,q) = {} & q\left[1+r-r^*+\tfrac{1}{2}\sigma^2\right]M^*(r)-M(r) \\ & +F_2(r)\sigma^2 q^2 M^*(r^*)^2 \hat{A}(r,r^*,q)\frac{\bar{J}_{kr}(k,r,r^*,q)}{\bar{J}_k(k,r,r^*,q)} \\ & -F_2^*(r^*)\sigma^2 M(r)M^*(r^*)\hat{A}^*(r^*,r,1/q)\frac{\bar{J}_{kr^*}(k,r,r^*,q)}{\bar{J}_k(k,r,r^*,q)} \\ & +\sigma^2 q^2 M^*(r^*)\frac{\bar{J}_{kq}(k,r,r^*,q)}{\bar{J}_k(k,r,r^*,q)}. \end{aligned} \tag{17.32}$$

Using (17.4), (17.5) and (17.8) together with (17.29)-(17.32) allows the HJB equation, implied by (17.28) and optimized from the viewpoint of the representative competitive agent, to be written as:

$$\begin{aligned} \delta\bar{J}(k,r,r^*,q) = {} & H(k,\bar{J}_k(k,r,r^*,q)) - \tfrac{1}{2}\left[\frac{B(k,r,r^*,q)}{\sigma q M^*(r^*)}\right]^2 \frac{\bar{J}_k(k,r,r^*,q)^2}{\bar{J}_{kk}(k,r,r^*,q)} \\ & +\Omega(k,r,r^*,q). \end{aligned} \tag{17.33}$$

17.4.2 A Reformulation of the HJB Equation

The HJB equation in the form (17.33) is a highly nonlinear partial differential equation, in view of (17.32). The objective of this section is to introduce a change of variables, from the state variable k to the costate variable λ, which enables the HJB equation to be rewritten in a form in which it can be solved to ultimately provide a source of endogenization of the costate variable. In order to complete this task, use is made of the optimal value function, treating it as "output" and defining its dual, optimal intertemporal profit.

Dual to the intertemporal utility maximizing problem (28) is the following "intertemporal profit" maximizing problem:

$$\Psi(\lambda,r,r^*,q) = \max_k \left\langle \bar{J}(k,r,r^*,q) - \lambda k \right\rangle, \tag{17.34}$$

which implies, consistently with (30) and (5), the first order condition:

$$\bar{J}_k(k,r,r^*,q)=\lambda\ . \tag{17.35}$$

Envelope results applied to (17.34) give:

$$\Psi_\lambda(\lambda,r,r^*,q)=-k, \tag{17.36a}$$

$$\Psi_r(\lambda,r,r^*,q)=\bar{J}_r(k,r,r^*,q), \tag{17.36b}$$

$$\Psi_{r*}(\lambda,r,r^*,q)=\bar{J}_{r*}(k,r,r^*,q), \tag{17.36c}$$

$$\Psi_q(\lambda,r,r^*,q)=\bar{J}_q(k,r,r^*,q). \tag{17.36d}$$

Result (17.36a), which may be interpreted as an intertemporal variant of Hotelling's Theorem, is important to highlight as providing the link between the costate and state variables:

$$k=-\Psi_\lambda(\lambda,r,r^*,q). \tag{17.37}$$

Together, (17.35) and (17.37) yield an identity:

$$\bar{J}_k(-\Psi_\lambda(\lambda,r,r^*,q),r,r^*,q)\equiv\lambda\ , \tag{17.38}$$

differentiation of which further implies:

$$\bar{J}_{kk}=-1/\Psi_{\lambda\lambda}, \tag{17.39a}$$

$$\bar{J}_{kr}=-\Psi_{\lambda r}/\Psi_{\lambda\lambda}, \tag{17.39b}$$

$$\bar{J}_{kr*}=-\Psi_{\lambda r*}/\Psi_{\lambda\lambda},\text{ and} \tag{17.39c}$$

$$\bar{J}_{kq}=-\Psi_{\lambda q}/\Psi_{\lambda\lambda}. \tag{17.39d}$$

Using (17.35) and (17.39), (17.32) may be rewritten as:

$$\begin{aligned}\tilde{B}(\lambda,r,r^*,q) = {} & q\left[1+r-r^*+\tfrac{1}{2}\sigma^2\right]M^*(r)-M(r)\\ & -F_2(r)\sigma^2 q^2 M^*(r^*)^2\hat{A}(r,r^*,q)\frac{\Psi_{\lambda r}(\lambda,r,r^*,q)}{\lambda\Psi_{\lambda\lambda}(\lambda,r,r^*,q)}\\ & +F_2^*(r^*)\sigma^2 M(r)M^*(r^*)\hat{A}^*(r^*,r,1/q)\frac{\Psi_{\lambda r*}(\lambda,r,r^*,q)}{\lambda\Psi_{\lambda\lambda}(\lambda,r,r^*,q)}\\ & -\sigma^2 q^2 M^*(r^*)\frac{\Psi_{\lambda q}(\lambda,r,r^*,q)}{\lambda\Psi_{\lambda\lambda}(\lambda,r,r^*,q)}.\end{aligned} \tag{17.40}$$

The solution technique to be employed extends the deterministic approach of Cooper (2001) to solve the HJB equation exactly for the case where the intertemporal profit function has the property of exogenous ratios of certain key elasticities (ie these are independent of λ). Specifically, we seek solutions in which Ψ has the structure:

$$\begin{aligned}&\left(\partial\ln\Psi_\lambda/\partial\ln r\right)/\left(\partial\ln\Psi_\lambda/\partial\ln\lambda\right)=\chi_r(r,r^*,q)\\ &\left(\partial\ln\Psi_\lambda/\partial\ln r^*\right)/\left(\partial\ln\Psi_\lambda/\partial\ln\lambda\right)=\chi_{r*}(r,r^*,q)\\ &\left(\partial\ln\Psi_\lambda/\partial\ln q\right)/\left(\partial\ln\Psi_\lambda/\partial\ln\lambda\right)=\chi_q(r,r^*,q)\end{aligned} \tag{17.41}$$

If these conditions hold then (17.40) is not a function of λ. In this case we may replace $\tilde{B}(\lambda,r,r^*,q)$ by the notation $\gamma\equiv\Gamma(r,r^*,q)$ where:

$$\Gamma(r,r^*,q)=q\left[1+\chi(r,r^*,q)\right]M^*(r)-M(r) \tag{17.42a}$$

$$\chi(r,r^*,q)=r-r^*+\tfrac{1}{2}\sigma^2+\chi_0(r,r^*,q) \tag{17.42b}$$

$$\begin{aligned}\chi_0(r,r^*,q) = {} & (1/r^*)F_2^*(r^*)\sigma^2 M(r)\hat{A}^*(r^*,r,1/q)\chi_{r*}(r,r^*,q)\\ & -(1/r)F_2(r)\sigma^2 q M^*(r^*)\hat{A}(r,r^*,q)\chi_r(r,r^*,q)\\ & -\sigma^2 q\chi_q(r,r^*,q).\end{aligned} \tag{17.42c}$$

Additionally, using (17.37), (17.35) may be rearranged to give a costate-variable-conditioned form of the optimal value function:

$$\tilde{J}(\lambda,r,r^*,q)=\Psi(\lambda,r,r^*,q)-\lambda\Psi_\lambda(\lambda,r,r^*,q). \tag{17.43}$$

Then, using (17.43) along with (17.17), (17.35), (17.37), (17.39a) and (17.42), the HJB equation (17.33) may be rewritten in terms of the intertemporal profit function Ψ as:

$$\begin{aligned}\delta\left[\Psi(\lambda,r,r^*,q)-\lambda\Psi_\lambda(\lambda,r,r^*,q)\right]&=\Pi(\lambda,r)-r\lambda\Psi_\lambda(\lambda,r,r^*,q)\\&+\left(\frac{1}{2\sigma_0{}^2}\right)\gamma^2\lambda^2\Psi_{\lambda\lambda}(\lambda,r,r^*,q)\\&+\Omega(r,r^*,q),\end{aligned}\tag{17.44}$$

where $\sigma_0 \equiv \sigma q M^*(r)$. Given the assumed structure (17.41), equation (17.44) is a second order linear differential equation in λ which may be explicitly solved for $\Psi(\lambda,r,r^*,q)$.

Assuming $0<\gamma<\infty$, the solution of (17.44) can be verified by direct examination to be:

$$\begin{aligned}\Psi(\lambda,r,r^*,q)&=\frac{\lambda^{\beta_1}\int_0^\lambda \zeta^{-\beta_1-1}\Pi(\zeta,r)d\zeta+\lambda^{\beta_2}\int_\lambda^\infty \zeta^{-\beta_2-1}\Pi(\zeta,r)d\zeta}{\frac{1}{2}(\beta_2-\beta_1)(\gamma/\sigma_0)^2}\\&+\frac{\Omega(r,r^*,q)}{\delta},\end{aligned}\tag{17.45}$$

where:

$$\beta_1=\frac{r-\delta+\frac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}-\sqrt{\left[\frac{r-\delta+\frac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}\right]^2+\frac{2\delta}{(\gamma/\sigma_0)^2}},$$

$$\beta_2=\frac{r-\delta+\frac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}+\sqrt{\left[\frac{r-\delta+\frac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}\right]^2+\frac{2\delta}{(\gamma/\sigma_0)^2}}$$

and $\bar{\Omega}(r,r8,q)\equiv\Omega(F_k^{-1}(r),r,r^*,q)$.

The requirement that the integrals in (17.45) converge to finite values effectively imposes the transversality condition for the optimization. This requirement imposes parameter restrictions on the function Π and hence on the underlying preferences and technology which need to be checked for any given functional form specification.

For development of the model it is useful to employ (17.43) together with (17.45) to generate an expression for the optimal value function. As an aid to simplification, it is first noted that, by an envelope theorem applied to (17.10) and use of (17.8):

$$\Pi_\lambda(\lambda,r)=\Phi_\lambda(\lambda)+F(\hat{k})-r\hat{k},$$

where $\hat{k}$ is the optimal solution of (17.8), itself a function of λ and r. Then, comparing this result with (17.8) and (17.10), it may be noted that:

$$\Pi(\lambda,r)-\lambda\Pi_{\lambda}(\lambda,r)=\Phi(\lambda)-\lambda\Phi_{\lambda}(\lambda)\,, \tag{17.46}$$

an expression which eliminates r and which, by virtue of (17.7), may be seen to represent costate conditioned, or Frischian, instantaneous utility. Now, integrating RHS (17.45) by parts and using (17.43) and (17.46), the optimal value function may be derived in a form conditioned on λ and r as:

$$\begin{aligned}&\tilde{J}(\lambda,r,r^*,q)\\&=\frac{\lambda^{\beta_1}\int_0^{\lambda}\zeta^{-\beta_1-1}[\Phi(\zeta)-\zeta\Phi_{\lambda}(\zeta)]d\zeta+\lambda^{\beta_2}\int_{\lambda}^{\infty}\zeta^{-\beta_2-1}[\Phi(\zeta)-\zeta\Phi_{\lambda}(\zeta)]d\zeta}{\tfrac{1}{2}(\beta_2-\beta_1)(\gamma/\sigma_0)^2}\\&+\frac{\bar{\Omega}(r,r^*,q)}{\delta}.\end{aligned} \tag{17.47}$$

In order to collect the complete set of model equations into a coherent form, it is useful to record some ancillary results. Firstly, defining a latent variable, j say, to represent the endogenous (to the decision maker) component of the optimal value function, we have:

$$j=\tilde{J}(\lambda,r,r^*,q)-\frac{\bar{\Omega}(r,r^*,q)}{\delta}=\tilde{J}^0(\lambda,r,r^*,q)\,. \tag{17.48}$$

In what follows (17.48) will be referred to as the truncated optimal value function. It may be noted for later reference that the truncated optimal value function is capable of explicit evaluation by integrating the first term on the right-hand side of (17.47). Moreover, by combining (17.43) and (17.48) with the HJB equation (17.44), in view of (17.8), (17.10), (17.38), and (17.40) a (non-linear) equation for the costate variable may be written as a rearrangement of the HJB equation:

$$\lambda=\frac{\delta j-\Phi(\lambda)}{F(k)+\tfrac{1}{2}\gamma a}, \tag{17.49}$$

where a denotes the risky asset choice (17.31). From (17.35), (17.37) and (17.39a), the Arrow-Pratt coefficient of relative risk aversion (conditional on r and denoted ρ) may be characterized as:

$$\rho=-k\hat{J}_{kk}/\hat{J}_k=\frac{k}{\lambda\Psi_{\lambda\lambda}}. \tag{17.50}$$

Then, in terms of the risk aversion coefficient, the risky asset choice (17.31) may be represented as:

$$a = \frac{\gamma k}{{\sigma_0}^2 \rho}\,. \tag{17.51}$$

In bringing together these results some further latent variables are defined to simplify expressions. These additional definitions, and their relationships to the results derived above, are explained following a listing of the full set of equations for the home country component of the model, which is set out below. This listing is really a partial equilibrium subset of the complete model in that it concentrates on the home country. The paths of q, k^* and by implication $f^* \equiv F^*(k^*)$ are treated as exogenously given in this description.

$$f = F(k), \tag{17.52a}$$

$$r = F_k(k), \tag{17.52b}$$

$$\gamma = q\left[1+\chi\right] f^*/k^* - f/k\,, \tag{17.52c}$$

$$\beta_1 = \frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2} - \sqrt{\left[\frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}\right]^2 + \frac{2\delta}{(\gamma/\sigma_0)^2}}\,, \tag{17.52d}$$

$$\beta_2 = \frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2} + \sqrt{\left[\frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}\right]^2 + \frac{2\delta}{(\gamma/\sigma_0)^2}}\,, \tag{17.52e}$$

$$\phi = \Phi(\lambda), \tag{17.52f}$$

$$j = \frac{\lambda^{\beta_1}\int_0^{\lambda} \zeta^{-\beta_1-1}\left[\Phi(\zeta)-\zeta\Phi_\lambda(\zeta)\right]d\zeta + \lambda^{\beta_2}\int_{\lambda}^{\infty} \zeta^{-\beta_2-1}\left[\Phi(\zeta)-\zeta\Phi_\lambda(\zeta)\right]d\zeta}{\sqrt{\left[r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2\right]^2 + 2\delta(\gamma/\sigma_0)^2}}\,, \tag{17.52g}$$

$$\rho = \left(\frac{-\sqrt{\left[r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2\right]^2 + 2\delta(\gamma/\sigma_0)^2}}{\beta_1\lambda^{\beta_1-1}\int_0^{\lambda} \zeta^{-\beta_1-1}\left[\Phi(\zeta)-\zeta\Phi_\lambda(\zeta)\right]d\zeta + \beta_2\lambda^{\beta_2-1}\int_{\lambda}^{\infty} \zeta^{-\beta_2-1}\left[\Phi(\zeta)-\zeta\Phi_\lambda(\zeta)\right]d\zeta}\right)k, \tag{17.52h}$$

$$a = \frac{\gamma}{\sigma_0{}^2 \rho} k, \tag{17.52i}$$

$$\lambda = \frac{\delta j - \phi}{f + \gamma a / 2}, \tag{17.52j}$$

$$z = -\Phi_\lambda(\lambda), \tag{17.52k}$$

$$dk = \left\{ f + \left[\gamma - \chi_0 f^* / k^* \right] a - z \right\} dt + \sigma_0 a dv. \tag{17.52l}$$

Equation (17.52a) represents some general specification of technology for the home country. Real output is represented by a latent variable f and is modeled as a function F of the home country's capital stock subject to the conditions $F_k > 0, F_{kk} < 0$. Equation (17.52b) constructs the marginal productivity of capital as a latent variable, r, from the predetermined variable, k, given the specification of technology, F.

A series of auxiliary latent variables are then constructed from r, k and the predetermined foreign capital stock variable, k^*. The first of these, given in (17.52c), is the normalized risk premium, γ, which is a rearranged form of (17.20) using (17.11), (17.11*), (17.19) and (17.19*) to write it more directly in terms of the predetermined capital stock variables k and k^*. In addition to its dependence on k and k^*, the normalized risk premium depends upon the real exchange rate, q, and a risk adjusted productivity premium, χ. While in principle each of these may be modeled as variable parameter functions of relevant explanatory variables, or alternatively by direct construction from relevant data where available, in order to proceed with an application involving only a short time series on limited available data, in the remainder of this chapter we assume that the real exchange rate follows a diffusion process while the risk adjusted productivity premium χ and the scaled volatility σ_0 are constants. In practice, these assumptions may be accommodated by constructing the real exchange rate as time series data and by treating χ and σ_0 as parameters. We further simplify the empirical specification by constraining $\chi_0 = 0$.

Equations (17.52d) and (17.52e) are the roots of the fundamental quadratic underlying the solution to the stochastic HJB equation. These negative and positive roots, β_1 and β_2 respectively, are functions of previously defined variables (r and γ) in addition to being functionally dependent on the consumer time preference parameter, δ and the risk parameter σ. The set of equations up to (17.52e) make up a recursive sub-group within the model.

Equation (17.52f) represents the specification of preferences via a consumer profit function. The latent variable ϕ denotes the value of consumer profit. The function Φ itself depends upon a latent variable, the costate variable, λ, which is ultimately determined simultaneously within the block of equations from (17.52f) down to the HJB equation (17.52j).

To aid in the determination of λ, and exhibiting its essential simultaneity, (17.52g) and (17.52h) require evaluation of some integral expressions which are conditional on a given value of λ. Equation (17.52g) calculates the truncated (endogenous to the competitive representative agent) optimal value, defined by substitution of (17.47) into (17.48), with the denominator rewritten making use of (17.45a) and (17.45b). Equation (17.52h) calculates an expression for the coefficient of relative risk aversion (17.50), using (17.44), (17.45a) and (17.45b). For a given specification of the consumer instantaneous profit function Φ, (17.52g) and (17.52h) essentially involve the integration of functions of Frischian instantaneous utility, $U^F = \Phi - \lambda\Phi_\lambda$. Finite evaluation of these integral expressions is necessary for satisfaction of the transversality condition for the intertemporal optimization.

A general form of the transversality condition for infinite horizon problems is (see, for example, Barro and Sala-i-Martin, 1995, p.505):

$$\lim_{t\to\infty} e^{-\delta t}\hat{H}(t) = 0, \tag{17.53}$$

where $\hat{H}(t)$ denotes the optimized value of the Hamiltonian at time t. To analyze this further, note that by combination of (17.17) and (17.42), for *k* optimal as defined by (17.37):

$$\hat{H}(k,\lambda,\rho) = \delta\left[\Psi - \lambda\Psi_\lambda - \omega/\delta\right] - \tfrac{1}{2}(\gamma/\sigma_0)^2\lambda^2\Psi_{\lambda\lambda}. \tag{17.54}$$

Evaluating this expression using (17.44) gives:

$$\hat{H}(k,\lambda,\rho) = \\ \frac{\left[\delta + \tfrac{1}{2}\beta_1(\gamma/\sigma_0)^2\right]\lambda^{\beta_1}\int_0^\lambda \zeta^{-\beta_1-1}U^F(\zeta)d\zeta + \left[\delta + \tfrac{1}{2}\beta_2(\gamma/\sigma_0)^2\right]\lambda^{\beta_2}\int_\lambda^\infty \zeta^{-\beta_2-1}U^F(\zeta)d\zeta}{\sqrt{\left[r - \delta + \tfrac{1}{2}(\gamma/\sigma_0)^2\right]^2 + 2\delta(\gamma/\sigma_0)^2}}. \tag{17.55}$$

Satisfaction of the transversality condition therefore requires two things in this context. First the integral expressions in (17.55) must be capable of finite evaluation. However, these are the same integral expressions involved in the model equations (17.52g) for the truncated optimal value function and (17.52h) for the coefficient of relative risk aversion. They need to be evaluated for choice of a specific functional form. Second, even if these expressions are finite for given t, there is a restriction on their growth over time along an optimal path. Clearly this will also depend upon the functional form specification. Recall that (17.52g) constructs the truncated optimal value, defined as that component of the value of the optimized objective which is under the control of the decision making economic agent, and denoted here by the latent variable *j*, so that equation (17.52g) is an equivalent expression to the truncated optimal value function (17.48). In a similar manner, the Arrow-Pratt measure of relative risk aversion is constructed in (17.52h) as the latent variable ρ, by first deriving $\Psi_{\lambda\lambda}$ from (17.44) and then applying this to (17.50). The risky asset choice is then given in (17.52i). Note that, no matter how complex the specification of preferences and technology, the proportion of resources held in the risky asset is simply the normalized risk premium relative to the risk aversion coefficient. Finally, in (17.52j) the simultaneous loop determining the costate variable is closed by use of the HJB equation in the form (17.49), generating λ as a latent variable.

Given a value of λ determined by simultaneous solution of (17.52f)-(17.52j), equation (17.52k) then generates optimal current expenditure, *z*. This amount is then used along with previously determined variables to evaluate optimal investment in (17.52l). It should be noted that while data may be available on *z*, an alternative is to treat it as a latent variable. The same applies to the risky asset choice *a* and the real output variable *f*. The key equation ultimately is the growth equation (17.52l) and this is the only equation for which a stochastic disturbance term naturally arises from the theory.

While this model is capable of examining the effect of volatility on investment, it should not be surprising that an unambiguous analytical answer is unlikely to be available. However, it is possible to at least trace the channels of influence through the model system (17.52). For this purpose it is useful to highlight a variation of (17.52l), using (17.52i), which may be written as:

$$\frac{dk}{k} = \left\{ \frac{F(k)}{k} + \left(\frac{\gamma}{\sigma_0} \right)^2 \frac{1}{\rho} - \frac{z}{k} \right\} dt + \left(\frac{\gamma}{\sigma_0} \right) \frac{1}{\rho} d\xi \,. \tag{17.52l'}$$

Consider an increase in exogenous volatility (that is, in σ_0), with a compensating adjustment in the risk premium γ such that the standardized risk premium γ / σ_0 is held constant. Looking firstly at the instantaneous effect, since k is predetermined (17.52l') shows that such an increase in risk has no immediate impact on growth unless either z or r is instantaneously affected. Inspection of the system (17.52)—specifically of the simultaneous subsystem (17.51f), (17.51g), (17.51h) and the RHS term in (17.51j)—shows that there is no immediate impact on either of these as the simultaneous subsystem is dependent on σ_0 only through the standardized risk premium term γ / σ_0.

Turning to short run growth path responses, however, it is clear that there is a possibility of very substantial effects as k changes provided that either $F(k)$ is nonlinear in k or that preferences are more general than isoelastic. Both of these specifications lead to a change in the drift of k, via a change in $F(k)/k$ or in z/k respectively as k changes. In addition, under either of these general specifications the coefficient of relative risk aversion is functionally dependent on λ and this will also generate a variation in the subsequent growth path. Furthermore, (17.52l') shows that this latter channel of influence flows through to the diffusion of k, not simply to the drift.

These capital stock path effects on growth and volatility are not present under the special and possibly empirically unrealistic combined specifications of linearly homogeneous technology and isoelastic preferences. Under general technology and preference specifications, whether the drift and diffusion effects can be signed unambiguously would appear to be dependent on the specification of technology (17.52a) and preferences (17.52f). However, for empirically realistic specifications, which the proposed methodology is designed to allow, the size and direction of the effect of volatility on growth is likely to be an empirical matter.

The equations for the foreign country are the mirror images of the set (17.52), with starred and unstarred variables interchanged. In the context of modeling two countries, the equation groups are linked through each country's dependence on the evolution of the capital stock of the other.

17.5 Specification of the Home Country Component of the Model

In this section the system (17.52) is set out for a particular specification of technology and preferences. The specification of preferences is chosen to contain the isoelastic utility specification as a special case. The technology specification contains the typical power function specification as a special case. The general specifications are structured to enable investigation of the variability of the intertemporal elasticity of substitution (IES) and the Arrow-Pratt coefficient of relative risk aversion (RRA) over time and to examine empirically the nature of the relationship between them. In the following section a preliminary investigation is undertaken with South African data.

17.5.1 Preferences

We choose the consumer profit function specification:

$$\Phi(\lambda) = \eta \frac{\varepsilon_1 - \lambda^{1-\varepsilon_1}}{1-\varepsilon_1} + (1-\eta)\frac{\varepsilon_2 - \lambda^{1-\varepsilon_2}}{1-\varepsilon_2} \tag{17.56}$$

where η, ε_1 and ε_2 are parameters, $0 < \eta < 1$ and without loss of generality it may be assumed that $0 < \varepsilon_1 \le \varepsilon_2 < \infty$. Implied optimal costate conditioned (Frischian) expenditure is, by (4):

$$z = -\Phi_\lambda(\lambda) = \eta\lambda^{-\varepsilon_1} + (1-\eta)\lambda^{-\varepsilon_2} \tag{17.57}$$

and the intertemporal elasticity of substitution (IES) is:

$$\text{IES} = -d\log z / d\log\lambda = \frac{\varepsilon_1\eta\lambda^{-\varepsilon_1} + \varepsilon_2(1-\eta)\lambda^{-\varepsilon_2}}{\eta\lambda^{-\varepsilon_1} + (1-\eta)\lambda^{-\varepsilon_2}}\,. \tag{17.58}$$

In this specification, the IES is variable, being a weighted arithmetic mean of ε_1 and ε_2, with weights which vary with the value of the costate variable. Given $\varepsilon_1 < \varepsilon_2$, when the representative consumer is poor and λ is large, $\lambda^{-\varepsilon_1}$ dominates $\lambda^{-\varepsilon_2}$ and IES $\to \varepsilon_1$. On the other hand, when the representative consumer is rich and λ is small, the reverse applies and IES $\to \varepsilon_2$.

From (7), Frischian utility can be expressed as:

$$U^F(\lambda) = \Phi(\lambda) - \lambda\Phi_\lambda(\lambda) = \eta\varepsilon_1\frac{1-\lambda^{1-\varepsilon_1}}{1-\varepsilon_1} + (1-\eta)\varepsilon_2\frac{1-\lambda^{1-\varepsilon_2}}{1-\varepsilon_2} \tag{17.59}$$

In the special case $\varepsilon_1 = \varepsilon_2 = \varepsilon$, (17.56) corresponds to the isoelastic specification:

$$\Phi(\lambda) = \frac{\varepsilon - \lambda^{1-\varepsilon}}{1-\varepsilon} \tag{17.56'}$$

which implies:

$$z = \lambda^{-\varepsilon}, \tag{17.57'}$$

$$\text{IES} = \varepsilon \tag{17.58'}$$

and

$$U^F(\lambda) = \varepsilon\frac{1-\lambda^{1-\varepsilon}}{1-\varepsilon}\,. \tag{17.59'}$$

In this case the instantaneous utility function (in Marshallian form) is explicitly recoverable as:

$$U(z) = \frac{z^{1-1/\varepsilon} - 1}{1-1/\varepsilon}\,.$$

Although an equivalent explicit expression for utility in Marshallian form is not recoverable under the more general specification (17.56), it is clear from

(17.59) and (17.59’), where the representations in Frischian form may be compared, that utility for the representative agent in the general case (17.59) may be interpreted as a weighted average of the utility which would apply separately to the poor and the rich if the extreme poor and the extreme rich had separate isoelastic preferences with IES parameters ε_1 and ε_2 respectively.

An important point about specification (17.56) is that the IES (17.58) is a variable and will ultimately be compatible, even in the intertemporally additive expected utility maximizing context, with an Arrow-Pratt coefficient of relative risk aversion (RRA) which is not linked to it rigidly by reciprocality. This would be true even if the technology were specified to be linear. Another useful feature is that (17.57), by contrast with (17.57’) provides the prospect for considerably enhanced empirical fit.

In two-country modeling, it would be natural to assume a similar preference specification for the foreign country. In the absence of evidence on cultural differences, there is no reason why preference parameters will differ at all and although this is trivial to generalize it would be interesting to assume in the first instance that the parameters ε_1 and ε_2 are identical in the two countries. That is, the extreme rich could be assumed to have similar preferences with respect to current versus future consumption in the two countries, and the same might reasonably apply across the two countries for the extreme poor. However, we allow the general preference specification of the representative consumer to be made up of a country-specific blend of these extremes, representing different wealth distributions. Of course, the costate variable will also, in general take a different value at any point in time in the two countries. The preference specification for the foreign country could therefore be represented as:

$$\Phi^*(\lambda^*) = \eta^* \frac{\varepsilon_1 - \lambda^{*1-\varepsilon_1}}{1-\varepsilon_1} + (1-\eta^*)\frac{\varepsilon_2 - \lambda^{*1-\varepsilon_2}}{1-\varepsilon_2}\,. \tag{17.56*}$$

17.5.2 Technology

We choose the specification:

$$F(k) = \frac{\alpha_2 k^{\theta_2} + \left[2\alpha_1 k^{\theta_1} - \alpha_2 k^{\theta_2}\right]e^{-k}}{1+e^{-k}}\,. \tag{17.60}$$

We assume $0 \le \theta_2 \le \theta_1 < 1$, $0 < \alpha_1 \le \alpha_2$. As $k \to 0$ the technology is dominated by $F(k) \to \alpha_1 k^{\theta_1}$, while as $k \to \infty$ it is dominated by $F(k) \to \alpha_2 k^{\theta_2}$. This function therefore nests standard specifications at its extremes and allows for more complex technology to apply for finite *k*. Under the parameter restrictions $\theta_1 = \theta_2 = \theta$ and $\alpha_1 = \alpha_2 = \alpha$, it reduces to the commonly employed power specification for all *k*,

$$F(k) = \alpha k^{\theta} \, . \tag{17.60'}$$

If $\theta_1 = \theta_2 = 0$, the function has the logistic form:

$$F(k) = \frac{\alpha_2 + (2\alpha_1 - \alpha_2)e^{-k}}{1 + e^{-k}} \tag{17.60''}$$

with $F(k)$ ranging from α_1 at $k = 0$ to α_2 as $k \to \infty$.

In the case of specification (17.60), the definition of ρ in (17.11) implies:

$$F_k(k) = \frac{\alpha_2\theta_2 k^{\theta_2 - 1} + 2\left[\alpha_1\theta_1 k^{\theta_1 - 1} - \alpha_1 k^{\theta_1} + \alpha_2 k^{\theta_2}\right]e^{-k} + \left[2\alpha_1\theta_1 k^{\theta_1 - 1} - \alpha_2\theta_2 k^{\theta_2 - 1}\right]e^{-2k}}{\left(1 + e^{-k}\right)^2} \, ,$$

a result which is used to endogenize r as a latent variable in (17.70b) below. Given the generality of specification (17.60), the positivity of F_k and, in an even more complex fashion, the negativity of F_{kk} (not displayed) will in general be parameter and data value dependent, and satisfaction of these restrictions may be checked as a test of the model empirically. Of course, in the special cases it is trivial to check that these restrictions will be automatically satisfied under the parameter value conditions specified above.

It is convenient at this point to record the mirror image specification of technology for the foreign country in the two-country modeling case (with, of course, potentially different technological parameters):

$$F^*(k^*) = \frac{\alpha_2^* k^{*\theta_2^*} + \left[2\alpha_1^* k^{*\theta_1^*} - \alpha_2^* k^{*\theta_2^*}\right]e^{-k^*}}{1 + e^{-k^*}} \, . \tag{17.60*}$$

17.5.3 Integral Evaluations and Transversality Conditions

In order to employ specifications (44) and (48) to construct a specific model of the equation set (40) it is necessary to provide expressions for the evaluation of the integral expressions which are involved in (40g) and (40h) under specifications (44) and (17.60). Utilizing (17.59), it can be seen that these involve evaluation of expressions, i_0 and i_∞ say, where:

$$i_0 = \int_0^{\lambda} \zeta^{-\beta_1 - 1}\left[\eta\varepsilon_1 \frac{1 - \zeta^{1-\varepsilon_1}}{1 - \varepsilon_1} + (1 - \eta)\varepsilon_2 \frac{1 - \zeta^{1-\varepsilon_2}}{1 - \varepsilon_2}\right] d\zeta \tag{17.61}$$

$$i_\infty = \int_\lambda^\infty \zeta^{-\beta_2 - 1}\left[\eta\varepsilon_1 \frac{1-\zeta^{1-\varepsilon_1}}{1-\varepsilon_1} + (1-\eta)\varepsilon_2 \frac{1-\zeta^{1-\varepsilon_2}}{1-\varepsilon_2}\right] d\zeta \tag{17.62}$$

Noting the definitions of β_1 and β_2 given in (17.52d) and (17.52e), it can be shown after considerable manipulation that:

$$\begin{aligned}
i_0 = {} & \tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{\beta_2}{\delta}\left[\eta\frac{\varepsilon_1}{1-\varepsilon_1} + (1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right]\lambda^{-\beta_1} \\
& + \tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{1-\beta_2-\varepsilon_1}{\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]}\left(\frac{\eta\varepsilon_1}{1-\varepsilon_1}\right)\left\{\zeta^{1-\beta_1-\varepsilon_1}\Big|_{\zeta=\lambda} - \zeta^{1-\beta_1-\varepsilon_1}\Big|_{\zeta=0}\right\} \\
& + \tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{1-\beta_2-\varepsilon_2}{\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}\left(\frac{(1-\eta)\varepsilon_2}{1-\varepsilon_2}\right)\left\{\zeta^{1-\beta_1-\varepsilon_2}\Big|_{\zeta=\lambda} - \zeta^{1-\beta_1-\varepsilon_2}\Big|_{\zeta=0}\right\}
\end{aligned} \tag{17.61'}$$

In order for i_0 to return a finite value it is necessary for the restrictions $1-\beta_1-\varepsilon_1 > 0$ and $1-\beta_1-\varepsilon_2 > 0$ to ensure that the terms $\zeta^{1-\beta_1-\varepsilon_1}\big|_{\zeta=0}$ and $\zeta^{1-\beta_1-\varepsilon_2}\big|_{\zeta=0}$ do not go to infinity. Since $\varepsilon_2 \geq \varepsilon_1 > 0$, a sufficient condition is the second of these, that involving ε_2, as this then implies satisfaction of the first. On manipulation of the definition of β_1 the condition may be written:

$$\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right] > 0\,. \tag{17.63}$$

If $\varepsilon_2 < 1$ this condition is necessarily satisfied. However, if $\varepsilon_2 > 1$ the condition may be violated. Evaluating the positive root of the above quadratic, the condition may be written as:

$$\varepsilon_2 < \tfrac{1}{2} + \frac{\delta - r}{(\gamma/\sigma_0)^2} + \sqrt{\left[\tfrac{1}{2} + \frac{\delta - r}{(\gamma/\sigma_0)^2}\right]^2 + \frac{2r}{(\gamma/\sigma_0)^2}}\,. \tag{17.64}$$

In a similar manner, it can be shown that:

$$\begin{aligned} i_{\infty} = & -\tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{\beta_1}{\delta}\left[\eta\frac{\varepsilon_1}{1-\varepsilon_1} + (1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right]\lambda^{-\beta_2} \\ & -\tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{1-\beta_1-\varepsilon_1}{\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]}\left(\eta\frac{\varepsilon_1}{1-\varepsilon_1}\right)\left\{\zeta^{1-\beta_2-\varepsilon_1}\Big|_{\zeta=\infty} - \zeta^{1-\beta_2-\varepsilon_1}\Big|_{\zeta=\lambda}\right\} \\ & -\tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{1-\beta_1-\varepsilon_2}{\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}\left((1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right)\left\{\zeta^{1-\beta_2-\varepsilon_2}\Big|_{\zeta=\infty} - \zeta^{1-\beta_2-\varepsilon_2}\Big|_{\zeta=\lambda}\right\} \end{aligned} \tag{17.62'}$$

In this case, to provide a finite value for i_{∞} it is necessary that the terms $\zeta^{1-\beta_2-\varepsilon_1}\big|_{\zeta=\infty}$ and $\zeta^{1-\beta_2-\varepsilon_2}\big|_{\zeta=\infty}$ do not explode. These conditions require $1-\beta_2-\varepsilon_1 < 0$ and $1-\beta_2-\varepsilon_2 < 0$. Now, given $\varepsilon_1 \le \varepsilon_2$, satisfaction of the first condition implies the second. On manipulation of the definition of β_2, the condition can be written as:

$$\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right] > 0\,. \tag{17.65}$$

This is the same condition as that for ε_2 which was given as (17.63) and re-expressed as an upper bound in (17.64). Since we may assume without loss of generality that $\varepsilon_1 \le \varepsilon_2$ it is sufficient, given this, to enforce (17.64) to ensure satisfaction of a finite evaluation of both i_0 and i_{∞} and hence of (17.52g). It can be shown that this condition, which ensures a finite value for the optimal value function, is equivalent to ensuring satisfaction of the transversality condition for the existence of a solution to the optimization problem. Given restriction (17.64), $\lambda^{\beta_1} i_0$ and $\lambda^{\beta_2} i_{\infty}$ may be evaluated as:

$$\begin{aligned} \lambda^{\beta_1} i_0 = & \tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{\beta_2}{\delta}\left[\eta\frac{\varepsilon_1}{1-\varepsilon_1} + (1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right] \\ & + \frac{\tfrac{1}{2}(\gamma/\sigma_0)^2(1-\beta_2-\varepsilon_1)}{\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]}\left(\eta\frac{\varepsilon_1}{1-\varepsilon_1}\right)\lambda^{1-\varepsilon_1} \\ & + \frac{\tfrac{1}{2}(\gamma/\sigma_0)^2(1-\beta_2-\varepsilon_2)}{\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}\left((1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right)\lambda^{1-\varepsilon_2} \end{aligned} \tag{17.66}$$

$$\begin{aligned}\lambda^{\beta_2} i_\infty = &-\tfrac{1}{2}\left(\frac{\gamma}{\sigma_0}\right)^2 \frac{\beta_1}{\delta}\left[\eta\frac{\varepsilon_1}{1-\varepsilon_1}+(1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right] \\ &-\frac{\tfrac{1}{2}(\gamma/\sigma_0)^2(1-\beta_1-\varepsilon_1)}{\varepsilon_1\delta+(1-\varepsilon_1)\left[r+\tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]}\left(\eta\frac{\varepsilon_1}{1-\varepsilon_1}\right)\lambda^{1-\varepsilon_1} \\ &-\frac{\tfrac{1}{2}(\gamma/\sigma_0)^2(1-\beta_1-\varepsilon_2)}{\varepsilon_2\delta+(1-\varepsilon_2)\left[r+\tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}\left((1-\eta)\frac{\varepsilon_2}{1-\varepsilon_2}\right)\lambda^{1-\varepsilon_2},\end{aligned} \tag{17.67}$$

and the finiteness of these expressions ensures the finiteness of the Hamiltonian (17.55) under specification (17.56) for all t. These expressions contain similar power terms in λ to the Frisch utility function, and a finite evaluation of the objective function for the intertemporal optimization is therefore equivalent to satisfaction of the transversality condition. However, the finite value of the truncated optimal value function may actually be constructed, given restriction (17.64), as the evaluation of (17.52g) under specification (17.56). Specifically, in view of definitions (17.61) and (17.62), (17.52g) may be written:

$$j=\frac{\lambda^{\beta_1} i_0+\lambda^{\beta_2} i_\infty}{\sqrt{\left[r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2\right]^2+2\delta(\gamma/\sigma_0)^2}}. \tag{17.68}$$

However, using (17.66) and (17.67) together with (17.52d) and (17.52e), (17.68) simplifies to:

$$\begin{aligned} j=&\,\eta\left(\frac{\varepsilon_1}{1-\varepsilon_1}\right)\left[\frac{1}{\delta}-\frac{\lambda^{1-\varepsilon_1}}{\varepsilon_1\delta+(1-\varepsilon_1)\left[r+\tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]}\right] \\ &+(1-\eta)\left(\frac{\varepsilon_2}{1-\varepsilon_2}\right)\left[\frac{1}{\delta}-\frac{\lambda^{1-\varepsilon_2}}{\varepsilon_2\delta+(1-\varepsilon_2)\left[r+\tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}\right].\end{aligned} \tag{17.68'}$$

In an analogous manner, for this specification equation (17.52h) becomes:

$$\rho=\frac{-\lambda k\sqrt{\left[r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2\right]^2+2\delta(\gamma/\sigma_0)^2}}{\beta_1\lambda^{\beta_1} i_0+\beta_2\lambda^{\beta_2} i_\infty}, \tag{17.69}$$

and once again using (17.66), (17.67), (17.52d) and (17.52e), this allows (17.69) to be simplified to:

$$\rho = \frac{k}{\dfrac{\eta\varepsilon_1\lambda^{-\varepsilon_1}}{\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]} + \dfrac{(1-\eta)\varepsilon_2\lambda^{-\varepsilon_2}}{\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}}. \tag{17.69'}$$

17.5.3 The Complete Model of the Home Country

To summarize the specification under (17.56) and (17.60), the home country model equation system (17.52), treating relevant foreign country information as exogenous, becomes:

$$f = \frac{\alpha_2 k^{\theta_2} + \left[2\alpha_1 k^{\theta_1} - \alpha_2 k^{\theta_2}\right]e^{-k}}{1+e^{-k}}, \tag{17.70a}$$

$$r = \frac{\alpha_2\theta_2 k^{\theta_2-1} + 2\left[\alpha_1\theta_1 k^{\theta_1-1} - \alpha_1 k^{\theta_1} + \alpha_2 k^{\theta_2}\right]e^{-k} + \left[2\alpha_1\theta_1 k^{\theta_1-1} - \alpha_2\theta_2 k^{\theta_2-1}\right]e^{-2k}}{\left(1+e^{-k}\right)^2}, \tag{17.70b}$$

$$\gamma = q[1+\chi]f^*/k^* - f/k, \tag{17.70c}$$

$$\beta_1 = \frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2} - \sqrt{\left[\frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}\right]^2 + \frac{2\delta}{(\gamma/\sigma_0)^2}}, \tag{17.70d}$$

$$\beta_2 = \frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2} + \sqrt{\left[\frac{r-\delta+\tfrac{1}{2}(\gamma/\sigma_0)^2}{(\gamma/\sigma_0)^2}\right]^2 + \frac{2\delta}{(\gamma/\sigma_0)^2}}, \tag{17.70e}$$

$$\phi = \eta\frac{\varepsilon_1 - \lambda^{1-\varepsilon_1}}{1-\varepsilon_1} + (1-\eta)\frac{\varepsilon_2 - \lambda^{1-\varepsilon_2}}{1-\varepsilon_2}, \tag{17.70f}$$

$$\begin{aligned} j = {} & \eta\left(\frac{\varepsilon_1}{1-\varepsilon_1}\right)\left[\frac{1}{\delta} - \frac{\lambda^{1-\varepsilon_1}}{\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \tfrac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]}\right] \\ & +(1-\eta)\left(\frac{\varepsilon_2}{1-\varepsilon_2}\right)\left[\frac{1}{\delta} - \frac{\lambda^{1-\varepsilon_2}}{\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \tfrac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}\right], \end{aligned} \tag{17.70g}$$

$$\rho = \frac{k}{\dfrac{\eta\varepsilon_1\lambda^{-\varepsilon_1}}{\varepsilon_1\delta + (1-\varepsilon_1)\left[r + \frac{1}{2}\varepsilon_1(\gamma/\sigma_0)^2\right]} + \dfrac{(1-\eta)\varepsilon_2\lambda^{-\varepsilon_2}}{\varepsilon_2\delta + (1-\varepsilon_2)\left[r + \frac{1}{2}\varepsilon_2(\gamma/\sigma_0)^2\right]}}, \tag{17.70h}$$

$$a = (1/\sigma_0)\left(\frac{\gamma}{\sigma_0}\right)\frac{k}{\rho}, \tag{17.70i}$$

$$\lambda = \frac{\delta j - \phi}{f + \gamma a/2} \tag{17.70j}$$

$$z = \eta\lambda^{-\varepsilon_1} + (1-\eta)\lambda^{-\varepsilon_2} \tag{17.70k}$$

$$dk = \{f + \gamma a - z\}dt + \sigma_0 a d\xi\,. \tag{17.70l}$$

In this system, (17.70d) and (17.70e) provide information only. The particular specification of preferences allows direct evaluation of the integrals in (17.52g) and (17.52h), allowing elimination of the latent variables β_1 and β_2 in the construction of (17.70g) and (17.70h).

17.6 Estimation

To date we have been able to investigate the model empirically only in single-economy form for South Africa. The structural model of the "home country" elaborated above comprises a single equation of motion - a stochastic differential equation (SDE), equation (17.70l) - and several zero-order equilibrium relationships, equations (17.70a)-(17.70k). To estimate the parameters of this system of equations from discrete observations, we work with the following transformation of (17.70l),

$$D(k/a) = f/a + \gamma - z/a - k \cdot Da/a^2 + \sigma_0 u(t)\,, \tag{17.71}$$

in which D is the differential operator, *d/dt*. This transformation serves to isolate a from the Wiener process. Equations (17.71) can be shown to possess the same properties as (17.70l) if the disturbances *u(t)* are treated as if they were well determined—i.e., if they were generated by a stationary process with constant spectral density so that the integral $\xi(t) = \int_0^t u(s)ds$ is such that $\xi(t)$ is a homogeneous random process with uncorrelated increments, where $\mathrm{E}\{\xi(t_1)\} = 0$ for all t and

$\mathrm{E}\{[\xi(t_1)-\xi(t_2)][\xi(t_3)-\xi(t_4)]'\}=0$ for $t_1>t_2\geq t_3>t_4$. (See Wymer (1972) for a precise interpretation of the disturbances $\xi(t)$ and their relationship to $u(t)$.)

The literature on estimating systems of linear SDEs, from Wymer (1972) onwards, is by now well developed. There is little besides Wymer (1993) in the econometrics literature on the estimation of systems of *non-linear* SDEs, however. In view of the nature of the model, an existing non-linear estimation program of Wymer, discussed in Wymer (1993), may be used to obtain serviceable approximations to quasi-full-information-maximum-likelihood (quasi-FIML) estimates of the model's parameters. The adequacy of this approach in the present context remains to be evaluated.

To estimate the model we employ published annual data from 1980-1998 on GDP, capital formation and consumption, aggregates of non-investment items in national income identities, direct investment abroad, the rand/US$ exchange rate, and price deflators, for South Africa and "the industrial economies", as reported on by the IMF, which stand in as "country two" or the rest of the world.[12]

The exact quasi-FIML non-linear continuous-time estimator of Wymer (1993) is implemented (in his program ESCONA) according to a two-step algorithm.

1. For a given set of parameter estimates (or initial values) the equation system is integrated forward over each observation interval by a numerical variable-order, variable-step Adams method, residuals are computed by comparing the one-period-forward solution values with the observed values for variables on which there are observations (for latent variables, of course, there will be none), and the variance covariance matrix is then formed.

2. The natural logarithm of the determinant of the variance-covariance matrix is minimized by a quasi-Newton method to update parameter estimates. Convergence criteria are then checked and, if not met, another iteration is begun.

Across-equation restrictions that are implied by theory are, of course, imposed in estimation. In estimating the model we have also imposed constraints on the values that parameters can assume. Estimates of all parameters must be non-

---

[12] The data were obtained from various IMF and OECD publications. All real quantities were deflated by appropriate 1990 GDP price deflators. The real exchange rate was constructed by dividing the nominal rand/US$ exchange rate by the South African 1990 GDP price deflator and multiplying by the industrialized economies' 1990 GDP price deflator. Series on capital stocks were constructed by the "perpetual inventory" method on base stocks that conform to the stylized facts of industrialized economies. Capital consumption is not reported for South Africa. We estimated this variable from the ratio of capital consumption to gross capital formation for the industrialized countries. This practice may introduce distortions in capital stock figures, but we deem it superior to the use of a constant percentage or an estimated depreciation parameter.

negative, $\hat{\alpha}_1$ may not exceed $\hat{\alpha}_2$, $\hat{\theta}_2$ may not exceed $\hat{\theta}_1$, $\hat{\varepsilon}_1$ may not exceed $\hat{\varepsilon}_2$ and $\hat{\eta}$ must be less than unity.

The estimated model tracks the historical data reasonably well, given the volatility of the period, with proportionate root-mean-square errors in static in-sample forecasts of real variables as presented in Table 17.1.

Table 17.1. Static in-sample proportionate root-mean-square errors (RMSEs) of estimates of observed variables

Variable	Proportionate RMSE
k/a	0.02008
f	0.07502
a	0.09877
z	0.19677

The quasi-FIML parameter estimates and the estimates of their asymptotic standard errors are given in Table 17.2.

Table 17.2. Quasi-FIML Parameter Estimates[13]

Parameter	Parameter Estimate	Asymptotic_Standard Error (Estimate)
α_1	0.19918	0.03291
α_2	0.19917	0.09151
θ_1	1.94029	0.14873
θ_2	0.99302	0.10115
ε_1	0.73326	0.04265
ε_2	1.36570	0.05223
χ	0.00647	0.00128
δ	0.00029*	0.00023
η	0.74494	0.01333
σ_0^2	3.02186	0.09006

There are several troubling aspects to the parameter estimates in Table 17.2. Of most concern is the estimate of the subjective time preference rate, $\hat{\delta}$, which is not significantly different from zero. At the same time, the estimate of the asymptotic elasticity of substitution for the extremely wealthy, $\hat{\varepsilon}_2$, is larger than 1.0. The

---

[13] * denotes an estimate not statistically discernible from zero at the 0.05 level of significance.

combination of these results contributes to the violation of the transversality condition (17.63) at all points in the sample (although, as illustrated in Table 17.3, the computed value of the optimized value function, *j*, is finite at every data point).

It would have been possible to maintain the more general preference specification compatibly with satisfaction of the transversality condition by enforcing a further restriction $\hat{\varepsilon}_2 < 1$. However, this would have done violence to what the data are saying. Our preferred option is to refine the data series and then consider alternative parameterizations. Recall that in order to simplify the specification in our initial investigations we are treating χ as a constant parameter, constraining $\chi_0 = 0$ and additionally treating $\sigma = \sigma_0 q M^*(r)$ with σ_0 constant, all of which are somewhat arbitrary (though these simplifications are intended to allow us to focus on the generality of the functional form specifications for preferences and technology). Subsequently, it will be important to investigate alternative generalizations of the isoelastic specification, and additionally conduct a more extensive search over a complex likelihood surface in what is a reasonably high dimensional parameter space. For present purposes, we continue the discussion conceding that the parameter estimate of ε_2 is too high for comfort, that of δ is too low for comfort and that the combination of these violates transversality.

Before closing the discussion of estimates of the utility function, we should note that the difference between $\hat{\varepsilon}_1$ at 0.73 and $\hat{\varepsilon}_2$ at 1.36 is of considerable interest, and points to a tendency in the data to support a variable elasticity of substitution which rises with wealth.

We turn now to the technology estimates and some additional overall implications of the combination of preference and technology estimates for the values of some key latent variables. We note firstly in passing (and discuss more fully below) that the results support the existence of a non-linear technology and hence (subject to the caveat of the preference parameter violations of optimality) the results tend to support a decoupling of strict reciprocality of the intertemporal elasticity of substitution (IES) and the Arrow-Pratt coefficient of relative risk aversion applied to resource risk (RRA), *r*. Results supporting the breaking of this link are reported in Table 17.3, where estimates of some key latent variables are presented for 1982-1997.

As pointed out above, the implication of the parameter estimates, is that the there is little variation in IES. Table 17.2 gives the calculation of this "constant" IES at approximately 0.98. This value is a weighted average of 0.73 and 1.37, using weights given in equation (17.57). These weights depend upon the value of the costate variable λ, which has some variation—ranging from 0.27 to 0.45—but is close to 0.37 on average over the examined time period. The large value of ε_2 at 1.37 would have lead to considerable variation in the IES were it not for the negation of this effect by the multiplicative term $(1-\eta)$ in (17.57).

Table 17.3. Time Series Estimates of Some Key Latent Variables

	λ	r	IES	ρ (RRA)	j
1982	0.37829	0.19404	0.97845	17.26141	3717.43612
1983	0.35743	0.19403	0.98386	22.17611	3717.79835
1984	0.29374	0.19404	1.00286	18.35782	3724.61645
1985	0.27000	0.19406	1.01113	18.59111	3727.91335
1986	0.31562	0.19403	0.99586	17.70923	3722.15049
1987	0.35262	0.19404	0.98516	16.79986	3719.03293
1988	0.31759	0.19402	0.99525	24.87270	3720.74674
1989	0.36718	0.19405	0.98129	10.23210	3720.10663
1990	0.39163	0.19406	0.97517	9.68007	3718.94711
1991	0.40016	0.19407	0.97314	9.48419	3718.62884
1992	0.40968	0.19407	0.97093	8.43943	3718.73879
1993	0.39655	0.19407	0.97399	8.76501	3719.16582
1994	0.40485	0.19407	0.97204	8.53310	3718.90065
1995	0.45388	0.19410	0.96140	6.54914	3718.35408
1996	0.40245	0.19409	0.97260	5.79011	3721.20025
1997	0.42284	0.19410	0.96797	5.14260	3720.98662

The equation for the RRA is given as one of the model's latent variable equations, (17.70h). It is the second term in the denominator of (17.70h) which is problematic, as this term's divisor, which should be positive in view of the transversality condition (17.63), is in fact negative due to the unfortunate combination of parameter estimates discussed above. Of course, in (17.70h) the multiplicative term $(1-\eta)$, being estimated as a small number, acts to reduce the influence of the transversality violating term (although, as conceded above, it does not do this completely successfully and consequently suggests that rationality is violated in the sense that the full set of necessary optimality conditions are not satisfied).

As noted above (and subject to the caveat of the violation of rationality), the implied IES is virtually constant. Despite this, the RRA is very clearly not equal to the reciprocal of the IES and fluctuates considerably over the sample period. Examination of the formulae for the IES and the RRA (and a rather complex expression for the capital elasticity of expenditure which is implied by (17.70k) together with the simultaneous solution of (17.70a)-(17.70j) which theoretically provides λ as a function of k) shows that this lack of reciprocality is due to the divergence of the technology from linearity rather than to the violation of rationality. That is, for the specifications considered here, the capital elasticity of expenditure (which is the natural analogue in the representative consumer-firm context of the wealth elasticity of consumption discussed in Section 2) is decidedly non-unity.

Subject to any caveats on the precision of estimation of the asymptotic standard errors of parameters, the technology appears to be significantly different from asymptotically linear. It is also non-linear for finite k.

It is this non-linearity which produces estimates of a relatively high degree of risk aversion compatibly with a high degree of intertemporal substitutability. These kinds of results can help in explaining the equity premium paradox because they can justify a high equity premium to provide compensation for high risk aversion at the same time as they allow a substantial willingness to substitute current for future consumption.

17.7 Conclusion

In this paper we have sought to provide the theoretical background to an examination of the relationship between volatility and growth from a representative agent modeling perspective. We have argued firstly that the issues may be fruitfully examined within the context of the intertemporally additive expected utility maximizing paradigm, provided that sufficiently general functional forms are employed. We have then proposed and developed at some length a methodology based on reasonably extensive use of results in duality theory to enable the complete derivation of estimating forms which are consistent with these microeconomic foundations. Finally, we have undertaken a preliminary empirical investigation using South African data.

The issue of the relative weight which policy makers should give to reduction in variability at the apparent expense of growth versus opting for growth at the apparent expense of variability is one of vital importance to many countries, but particularly to countries which are at crucial stages of economic and social development such as South Africa. The microeconomic perspective enables the conflicting criteria to be examined in an overall consistent framework under a utility maximizing objective in which welfare variations due to policy changes can be explicitly examined. The preliminary estimates obtained and discussed in this paper suggest that the proposed methodology can provide a viable means of addressing these important issues. However, problems with the current preliminary set of parameter estimates do not warrant their use at this stage in providing empirical evidence on the link between volatility and growth. The results to date point to the need for an ongoing effort to refine the data and to generalize to a point of empirical satisfaction the preference relationships and technologies which are specified and maintained behind these types of relatively sophisticated microeconomic foundations-based analyses.

The methodology described in this paper has been researched and purpose-built precisely to enable generalizations of the tightly theoretically specified functional forms which are now current in the theoretical growth literature in order to align them more closely with the probably more flexible specifications which will be required to offer practical input to the policy debate. The reported results are mixed but they point clearly to a promising research path.

Acknowledgement

We wish to acknowledge invaluable advice and assistance from Clifford Wymer in estimating the model, including custom software programming.

References

Abel AB (1991) The equity premium puzzle. Business Review: Federal Reserve Bank of Philadelphia, Sept/Oct:3-14

Arrow KJ (1965) Aspects of a theory of risk bearing. Yrjo Jahnsson Lectures, Helsinki. Reprinted in: Arrow KJ (1971) Essays in the Theory of Risk Bearing. Markham, Chicago

Attanasio OP, Weber G (1989) Intertemporal substitution, risk aversion and the Euler equation for consumption. The Economic Journal 99:59-73

Barro RJ, Sala-i-Martin X (1995) Economic growth. McGraw-Hill, New York

Browning M (1982) Profit function representations for consumer preferences. Discussion paper 82/125, Bristol University

Browning M, Deaton A, Irish M (1985) A profitable approach to labor supply and commodity demands over the life cycle. Econometrica 53(3):503-543

Constantinides GM (1990) Habit formation: A resolution of the equity premium puzzle. Journal of Political Economy 98(3):519-543

Cooper RJ (2001) General structural dynamic economic modeling. Macroeconomic Dynamics 5:647-672

Deaton A (1992) Understanding consumption. Clarendon lectures in economics. Oxford University Press, Oxford

Epstein LG, Zin SE (1989) Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework.. Econometrica 57(4):937-969

Epstein LG, Zin SE (1991) Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. Journal of Political Economy 99(2):263-286

Grossman SJ, Laroque G (1990) Asset pricing and optimal portfolio choice in the presence of illiquid durable consumption goods. Econometrica 58(1):25-51

Kocherlakota NR (1996) The equity premium: It's still a puzzle. Journal of Economic Literature XXXIV:42-71

Lucas RE (1987) Models of Business Cycles. Yrjo Johansson Lectures, Helsinki. Basil Blackwell, Oxford

Lucas RE (1988) On the mechanics of economic eevelopment. Journal of Monetary Economics 22:3-42

Mankiw GN (1986) The equity premium and the concentration of aggregate shocks. Journal of Financial Economics 17:211-219

Mehra R, Prescott EC (1985) The equity premium: A puzzle. Journal of Monetary Economics 15:145-161

Naik NY, Moore MJ (1996) Habit formation and intertemporal substitution in individual food Consumption. Review of Economics and Statistics LXXVIII:321-328

Obstfeld M (1994) Evaluating risky consumption paths: The role of intertemporal substitutability. European economic review 38:1471-1486

Pratt JW (1964) Risk aversion in the small and in the large. Econometrica 32:122-136

Ramey G, Ramey VA (1995) Cross-country evidence on the link between volatility and growth. American Economic Review 85(5):1138-1151

Romer PM (1986) Increasing returns and long-run growth. Journal of Political Economy 94(5):1002-1037

Turnovsky SJ (1995) Methods of Macroeconomic Dynamics. MIT Press, Cambridge, MA

Weil P (1989) The equity premium puzzle and the risk-free puzzle. Journal of Monetary Economics 24:401-421

Weil P (1990) Nonexpected utility in macroeconomics. Quarterly Journal of Economics CV(1):29-42

Wymer CR (1972) Econometric estimation of stochastic differential equation systems. Econometrica 40:565-577

Wymer CR (1993) Estimation of non-linear continuous time models from discrete data. In: Phillips P, Hall V (eds) Models, methods, and applications of econometrics. Basil Blackwell, Oxford

List of Contributors

Anderson, William P.
Department of Geography and Center for Transportation Studies
Boston University
Stone Science Building,
675 Commonwealth Avenue
Boston, MA 02215
USA
bander@bu.edu

Balta-Ozkan, Nazmiye
Policy Studies Institute
50 Hanson Street
London, W1W 6UP
UK
n.ozkan@psi.org.uk

Cooper, Russel J.
Department of Economics
Macquarie University
Sydney NSW 2109
Australia
rcooper@efs.mq.edu.au

Cruz, Jose
Operations and Information Management Department
School of Business
University of Connecticut
2100 Hillside Road
Storrs, CT
USA
Jose.Cruz@business.uconn.edu

Donaghy, Kieran P.
Department of City and Regional Planning
Cornell University
106 West Sibley Hall
Ithaca, NY 14853
USA
kpd23@cornell.edu

van Duin, Jan H.R.
Delft University of Technology
Faculty of Technology, Policy and Management, Room 4.3.090
P.O. Box 5015, 2600 GA Delft
The Netherlands
j.h.r.vanduin@tbm.tudelft.nl

Feser, Edward
Department of Urban and Regional Planning
University of Illinois at Urbana-Champaign
111 Temple Hoyne Buell Hall
611 Lorado Taft Drive
Champaign, IL 61820
USA
feser@uiuc.edu

Fingleton, Bernard
Department of Land Economy
University of Cambridge
Silver Street
Cambridge CB3 9EP
UK
bf100@cam.ac.uk

Giesecke, James A.
Centre of Policy Studies
Monash University
Melbourne VIC 3800
Australia
James.Giesecke@BusEco.monash.edu.au

Guo, Dong
Laboratoire d'Economie et de Gestion
Université de Bourgogne
Pôle d'Economie et de Gestion
2, boulevard Gabriel
BP 26611 -21066 Dijon Cedex
France
dguo@uiuc.edu

Hewings, Geoffrey J.D.
Regional Economics Applications Library
Davenport Hall
University of Illinois at Urbana-Champaign
Urbana, IL 61801-3671
USA
hewings@uiuc.edu

Higano, Yoshiro
Institute of Agricultural and Forest Engineering
University of Tsukuba
Japan
higano@jsrsai.envr.tsukuba.ac.jp

Li, Bin
Department of Biology
University of Jiaxing
China
lilv963@hotmail.com

Madden, Gary
Communications Economics and Electronic Markets Research Centre
Department of Economics
Curtin University of Technology
GPO Box U1987
Perth WA 6845
Australia
maddeng@cbs.curtin.edu.au

Madden, John R.
Centre of Policy Studies
Monash University
Melbourne, VIC 3800
Australia
John.Madden@BusEco.monash.edu.au

McCann, Philip
Waikato Management School
University of Waikato
Private Bag 3105, Hamilton
New Zealand
pmccann@ waikato.ac.nz

McQuaid, Ronald W.
Employment Research Institute,
Napier University,
Craiglockhart Campus,
Edinburgh, EH14 1DJ,
Scotland, UK
r.mcquaid@napier.ac.uk

Munroe, Darla K.
Department of Geography
The Ohio State University
154 N Oval Mall
Columbus, OH 43210-1361
USA
munroe.9@osu.edu

Nagurney, Anna
Department of Finance and Operations Management
Isenberg School of Management
University of Massachusetts
Amherst, MA 01003
USA
nagurney@gbfin.umass.edu

Nijkamp, Peter
Department of Spatial Economics
Free University
De Boelelaan 1105 1081 HV
Amsterdam
The Netherlands
pnijkamp@vuw.nl.ac

Sue Wing, Ian
Department of Geography and Center for Energy and Environmental Studies
Boston University
Stone Science Building, 675 Commonwealth Avenue
Boston, MA 02215
USA
isw@bu.edu

Vickerman, Roger
Department of Economics
University of Kent
Canterbury, Kent CT2 7NF
UK
R.W.Vickerman@ukc.ac.uk

Wakolbinger, Tina
Department of Finance and Operations Management
Isenberg School of Management
University of Massachusetts
Amherst, MA 01003
USA
wakolbinger@som.umass.edu

Wang, Shunli
Department of Spatial Economics
Free University
De Boelelaan 1105 1081 HV
Amsterdam
The Netherlands
swang@bncconsulting.com

van Wee, Gijsbertus P.
Delft University of Technology
Faculty of Technology, Policy and Management
Jaffalaan 5, 2628 BX, Delft
The Netherlands
g.p.vanwee@tbm.tudelf.nl

Wixted, Brian
Centre for Policy Research on Science and Technology
Simon Fraser University
515 West Hastings Street
Vancouver BC V6B 5K3
Canada
wixted@shaw.ca

Wymer, Clifford R.
6 Muritai Road
Takapuna
Auckland 9
New Zealand
wymer@mail.com

Subject Index