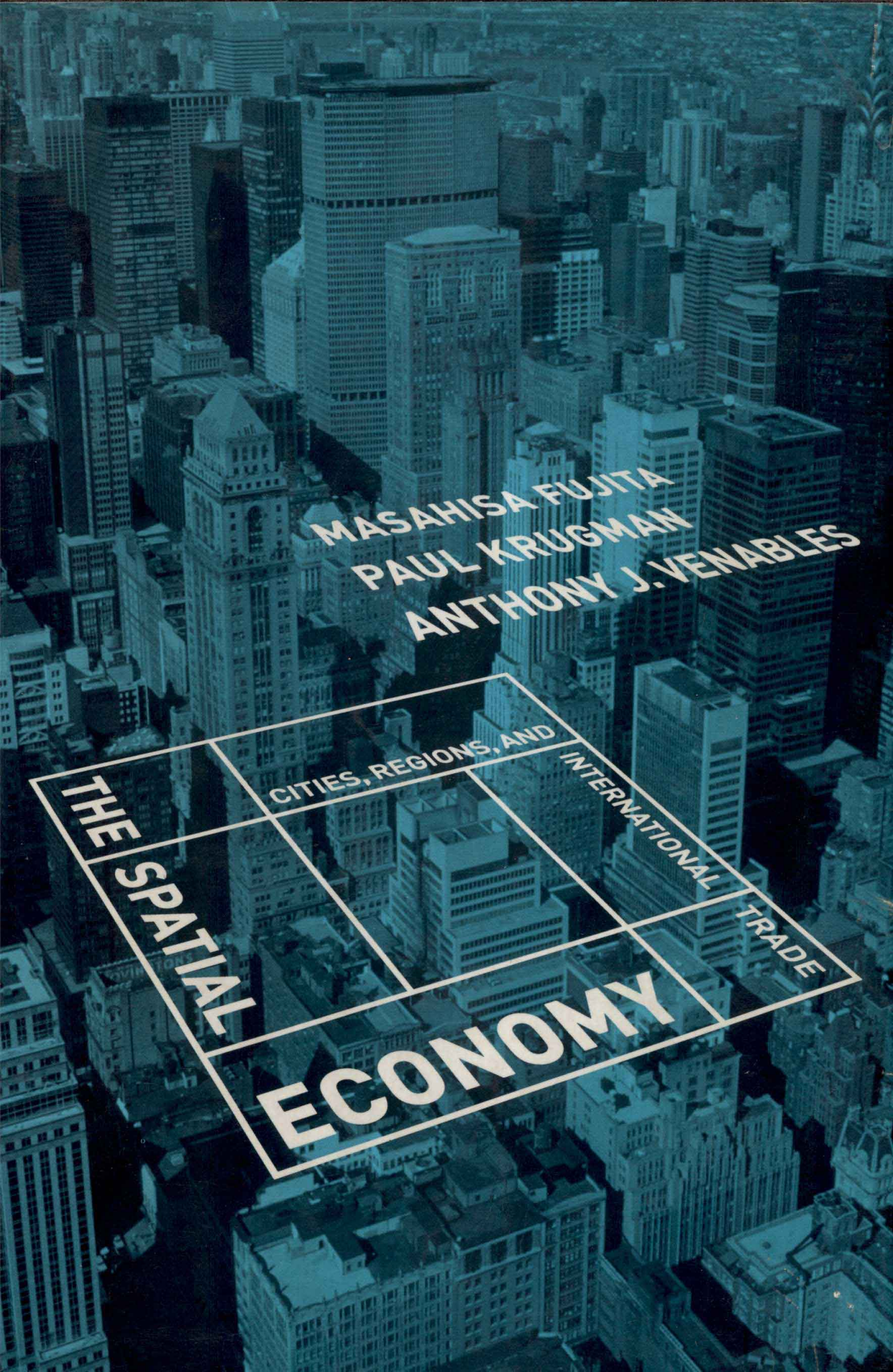
MASAHISA FUJITA
PAUL KRUGMAN
ANTHONY J. VENABLES
THE SPATIAL
CITIES, REGIONS, AND
INTERNATIONAL
TRADE
ECONOMY

The Spatial Economy

Cities, Regions, and International Trade

Masahisa Fujita,
Paul Krugman,
Anthony J. Venables

The MIT Press
Cambridge, Massachusetts
London, England

This book was set in Palatino by Achorn Graphic Services, Inc.

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Fujita, Masahisa.
The spatial economy : cities, regions, and international trade / M. Fujita, P. Krugman, A.J. Venables.
p. cm.
Includes bibliographical references and index.
ISBN 0-262-06204-6 (hc. : alk. paper)
1. Economic geography. 2. Space in economics. 3. Urban economics. 4. Regional economics. 5. International trade. I. Krugman, Paul R. II. Venables, Anthony. III. Title.
HF1025.F973 1999
330.9—dc21 99-10744
CIP

Preface

Mainstream economics has traditionally paid remarkably little attention to the location of economic activity—to the choices firms and households make about where to produce and consume, and about how these choices interact. The most recent edition of Mark Blaug's *Economic Theory in Retrospect* (1997) speaks of a "curious disdain of location theory on the part of mainstream economics," and asserts that "this neglect largely continues to this day."

But these remarks are, it turns out, a bit out of date. Since about 1990 there has been a renaissance—or perhaps simply a naissance, because the field has always been neglected—of theoretical and empirical work on the spatial aspects of the economy. Relying on new theoretical tools, this "new economic geography" has quickly emerged as one of the most exciting areas of contemporary economics.

Experience shows that a few years into such a new movement, it is often helpful if someone provides a synthesis—typically a book that shows how many seemingly disparate models can be viewed as variations on a few main themes, that develops a common "grammar" for discussing a range of issues. Such books as Helpman and Krugman 1985, on imperfect competition and international trade, or Grossman and Helpman 1991, on endogenous growth, helped give shape and direction to the new fields they surveyed. We believe that the time has come for a similar effort on the theory of economic geography. This book shows, in particular, how a common approach—one that emphasizes the three-way interaction among increasing returns, transportation costs, and the movement of productive factors—can be applied to a wide variety of issues in regional, urban, and international economics.

Not everyone will want or need to read all of the book. Here is a brief guide to its contents. Part I is essentially background material: a

review of the motivations for doing this kind of economic theory, and of some themes in earlier work that bear directly on our approach. The base-multiplier model of chapter 3 and the discussion of bifurcations in that chapter's appendix will probably prove useful as warm-up exercises for subsequent discussions. Part II then develops the basic approach, applying it to "regional" models, by which we mean models in which some factors of production are free to move among locations. Even for those whose principal interest is in either urban or international economics, chapters 4 and 5 are essential reading: The former sets out the market structure we use throughout the book; the latter, in the course of developing a basic core-periphery model, also develops a number of concepts and algebraic results that recur repeatedly. Chapters 6 and 7 are more optional (although each introduces concepts that are used in part IV; in particular, chapter 6 is a prerequisite for the similar discussion in chapter 17).

With these preliminaries under his or her belt, the reader has more options. The order of parts III and IV is arbitrary: You can proceed from regional directly to international economics rather than via urban economics, if you like. Within part III, the heuristic introduction in chapter 8 provides a road map to the material; from then on the development is sequential, except for the empirically motivated digression in chapter 12. In part IV, chapter 14 is essential background for the remaining chapters, but thereafter they can be taken on a stand-alone basis.

Some of this book is based on earlier publications by the authors, in some cases in collaboration with others. We would like to give particular mention to Tomoya Mori's role as a coauthor of the original papers on which much of chapters 10, 11, and 13 is based; to Diego Puga's role as coauthor of the basis paper for chapter 15; and to Raul Livas-Elizondo's corresponding role vis-à-vis chapter 18.

The book also benefited immensely from comments from many people. Portions of the manuscript have been used as the basis of courses at both the Massachusetts Institute of Technology and the London School of Economics, and students in these courses provided important input. Among those who have read draft versions of the manuscript and provided valuable suggestions are Jacques Thisse at the Center for Operations Research and Econometrics (CORE), J. Vernon Henderson at Brown University, Yannis Ioannides at Tufts University, Gianmarco Ottaviano at the University of Bologna, Martin Wagner at the Vienna

University of Technology, and Hiroyuki Koide at Nagoya Economics University.

Thanks also go to Hiroyuki Koide and Tomoya Mori for their excellent work in editing parts of this book and, for research support, to the U.K. Economic and Social Research Council–funded Centre for Economic Performance at the London School of Economics and to the British Taiwan Cultural Institute.

Contents

1 Introduction

1.1 The Rediscovery of Geography

Around the corner from the English National Opera lies St. Martin's Court, a short street occupied mainly by sellers of secondhand books and prints. It is a reasonable location for such shops, but there are no doubt other locations that would serve as well. Why, then, have the shops' owners chosen to be there? To be near each other. No doubt there is some interesting story about how that cluster of book and print shops originally became established, but what sustains it now is a sort of circular logic: Potential customers come to St. Martin's Court because they expect to find a range of shops to browse in, and shops locate there because they know they will have access to a large pool of potential customers.

The phenomenon that St. Martin's Court illustrates in microcosm pervades every economy. Agglomeration—the clustering of economic activity, created and sustained by some sort of circular logic—occurs at many levels, from the local shopping districts that serve surrounding residential areas within cities to specialized economic regions like Silicon Valley (or the City of London) that serve the world market as a whole. The distribution of population and activity across the landscape is radically uneven; in advanced countries the majority of the population lives in large metropolitan areas, and these metropolises are themselves clustered into regions like the Boston-Washington corridor. Yet although agglomeration is clearly a powerful force, it is not all-powerful: London is big, but most Britons live elsewhere, in a system of cities with widely varying sizes and roles.

It should not, in other words, be hard to convince economists that economic geography—the study of where economic activity takes place and why—is both an interesting and an important subject. Yet

until a few years ago it was a subject mainstream economics largely neglected. Even now, introductory textbooks seem to describe a curiously disembodied economy, without cities or regions. (Most such texts, indeed, make literally no mention at all of such questions as the reasons for urbanization or the role of location in economic decisions).

In the last few years, however, research on economic geography—that is, on where economic activity occurs and why—has increased dramatically. Real-world concerns have, to some extent, driven this surge of interest: The field has been given a big boost in particular by plans to unify the European market and the attempt to understand how this deeper integration will work by comparing international economics within Europe with interregional economics within the United States. But economic geography has always been important; if the economics profession has notably neglected it, this is not because economists have been uninterested in the subject, but because they have regarded it as intractable. Their new willingness to work on economic geography comes from their sense that new tools—in particular, modeling tricks that have been developed to analyze industrial organization, international trade, and economic growth—have removed crucial technical barriers and transformed a once inhospitable field into fertile ground for theorists.

The basic problem with doing theoretical work in economic geography has always been that any sensible story about regional and urban development hinges crucially on the role of increasing returns. Suppose that we really lived in the constant-returns world that much economic theory still assumes. Then it would be hard to understand why the economy is not characterized by "backyard capitalism," in which each household or small group produces most items for itself. There would, admittedly, be some unevenness in population density and some trade among locations because of the variation in the natural environment: Land differs in fertility, and differences in soil, climate, and resources mean that no one locality would produce all goods even under constant returns. Nonetheless, the dramatic spatial unevenness of the real economy—the disparities between densely populated manufacturing belts and thinly populated farm belts, between congested cities and desolate rural areas; the spectacular concentration of particular industries in Silicon Valleys and Hollywoods—is surely the result not of inherent differences among locations but of some set of cumulative processes, necessarily involving some form of increasing returns, whereby geographic concentration can be self-reinforcing.

Unfortunately, increasing returns have always posed difficulties for economic theorists. Except under very special circumstances they lead to a breakdown of perfect competition; even if this problem can somehow be finessed, they pose problems for the existence and uniqueness of equilibria. For the theorist determined to make some headway in understanding the location of economic activity, these difficulties have not been insurmountable. For example, one can, like much of urban economics, simply take the existence of cities (or central business districts within cities) as a given and trace out the consequences for land rents and land use; this is the basis of the famous von Thünen model, which has given rise to a rich and productive literature. Or one can, like urban systems theorists (above all Henderson (1974, 1980, 1988)), represent increasing returns in a somewhat black-box way as localized production externalities; this approach sidesteps some important questions but opens the door to a powerfully insightful analysis of others. Still, until a few years ago these efforts remained peripheral to the main body of economic theory.

In the last few years, however, a "new economic geography" has emerged, the fourth wave of the increasing-returns revolution in economics. The revolution began in the 1970s in the field of industrial organization, when theorists began for the first time to develop tractable models of competition in the presence of increasing returns; in particular, Dixit and Stiglitz (1977) developed a formalization of Chamberlin's concept of monopolistic competition that, though admittedly a very special case, has turned into the workhorse of theoretical modeling in a number of fields. Beginning at the end of the 1970s, a number of theorists applied the analytical tools of the new industrial organization theory to international trade; a few years later the same tools were applied to technological change and economic growth. In each case it was, of course, necessary to do much more than mechanically apply the Dixit-Stiglitz model to the subject at hand: New concepts needed to be developed, and at first seemingly inconsistent models and approaches proliferated, in which each author appeared to be inventing his or her own private language and notation. In time, however, it became clear in each case that a core set of useful insights had emerged; indeed, in retrospect it is remarkable how tightly integrated, how classical in feel, both the "new trade" and "new growth" theory have turned out to be.

Our sense is that the state of the "new economic geography" is currently similar to that of the new trade theory circa 1984, or the new

growth theory circa 1990. That is, an exuberant and initially exhilarating growth of theory has reached the point at which it has become difficult to see the forest for the trees; and yet there is, if one looks for it, a strong element of commonality among many if not all of the analyses. The integration of new trade and new growth theory was, we believe, powerfully aided by the appearance of judiciously timed monographs that endeavored to synthesize each field into a coherent whole: Helpman and Krugman's *Market Structure and Foreign Trade* (1985) and Grossman and Helpman's *Innovation and Growth in the World Economy* (1991). This book is, of course, an effort to do the same with the new economic geography.

In the remainder of this chapter, we describe what we regard as the unifying themes, methods, and questions of this new field and set out the plan of the book.

1.2 Linkages and Circular Causation

We would argue that the defining issue of economic geography is the need to explain concentrations of population and of economic activity: the distinction between manufacturing belt and farm belt, the existence of cities, the role of industry clusters. Broadly speaking, all these concentrations form and survive because of some form of agglomeration economies, in which spatial concentration itself creates the favorable economic environment that supports further or continued concentration. And for some purposes, as in the urban systems literature described in chapter 2, it may be enough simply to posit the existence of such agglomeration economies. But the main thrust of the new geography literature has been to get inside that particular black box and derive the self-reinforcing character of spatial concentration from more fundamental considerations. The point is not just that positing agglomeration economies seems a bit like assuming one's conclusion; as a sarcastic physicist remarked after hearing one presentation on increasing returns, "So you're telling us that agglomerations form because of agglomeration economies." The larger point is that by modeling the sources of increasing returns to spatial concentration, we can learn something about how and when these returns may change, and then explore how the economy's behavior changes with them.

How should the returns to spatial concentration be modeled? More than a century ago Alfred Marshall suggested a threefold classification (1920, p. 271). In modern terminology, he argued that industrial dis-

tricts arise because of knowledge spillovers ("the mysteries of the trade become no mysteries; but are as it were in the air"), the advantages of thick markets for specialized skills, and the backward and forward linkages associated with large local markets. Although all three of Marshall's forces are clearly operating in the real world, the new geography models have generally downplayed the first two, essentially because they remain hard to model in any explicit way. Instead, they have focused on the role of linkages.

The linkage story is easy to tell if one is willing to be a bit vague about the details. Producers, so the story goes, want to choose locations that have good access to large markets and to supplies of goods that they or their workers require. However, a place that for whatever reason already has a concentration of producers tends to offer a large market (because of the demand the producers and their workers generate) and a good supply of inputs and consumer goods (made by the producers already there). These two advantages correspond precisely to the backward linkages and forward linkages of development theory. Because of these linkages, a spatial concentration of production, once established, may tend to persist, and a small difference in the initial economic size of two otherwise equivalent locations may grow over time.

Discussions of linkage-based spatial concentration that embody more or less this story have been familiar to regional scientists for many years. In chapter 3, we describe in particular two such stories: the dynamic extension of the base-multiplier approach largely identified with Pred (1966) and the widely used concept of market potential associated with such authors as Harris (1954). And provided that one is prepared to be strategically sloppy about details, it is possible to jump straight from such stories into heuristic models that are quite useful both for quick and dirty discussions of real-world issues and as guides to the results of more careful modeling. Such loose-jointed modeling is, we believe, underappreciated in economics; we try to give it its due.

Nonetheless, traditional discussions of linkages and economic geography do not address certain questions that nevertheless become crucial once one tries to get beyond the simplest stories. Most important of these is the nature of competition. Linkage stories work only if there are increasing returns to production at the level of the individual firm; otherwise, the firm would not concentrate production where the market is largest, but rather establish a separate facility to serve each market. But if there are increasing returns, competition must be imperfect;