ABSTRACT

Title of Dissertation: THREE ESSAYS ON NON-BALANCED ECONOMIC GROWTH, ECONOMIC GEOGRAPHY, AND THE REGULATION OF PUBLIC LAND IN THE UNITED STATES

Michael H. Taylor, Doctor of Philosophy, 2014

Dissertation directed by: Professor Robert G. Chambers
Department of Agricultural and Resource Economics

This dissertation is comprised of three essays. The first essay, "The Spatial Consequences of Non-Balanced Growth," develops a theoretical model of regional development to analyze the consequences of non-balanced growth for the spatial distribution of population and production. The essay focuses on the spatial consequences of deindustrialization, where deindustrialization refers to the increase in the service sector's share of total employment over time. This essay demonstrates that the presence of non-balanced growth at the national level has implications for both regional population movements and the long-run distribution of population between regions, as well as for several of the relationships between regional economic activity and regional population growth that are emphasized in the previous literature.

The second essay, "Non-Balanced Growth in the United States: Evaluating Supply-Side versus Demand-Side Explanations," develops and calibrates a dynamic general equilibrium model that integrates supply-side and demand-side explanations for NBG,

and uses it to evaluate the extent to which the two explanations for NBG can account for patterns of NBG consistent with the "Kuznets" and "Kaldor" facts in the post-war United States. This essay demonstrates that it is necessary to consider both explanations for non-balanced growth to generate patterns of sectoral output and employment growth that are qualitatively consistent with the Kuznets facts as they occurred in the United States over the study period. In addition, the model generates equilibrium dynamics that are broadly consistent with the Kaldor facts for a wide range of different parameterizations.

The third essay, "Regulatory Policy Design for Agroecosystem Management on Public Rangelands," analyzes regulatory design for agroecosystem management on public rangeland. It presents an informational and institutional environment where three of the most prominent regulatory instruments on public rangelands – input regulation, cost-sharing/taxation, and performance regulation – can be defined and compared. The essay examines how the optimal regulation is shaped by informational asymmetries between ranchers and regulators within federal land management agencies, limitations on the ability of regulators to monitor ranch-level ecological conditions, and constraints on regulators' actions due to budget limitations and restrictions on the level of penalties they can assess.

THREE ESSAYS ON NON-BALANCED ECONOMIC GROWTH, ECONOMIC GEOGRAPHY, AND THE REGULATION OF PUBLIC LAND IN THE UNITED STATES

by

Michael H. Taylor

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:

Professor Robert G. Chambers, Chairman/Advisor
Professor Lint Barrage
Professor Tigran Melkonyan
Professor Lars J. Olson
Professor John Shea

Preface

The third essay in this dissertation, "Regulatory Policy Design for Agroecosystem Management on Public Rangelands," is co-authored with Tigran Melkonyan and has been published in the *American Journal of Agricultural Economics* (Volume 95, Issue 3; pages 606-627). It is the opinion of the Dissertation Committee that Michael H. Taylor made a substantial contribution to the relevant aspects of this essay.

DEDICATION

To Lydia

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of several people. First off, of course, I would like to thank my advisor, Prof. Robert G. Chambers, for his guidance and patience throughout the writing of this dissertation. His example of hard work and professionalism and his passion for economics give me something to aspire to in my professional career.

In addition to Prof. Chambers, I have benefited from the help and encouragement of several other professors at the Department of Agricultural and Resource Economics. To begin, I would like to thank Prof. Andreas Lange, Prof. Ramón E. Łópez, and Prof. Carol McAusland for their thoughtful comments on earlier drafts of this dissertation. I would also like thank Prof. John A. List and Prof. Lars J. Olson for giving me the opportunity to learn from them through collaboration.

In addition to professors, there are many other staff members at the Department of Agricultural and Resource Economics whose help I would like to acknowledge. To begin, I would like to thank the entire front office staff - Barbara Burdick, Jane Doyle, Chris Aggour, Liesl Koch, Peggy Gazelle, Linda Wilkinson, and Curtis Henry - for their help throughout my time at Maryland. I would also like to thank Jeff Cunningham and Chuck McCormick for having the patience to put up with all of my inane computer and software questions, and Katherine Faulkner for her help and friendship.

I would also like to acknowledge the tremendous role that friends and family have played in my life over my time at Maryland. First and foremost, I would like to thank my parents, David and Eithne, for their love and support. I would also like to thank my fellow graduate students, whose friendship and camaraderie have made my time at Maryland enjoyable.

Last, but certainly not least, I would like to acknowledge the tremendous role that my wife, Lydia, has played in making this dissertation possible. Words cannot describe how lucky I am to have her in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

The Spatial Consequences of Non-Balanced Growth

1.1 Introduction

This essay develops a theoretical model of regional development in the presence of *non-balanced growth* (NBG) and uses it to analyze the consequences of NBG at the national level for the spatial distribution of population and production. The term NBG refers to systematic changes in the employment and output shares of the major sectors of the economy over time in the process of economic development (see Kongsamut, Rebelo, and Xie 2001; Acemoglu and Guerrieri 2008). This essay focuses on the spatial consequences of *deindustrialization* – the increase in the service sector's share of total employment over time relative to goods producing sectors – which is the most prominent form of NBG in advanced economies. The analysis considers the three explanations for NBG most emphasized in the literature: *supply-side* NBG, where NBG results from differential rates of productivity growth across sectors; *demand-side* NBG, where NBG results from differences in the income elasticities of demand for sectoral output; and NBG in an open economy, where change in the terms of trade between countries results in NBG.[1] As these alternative explanations for NBG are neither mutually exclusive nor contradictory, all three causal explanations for NBG may be simultaneously influencing the spatial development of an economy undergoing NBG.

It is shown in this essay that the presence of NBG on the national level influences how the spatial distributions of population and production change as a consequence of economic growth. How the spatial distribution of the economy evolves as a consequence of economic growth (and how the spatial distribution of the economy influences the process of economic growth) is one of the central questions in economic geography (Gabaix 1999; Duranton 2007). The theoretical analysis suggests a simple mechanism for how NBG may influence the spatial distribution of the economy: when there are differences in the relative importance of *regional economic characteristics* between the major sectors of the economy, NBG will cause population to shift towards regions whose economic characteristics give them a comparative advantage in the sector(s) whose share of total employment is expanding. In this essay, the term regional economic characteristics refers both to regions' innate geographic characteristics – or *first-nature features* – and to regions' endogenous economic characteristics – or *second-nature features*.

In addition to influencing the spatial distribution of the economy, NBG is shown to have implications for how a region's economic characteristics influence its rate of population growth. The relationship between a region's economic characteristics and its population growth rate is another major question in economic geography (e.g., Glaeser, Scheinkman, and Shleifer 1995). It is shown that when NBG influences regional population movements, the presence of NBG can undermine two important

[1] Supply-side non-balanced growth is also referred to as *technology-based* non-balanced growth. Demand-side non-balanced growth is also referred to *preference-based* non-balanced growth.

predictions of previous models of urban and regional development concerning the relationship between a region's economic characteristics and its rate of population growth. First, NBG can disrupt the self-reinforcing relationship between the spatial agglomeration of economic activity and regional population growth emphasized in previous models (e.g., Krugman 1991; Venables 1996; Baldwin and Forslid 2000). In particular, it is shown that when economic growth is non-balanced, it can lead to a reduction in the spatial agglomeration of economic activity at both the regional level, leading to (relative) population decline in agglomerated regions, and the national level, leading to a more even distribution of population across regions in the economy. Moreover, this result holds in a theoretical model where *agglomeration externalities* benefit both regional productivity and regional productivity growth. Agglomeration externalities refer to the advantages that firms and workers receive from locating close to one another spatially that are external to the firm or worker, but internal to the region in which the firm or worker operates.[2]

Second, it is shown that the presence of NBG on the national level can attenuate the positive relationship between regional productivity growth and regional population growth emphasized in previous models of urban and regional development. (The term productivity growth in this essay refers to total factor productivity growth unless otherwise specified.) In many previous models of urban and regional development, such as Glaeser, Scheinkman, and Shleifer (1995) and Palivos and Wang (1996), above-average productivity growth in a region increases the return to regional factors of production, including labor, and encourages in-migration. The analysis in this

[2]Duranton and Puga (2004) survey the theoretical literature on agglomeration externalities. Rosenthal and Strange (2004) survey this literature's empirical counterpart. Berliant and Ping (2004) and Henderson (2006) survey the literature that examines the relationship between agglomeration externalities and urban and regional economic and population growth. In addition to the term agglomeration externalities, the terms *agglomeration economies* and *localized production externalities* are also used.

essay demonstrates that in certain circumstances, NBG can cause this positive relationship between regional productivity and population growth to be reversed, with regions concentrated in relatively stagnant sectors experiencing stronger population growth. This result is in agreement with recent patterns of regional development in the United States and other advanced economies, where many of the regions that have experienced the strongest population growth have been concentrated in relatively stagnant sectors of the economy (Glaeser and Tobio 2007).

The spatial consequences of deindustrialization have received little formal theoretical analysis in the previous literature. Indeed, theoretical studies that have jointly analyzed location and economic growth, such as Baldwin and Forslid (2000) and Fujita and Thisse (2003), have focused on the relationship between economic growth and the spatial distribution of the economy, and do not consider the influence of changes in sectoral employment shares over the process of development. The lack of formal analysis of the spatial consequences of deindustrialization is in marked contrast to the literature on urban-rural transformation, where numerous studies have examined the spatial consequences of *industrialization* – the movement of workers from the agricultural sector to non-agricultural sectors (Becker, Mills, and Williamson 1986; Davis and Henderson 2003). This lack of attention is surprising given that deindustrialization has been observed in almost every advanced economy – in the United States since the mid-1960s and in Western Europe and Japan since the early 1970s – and that the shift in the sectoral composition of employment towards services in many of these economies has been profound. Consider the recent transformation of the U.S. economy: in 1957, the service sector comprised 57% of total employment; by 2000, it comprised 75% (Lee and Wolpin 2006). Further, Rowthorn and Ramaswamy (1999) find, when looking at changes in the service sector employment relative to manufacturing, that between 1960 and 1994 the total number of workers engaged in manufacturing across advanced economies as a whole remained roughly constant, while employment in the service sector grew at an average annual rate of 2.2%.

The magnitude of theses changes in the sectoral composition of employment suggests that deindustrialization is likely to have consequences for many aspects of the economy, including the spatial distribution of population and production. Indeed, several empirical studies of regional development in advanced economies have found that deindustrialization is concurrent with the population decline of regions whose economies are relatively less concentrated in services.[3] For example, Glaeser, Scheinkman, and Shleifer (1995), when considering the growth of cities in the United States between 1960 and 1990, and Combes (2000), when considering the growth of French employment regions between 1984 and 1993, find that regions significantly involved in manufacturing grew more slowly in terms of both population and per capita income than regions less involved in manufacturing.

This essay's focus on deindustrialization is reflected in the importance given to endogenous regional economic characteristics in the analysis. Previous models of urban-rural transformation, such as Davis and Henderson (2003), have exogenously assumed that certain regions are more suitable for agricultural production (e.g., rural regions) and certain regions are more suitable for non-agricultural production (e.g., urban regions). In the case of deindustrialization, however, it is less defensible to assume that there are exogenous differences in the suitability of regions for production in services relative to goods production. The assumption of exogenous differences is less defensible because both the service and goods-producing sectors are largely ur-

[3] It is only the subset of industries in the manufacturing and service sectors that generate export revenue for a region whose location decisions will influence the spatial distribution of the economy. Black and Henderson (2003) find that 65% of employment – some manufacturing (wholesale trade, construction, etc.) and some service (retail, restaurants, etc.) – is relatively constant across regions in the United States. It is industries that comprise the remaining 35% of employment that account for regional heterogeneity in industrial composition and whose location decisions influence the distribution of economic activity across regions.

ban. Both sectors being largely urban is important because the benefits from urban production are thought to be related to the spatial concentration of workers and firms in urban areas – i.e., to agglomeration externalities – which is endogenous. Because of the importance of agglomeration externalities, endogenous regional economic characteristics must be explicitly taken into account in the model presented in this essay in order to make the analysis relevant to understanding the spatial consequences of deindustrialization.

As is mentioned above, the term NBG in this essay refers to systematic changes in the employment and output shares of the major sectors of the economy in the process of economic development. The analysis in this essay focuses by and large on NBG as it applies to changes in sectoral employment shares because it is changes in employment shares that are relevant for regional population movements. Certain studies have used the term "structural change" to refer to changes in sectoral employment shares that occur over long periods of time (e.g., Ngai and Pissarides 2007). This essay uses the term NBG rather than structural change because, as has been pointed out in Matsuyama (2007), the term structural change is often used more broadly to refer to changes in all aspects of the economy brought about by economic growth and the accompanying increases in per capita income.[4]

The term deindustrialization in this essay refers to the increase in the service sector's share of total employment over time relative to a broadly defined goods-producing sector, which includes both manufacturing and agriculture.[5] This essay defines deindustrialization in this way because its intent is to explore the spatial con-

[4] According to Matsuyama (2007), structural change includes changes in the sectoral composition of output and employment, changes in the distribution of income and wealth, changes in demographics (population age distribution, etc.), and changes in institutions, such as the financial system, the organization of industry, and political institutions.

[5] Note that together, services, manufacturing, and agriculture comprise all of private-sector employment.

sequences of NBG where the economic forces driving NBG are those associated with the rise in the service sector's share of total employment in advanced economies. The rise in the service sector's share of total employment in the later stages of development is pervasive across advanced economies, and the primary explanations offered for this rise – slower productivity growth in services (e.g., supply-side NBG) and higher income elasticities of demand for service sector output (e.g., demand-side NBG) – apply to the increase in the service sector's share of total employment relative to both manufacturing and agriculture. This use of the term deindustrialization is different from that of several previous studies, which use deindustrialization to refer to the decline in the manufacturing sector's shares of total employment (see Eltis 1996; Rowthorn and Ramaswamy 1999; Pitelis and Antonakis 2003).[6]

In the theoretical model, a region is defined by the geographic extent of its labor market. In particular, it is assumed that the geographic range of workers is limited so that the physically immobile factors of production (e.g., land) in a region can only be used by workers located in the region. It is also assumed that agglomeration externalities are limited to workers and firms in a given region. This latter assumption agrees with previous empirical research that suggests that agglomeration externalities have limited geographic reach (Jaffe, Trajtenberg, and Henderson 1993; Ellison and Glaeser 1997; Wallsten 2001).

On a formal level, the theoretical model developed in this essay combines a Romer (1990)-type model of endogenous growth with horizontally differentiated inputs and a core-periphery model of economic geography. The model extends previous multi-sector, multi-region, general-equilibrium models of regional development by allowing endogenous labor allocation across both sectors and regions within an analytically

[6] These previous studies attempt to explain the relative decline of the manufacturing sector in specific countries (e.g., the United Kingdom in Eltis (1996) and Greece in Pitelis and Antonakis (2003)) or across advanced countries as a whole (e.g., Rowthorn and Ramaswamy, 1999).

tractable model of endogenous growth. This extension is essential to analyzing both NBG (intersectoral labor mobility) and its spatial consequences (interregional labor mobility). The theoretical framework builds off previous models, such as Baldwin and Forslid (2000) and Fujita and Thisse (2003), that have successfully incorporated agglomeration externalities into a tractable, multi-region growth model with factor mobility.

The remainder of this essay is structured as follows. Chapter 1.2 surveys the literature on NBG and the literature in economic geography on the determinants of urban and regional growth. Chapter 1.3 discusses the elements that the theoretical model must contain in order to analyze the implications of NBG for regional development. Chapter 1.4 develops the baseline theoretical model of regional development in the presence of supply-side NBG. The baseline model considers supply-side NBG because it has been suggested that it is the explanation for NBG that most closely matches the experience of the United States and other advanced economies over the period of deindustrialization (Ngai and Pissarides 2007; Acemoglu and Guerrieri 2008). In Chapters 1.5 and 1.6, the static and dynamic equilibria for the baseline model of supply-side NBG are defined and analyzed. In Chapter 1.7.1, the baseline model is reconsidered for demand-side NBG; in Chapter 1.7.2, the baseline model is further extended to consider the spatial consequences of NBG in an open economy. Finally, in Chapter 1.8, conclusions are given and directions for future research are discussed.

1.2 Literature Review

This section reviews the literature on NBG and the literature in economic geography on the determinants of urban and regional growth and explains how concepts from these literatures underlie the main economic forces in the analysis.

1.2.1 Non-Balanced Growth

This section reviews the literature on NBG. As stated in the introduction, NBG refers to systematic change in the relative importance of the major sectors in the economy, in terms of output and employment shares, in the process of economic development. In particular, NBG refers to what Kongsamut, Rebelo, and Xie (2001) refer to as the "Kuznets facts" of economic development: the decline in the relative importance of the agricultural sector in the early stages of development, and the successive rise in the relative importance of the manufacturing and service sectors.

The literature on NBG can usefully be divided into two branches. The first branch consists of descriptive studies that seek to establish stylized facts concerning the patterns of NBG followed by most countries. Without exception, studies from this first branch of the literature provide empirical support for the Kuznets facts. For example, Maddison (1980) presents evidence on NBG from 16 countries from 1870 to 1976.[7] Over this period, the average portion of employment in agriculture across countries fell from 49% to 8%, while the portion of employment in manufacturing rose from 28% to 36% and the portion of employment in services rose from 24% to 56%. Similarly, Kongsamut, Rebelo, and Xie (1997) present evidence that through the period from 1970 to 1989 patterns of NBG consistent with the Kuznets facts have occurred in a cross-section of 123 non-socialist countries, where the agricultural sectors comprise the largest share of output and employment in poorer countries, and the service sectors comprise the largest share of output and employment in richer countries. Kongsamut, Rebelo, and Xie (1997) also find evidence consistent with the Kuznets facts in time-series data from 22 countries from 1970 to 1989, in which the relative importance of the agricultural sector declines and that of the service sector

[7] The 16 countries in Maddison (1980) are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, Switzerland, the United Kingdom, and the United States.

increases as countries become richer.

The second branch of the literature consists of analytical studies that use formal theoretical models to analyze how NBG is systematically related to the process of economic growth.[8] This second branch of the literature can be further divided into studies that emphasize supply-side explanations and studies that emphasize demand-side explanations for NBG. Studies that emphasize supply-side explanations for NBG posit that NBG occurs primarily as the result of differential rates of productivity growth across sectors. To explain deindustrialization, these studies contend that productivity growth has proceeded more slowly in services and that the elasticity of substitution between service-sector output and output from goods-producing sectors is low (less than 1).[9] When the elasticity of substitution is low, the increase in the relative price of service-sector output that results from slower sectoral productivity growth more than compensates for the decrease in the relative return to factors of

[8] There are also studies that emphasize the expansion of North-South trade in recent decades in their explanations of non-balanced growth and deindustrialization (Sacks and Schatz 1994; Wood 1995). According to these studies, deindustrialization has occurred in part because the number of low-skilled jobs lost in import-competing industries in advanced countries as a result of the expansion of North-South trade has exceeded the number of skill-intensive jobs created in the export sector. Rowthorn and Ramaswamy (1999) have found that, in a sample of advanced economies that have experienced deindustrialization, North-South trade plays a comparatively smaller role in explaining deindustrialization than do domestic factors such as slower productivity growth in services.

[9] While there is not information available in the literature on the elasticity of substitution between the service sector and the broadly-defined goods-producing sector considered in this essay, several previous studies have found an elasticity of substitution between service and non-service sectors of less than 1 using different definitions of the two sectors. For example, Acemoglu and Guerrieri (2008) estimate an elasticity of substitution between output from the more capital-intensive sector (which corresponds closely to the goods-producing sector) and the less capital-intensive sector (which corresponds closely to the service sector) of 0.76 with a two standard error confidence interval of 0.73 to 0.79.

production in the service sector, causing the sector's share of total employment to increase as the economy evolves. Differences in productivity growth between sectors drives NBG in the models of Baumol (1967), Ngai and Pissarides (2007), and Acemoglu and Guerrieri (2008).

Previous studies have put forward a variety of explanations for why productivity growth might proceed faster in goods production than in services. Baumol (1967) attributes differential rates of productivity growth across sectors to differences in the inherent technological character of different activities in the economy. In support of this claim, Baumol points out that productivity in many service-sector activities, such as education or health care, are unlikely to benefit greatly from technological advances or from scale economies. Alternatively, Acemoglu and Guerrieri (2008) identify capital deepening – increases in the capital-to-labor ratio as the economy evolves – as a potential source of differential productivity growth across industries, because capital-intensive industries derive greater benefit from capital deepening in the economy. The authors argue that sorting industries by capital intensity maps closely, though imperfectly, to the service and goods-producing sectors, which are less capital-intensive and more capital-intensive, respectively.

The empirical evidence suggests that, on average, productivity growth has proceeded more slowly in services. Maddison (1980) reports figures on output growth per worker (labor productivity growth) for 4 countries for the period 1870 to 1950 and for 12 countries for the period 1950 to 1976.[10] In all cases, Maddison finds that labor productivity has proceeded most rapidly in agriculture, followed by manufacturing, followed, in turn, by services. Baumol, Blackman, and Wolff (1985) also present evidence consistent with slower labor productivity growth in the service sector using

[10]The four countries for the period 1870 to 1950 are Germany, Italy, Japan, and the United Kingdom. The 12 countries for the period 1950 to 1976 are Austria, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, the United Kingdom, and the United States.

U.S. data from 1947 to 1976 at the two-digit industry level according to the *Standard Industrial Classifications.* Jorgenson and Stiroh (2000) provide evidence that both total factor productivity growth and labor productivity growth have proceeded more slowly in services than in goods production in the United States in the period from 1958 to 1996. In particular, Jorgenson and Stiroh find that the two fastest growing subsectors of the service sector in terms of total employment – *Finance, Insurance, and Real Estate* and *General Services* – both experienced negative annual total factor productivity growth (-0.176 and -0.190, respectively) and weak annual labor productivity growth (0.664 and 0.920, respectively) over the period.[11] More recently, Triplett and Bosworth (2003) have shown that the productivity resurgence beginning in the mid-1990s can be attributed in large measure to productivity growth in the service sector catching up to, though not surpassing, productivity growth in manufacturing.

Studies that emphasize demand-side explanations for NBG argue that differences in the income elasticities of demand for sectoral output caused by non-homothetic consumer preferences are the primary drivers of NBG. According to these studies, differences in the income elasticities of demand for sectoral output cause sectoral output and employment shares to change as the economy becomes wealthier on a per capita basis. Non-homothetic consumer preferences drive NBG in the models of Murphy, Shleifer, and Vishny (1989), Caselli and Coleman (2001), and Gollin, Parente, and Rogerson (2002).

The empirical evidence suggests that services are slightly income elastic (i.e., have an income elasticity of demand slightly greater than 1), though this evidence is mixed (see Kravis, Heston, and Summers (1983) and Flavey and Gemmell (1996)). While there are numerous studies that have documented the rise in the share of total expenditure devoted to services as per capita incomes rise (e.g., Houthakker and Taylor

[11] The subsectors Finance, Insurance, and Real Estate and General Services are defined according to the Standard Industrial Classifications.

1970; Kravis, Heston, and Summers 1983), the evidence that real expenditure on services increases with per capita expenditure is less strong.[12] For example, Kravis, Heston, and Summers (1983), using time series data from France (1959 to 1978), the United Kingdom (1957 to 1978), and the United States (1947 to 1978), show that while the share of services in total expenditure rises sharply as incomes rise when current-period prices are used, when services are measured in constant prices, the share of expenditure devoted to services is stable over time for France and the United Kingdom and only rises slightly for the United States.[13] The evidence from empirical studies that have directly estimated the aggregate income elasticity of demand for services is similarly mixed. For example, Bergstrand (1991), using a cross-section of 21 countries in 1975, and Flavey and Gemmell (1991), using a cross-section of 52 countries from 1980, find that the aggregate income elasticity of demand for services is slightly greater than 1. Alternatively, using cross-sectional data on consumer expenditure for 11 service-sector industries for 60 countries in 1980, Flavey and Gemmell (1996) find that while certain services are income-elastic, the aggregate income elasticity of demand for services is not significantly different from 1.

The majority of the literature on NBG focuses on explaining NBG in a closed econ-

[12] Houthakker and Taylor (1970) break household consumption in the United States into durable goods, non-durable goods, and services. Using this taxonomy, they show that over the period 1954 to 1970 the share of durable goods was roughly constant, while the share of services has been increasing at the expense of non-durables. Kravis, Heston, and Summers (1983) present evidence that the share of total expenditure devoted to services is increasing in per capita income from a cross-section of 34 countries in 1975 using 151 expenditure categories, where the service sector comprises 22% of household spending in developing economies and 35% to 45% of household spending in advanced economies.

[13] Ngai and Pissarides (2007) cite the increase in the relative prices of services that has accompanied deindustrialization as evidence supporting their claim that slower productivity growth in services is the primary driver of deindustrialization.

omy. Limiting the analysis to a closed economy is problematic because the advanced countries in which deindustrialization has taken place are all, to varying degrees, open economies, and because international trade will cause the process of NBG to unfold differently across countries. As has been pointed out by Matsuyama (2007), faster productivity growth in a given sector will shift a country's comparative advantage towards that sector. So, while faster productivity growth in goods production within a country may lead to deindustrialization globally, it may not cause employment in goods-producing sectors in that country to decline. This provides an explanation for why deindustrialization has been uneven across advanced economies. For example, Obstfeld and Rogoff (1996), when considering declines in manufacturing employment, show that many advanced economies, such as Germany and Japan, have experienced smaller declines in manufacturing employment than the United States, and that certain advanced Pacific Rim economies, such as Hong Kong, Taiwan, and South Korea, have seen their manufacturing sectors' share of employment continue to rise through the early 1990s.

1.2.2 The Determinants of Regional Growth

This section surveys the economic geography literature on the determinants of urban and regional growth, paying particular attention to concepts that underlie the explanation for how NBG influences spatial change in the theoretical model.

In the economic geography literature, the extent of economic development and the pace of economic growth in a given region is determined by its economic characteristics, or its first- and second-nature features (Krugman 1993). In this section, the literature in economic geography on the determinants of regional growth is divided into studies that emphasize the role of first-nature features and studies that emphasize the role of second-nature features in regional economic development. A region's first-nature features relate to its intrinsic geographic characteristics and are independent of previous economic development in the region. First-nature features are what led economic activity to become established in its current locations. Examples of first-nature features are a region's proximity to the ocean and other navigable bodies of water, a region's climate, and a region's proximity to deposits of natural resources. The term first-natures also denotes features, such as the amount of available land, that limit the extent of economic development in a region and encourage the dispersion of economic activity.

A region's second-nature features are dependent on previous economic development in the region. Second-nature features include a region's physical infrastructure, such as roadways and housing, which give more developed regions a competitive advantage over less developed regions, as well as the benefits that firms and workers receive from locating close to one another spatially (agglomeration externalities). Agglomeration externalities are the second-nature features most emphasized in the economic geography literature and are what give rise (along with indivisible public goods and indivisible capital) to increasing returns to scale on the regional level – the observed positive relationship between a region's size and the productivity of its workers and firms. Agglomeration externalities encompass all positive externalities, both pecuniary (such as the benefits of large labor markets) and technical (such as knowledge spillovers), that are external to the worker or firm but internal to the region in which the worker or firm operates. Agglomeration externalities are both static and dynamic – benefiting both current regional productivity and regional productivity growth.

Among the sources of agglomeration externalities emphasized in the literature, three factors are most prominent. First, there are the advantages of regional size that arise from labor-market pooling. Labor-market pooling emphasizes the benefits that follow from the better matches, on average, between heterogeneous workers and firms that occur in regions with large populations and large labor markets. Better matching allows workers to be more productive in the jobs they perform, which attracts firms to large regions, and allows firms located in large regions to pay higher wages, which in turn attracts workers.[14] Second, there are the advantages that firms located in large regions receive from having access to a larger variety of local suppliers of specialized intermediate inputs.[15] Third, there are the benefits workers and firms receive from exchanging industry-relevant knowledge with other workers and firms, often termed knowledge spillovers. The idea is that because the exchange of knowledge is facilitated by geographic proximity, workers and firms located in large regions have a greater ability to access and benefit from the knowledge of others.

In addition to agglomeration externalities, which are related to the economic conditions within a region, there are second-nature features that relate to the spatial structure of the economy as a whole. The most notable inclusion in this category is market potential. A region is said to have good market potential if, in addition to having a large domestic market, it is also geographically close to other regions with large markets. Good access to markets with large numbers of potential customers is thought to promote regional growth by encouraging export industries. Market potential has been shown to be an important factor in explaining why manufacturing production in the United States was initially concentrated in the Northeast and the Midwest (Harris 1954). There are also second-nature features that discourage

[14]This line of research was first developed by Helsley and Strange (1990). A variant of this explanation has been proposed by Duranton (1998). In Duranton's explanation, large labor markets confer benefits on firms and workers by allowing workers to become more specialized, and therefore more productive, in the tasks they perform.

[15]There is strong empirical evidence of the importance to local development of a large diversified base of non-tradable inputs (Saxenian, 1996). Examples of such inputs include legal and communication services, maintenance and repair services, financial services, and large non-tradable industrial inputs such as industrial waste disposal. For theoretical research on this phenomenon, see Abdel-Rahman and Fujita (1990).

economic development. For example, the spatial concentration of population often results in congestion of local public goods, such as roadways, and creates disamenities such as noise and air pollution.

1.3 Elements of the Model

This section describes the elements that the theoretical model must contain in order to analyze the implications of NBG for regional development and explains how these elements are included in the model.

1.3.1 Non-Balanced Growth

In the theoretical model, it is necessary that there is economic growth and that the growth be non-balanced, i.e., that it results in change in sectoral output and employment shares. In the theoretical model, economic growth is endogenous and results from purposeful innovative activity, as in a Romer (1990)-type model of endogenous growth with horizontally differentiated inputs. In particular, purposeful innovative activity in these models leads to Hicks-neutral technical change. Hicks-neutral technical change is a form of total factor productivity growth where technical change does not influence the marginal rate of substitution between inputs.[16] It is necessary that economic growth in the model be endogenous for agglomeration externalities and the spatial distribution of production to influence the pace of economic growth.

As is mentioned in the Introduction, this essay considers three different explanations for NBG. The baseline model presented in Chapter 1.4 considers supply-side

[16]Formally, total factor productivity is the ratio of output to an index of inputs (Chambers 1988). Growth in total factor productivity is a useful measure of technical change because the ratio of total factor productivity from two different time periods provides a measure of the change in the effectiveness of the index of inputs in producing output.

NBG, where NBG results from differences in productivity growth between sectors. It is assumed in the baseline model that one sector (i.e., the goods-producing sector) benefits from technological advance, and hence from Hicks-neutral technical change, while the other sector (i.e., the service sector) does not. The direction of supply-side NBG in the baseline model is shown to be determined by the elasticity of substitution between the two sectors' output in household preferences. When the elasticity of substitution is less than 1, economic growth causes factor shares in the less progressive sector to increase; when the elasticity of substitution is greater than 1, the opposite result holds, and economic growth causes factor shares in the progressive sector to increase. The intuition for the role of the elasticity of substitution in determining the direction of supply-side NBG is given in Chapter 1.5.3.

In Chapter 1.7.1, supply-side NBG in the model results from the assumption of non-homothetic consumer preferences. In particular, consumer preferences in Chapter 1.7.1 are non-homothetic with constant elasticity of substitution equal to one. (In contrast, consumer preferences in the baseline model of supply-side NBG are homothetic and additively separable with a non-negative constant elasticity of substitution.) Non-homothetic consumer preferences imply different income elasticities of demand for sectoral output and lead to uneven patterns of employment and output growth between sectors. In Chapter 1.7.2, which considers the spatial consequences of NBG in an open economy, NBG occurs as a result of changes in the terms of trade between countries.

1.3.2 Regional Economic Characteristics

In order for NBG to influence the spatial distribution of the economy in the model, it is necessary that regions' economic characteristics (e.g., population, climate, proximity to natural resources, and industrial composition) influence production on at least two dimensions and that their relative importance differs between the two sectors in the model. Regional economic characteristics must influence production on at least two dimensions for the spatial distribution of the economy to reflect the trade-off between competing economic forces. Furthermore, it is necessary that the relative importance of these regional economic characteristics differ between sectors for this trade-off to be altered by NBG and lead to spatial change in the model. The two regional economic characteristics included in the theoretical model are agglomeration externalities and an interregionally immobile factor of production.

Agglomeration externalities are included in the model in a manner consistent with core-periphery models of economic geography (e.g., Krugman 1991). In core-periphery models, the incentive for agglomeration arises from a combination of increasing returns to scale in production and transportation costs. The technical details for how agglomeration externalities are included in the theoretical model are given in Chapter 1.5. Agglomeration externalities represent a region's endogenous economic characteristics, or second-nature features, in the model and are both static, providing incentive for the spatial concentration of population and production, and dynamic, influencing the pace of regional productivity growth.

The interregionally immobile factor of production, which is referred to as the "fixed factor," is taken to represent all of a region's first-nature features that influence either firm production decisions. As such, the immobile factor represents regional economic characteristics such as the amount of available land, climate, and proximity to deposits of natural resources. As the immobile factor is scarce, it provides an incentive for the spatial dispersion of economic activity in the model.

In the theoretical model, it is assumed that agglomeration externalities only influence production in the goods-producing sector and that the fixed factor is only used in production in the service sector. These assumptions simplify the analysis while allowing the relative (and absolute) importance of regional economic characteristics to differ between the two sectors, which is required for NBG to influence the spatial distribution of the economy in the model. These assumptions, however, are arbitrary in that regions' first (the fixed factor) and second nature (agglomeration externalities) features influence production in both sectors, and that there is no empirical evidence in the published literature that suggests that the relative importance of regions' first and second nature features differs systematically between the two sectors.

1.3.3 Factor Mobility and General Equilibrium

It is necessary to assume in the theoretical model that regions share in national markets for factors of production, including labor. The assumption of national factor markets is defensible in a model of regional development because there are fewer legal, cultural, and technological constraints on the movement of factors across regions within a country than between countries. Fewer constraints mean that factors of production, including labor, will migrate quickly between regions within a country in response to interregional price (wage) differentials. Blanchard and Katz (1992) show that interstate labor migration is the main channel of adjustment to local economic shocks for U.S. states.

The assumption of national factor markets has three important implications for the analysis in this essay. First, a national factor market means that interregional differences in savings rates should not influence the pace of regional development in the analysis. Differences in savings rates are thought to be an important determinant of cross-country differences in economic growth. Second, labor mobility means that interregional differences in population growth rates are not related to differences in exogenous labor endowments or regional fertility. Third, national factor markets mean that the distribution of employment, and hence population, across regions is determined in general equilibrium. It is necessary that the interregional distribution of employment be determined in general equilibrium for NBG to influence the spatial distribution of population in the model.

1.4 The Model

This section defines the preferences and technology in the baseline model of the spatial consequences of supply-side non-balanced growth. The term baseline model is used because the assumptions about the technology in the economy in this chapter are maintained in the model of the spatial consequences of demand-side non-balanced growth presented in Chapter 1.7.1 and in the model of the spatial consequences of non-balanced growth in an open-economy presented in Chapter 1.7.2.

The model presented in this section exploits the analytical similarities between Romer (1990)-type models of endogenous growth with horizontally differentiated inputs and Krugman (1991)-type core-periphery models of economic geography. The analytical similarities between these models stem from their reliance on the monopolistic competition framework introduced by Dixit and Stiglitz (1977). Previous authors, such as Walz (1996) and Fujita and Thisse (2003), have demonstrated that the monopolistic competition framework is amenable to the joint study of growth and location because it allows for (i) endogenous growth, (ii) agglomeration externalities, and (iii) interregional factor movement to be dealt within a unified theoretical model.

The main theoretical innovation in the theoretical model is that it extends previous models in the literature to allow endogenous labor allocation across both sectors and regions within an analytically tractable growth model under perfect foresight. This extension is essential for analyzing both non-balanced growth (intersectoral labor mobility) and its spatial consequences (interregional labor mobility). Previous multi-region growth models have either constrained the interregionally mobile factor of production (labor) to employment in one sector (Baldwin and Forslid 2000; Fujita and Thisse 2003) or do not have fully developed dynamics (Walz 1996).

As is standard in Romer (1990)-type models of growth, endogenous economic growth in the model results from the expansion of the number of horizontally-differentiated intermediate inputs available for production. The model is also in keeping with the standard core-periphery assumption that agglomeration externalities arise from the combination of increasing returns to scale and transportation costs, which together provide an incentive for demanders and suppliers of intermediate inputs to locate in the same region. However, the model departs from the standard core-periphery framework by assuming that both suppliers and demanders of intermediate inputs are free to move between regions at zero cost. Assuming that firms are "footloose" avoids some of the analytical complexity of the core-periphery framework and allows a factor of production (labor) that is mobile between both sectors and regions to be included in a model that is analytically tractable.

1.4.1 Preferences

There are two regions in the economy, a and b, three sectors, the manufacturing or goods-producing sector (M-sector), the innovative sector (I-sector), and the service sector (S-sector), and H households. Each household member chooses their location (region a or b) and sector of employment (the M-, I-, or S-sector) to maximize their wage.

A typical household, household j, is of size $l_j(t)$ at time t and grows according to

$$l_j(t) = e^{nt} l_j(0), \tag{1.1}$$

where $l_j(0) > 0$ is household j's size at $t = 0$ and $n \in [0, \delta)$, where δ is its rate of time preference. Each household member supplies inelastically one unit of labor per unit of time, so that $l_j(t)$ also denotes household j's labor supply at time t.[17] It is assumed

[17] The assumption that each household member supplies inelastically one unit of labor per unit of time implies that neither involuntary unemployment nor labor/leisure trade-offs are considered in the model. These issues are ignored because the focus of the analysis is on how change in the distribution of employment across sectors in the process of economic growth influences the spatial development of the economy.

that all population growth takes place within existing households and that both n and δ are constant over time and identical for all households in the economy.[18,19]

Household members receive utility from the consumption of M- and S-sector output. Each member h of household j has a utility function given by

$$u\left(Y_{hjM}(t), Y_{hjS}(t)\right) = \ln\left[\mu Y_{hjM}(t)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\mu)\, Y_{hjS}(t)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \qquad (1.2)$$

where $Y_{hjM}(t)$ and $Y_{hjS}(t)$ are household member h's consumption of M- and S-sector output at time t, $\varepsilon \in (0, \infty)$ is the elasticity of substitution between M- and S-sector output, and $\mu \in (0, 1)$ is a parameter that determines household member h's expenditure shares for M- and S-sector output.

Household j's utility at time t is the summation of the utilities of all household members at time t. (1.2) implies that household j maximizes utility at time t by spreading consumption evenly across all of its members.[20] As such, household j's

[18] That households grow in size at an exponential rate, as is described by (1.1), implies that while the number of households in the economy is a fixed integer, household size will take non-integer values. A households size at time t determines its effective labor supply and, from (1.3), its utility from the consumption of a given amount of M- and S-sector output. Allowing household size to take non-integer values does not create problems for the analysis because neither the technology in the model, which described in Chapter 1.4.2, nor household preferences require that household size be an integer.

[19] The assumption that n and δ are the same for every household in economy, while unrealistic, allows the dynamics of economic growth in the economy to be analyzed without having to consider the distribution of income across households. This is the case because the assumption n and δ are the same for every household implies that expenditure increases at the same rate for each household at a given time t regardless of their level of expenditure (see (1.13) below). The fact that expenditure increases at the same for each household implies that only the level of total societal expenditure, and not the distribution of total societal expenditure across households, influences the rate of increase in societal expenditure at time t.

[20] Appendix A gives a formal proof that household member h's utility function (1.2) implies that household j maximizes utility at time t by spreading consumption evenly across all of its members.

utility at time t can be written as

$$U\left(Y_{jM}(t), Y_{jS}(t); l_j\left(t\right)\right) = l_j\left(t\right) u\left(\frac{Y_{jM}(t)}{l_j\left(t\right)}, \frac{Y_{jS}(t)}{l_j\left(t\right)}\right), \tag{1.3}$$

where $Y_{jM}(t)$ and $Y_{jM}(t)$ are household j's total consumption of M- and S-sector output at time t (hence, $Y_{jM}(t)/l_j\left(t\right)$ and $Y_{jM}(t)/l_j\left(t\right)$ are consumption of M- and S-sector output per household member at time t). (1.3) implies that $\varepsilon \in (0, \infty)$ is the elasticity of substitution between M- and S-sector output for household j as a whole, and that $\mu \in (0, 1)$ determines household j's expenditure shares for M- and S-sector output.

The household j's indirect utility function at time t corresponding to the direct utility function in (1.3) is

$$\mathbb{V}_j\left(t\right) = l_j\left(t\right) \ln\left\{\frac{E_j(t)}{l_j\left(t\right)}\left[\mu^{\varepsilon} + (1-\mu)^{\varepsilon} p(t)^{1-\varepsilon}\right]^{\frac{1}{\varepsilon-1}}\right\}, \tag{1.4}$$

where $E_j(t)$ is household j's expenditure at time t, $p\left(t\right)$ is the price of S-sector output at time t, and the price of M-sector output normalized to 1. Notice that $E_j(t) > 0$ for all $t \in [0, \infty)$ because $\lim_{E_j(t)\to 0} \frac{\partial \mathbb{V}_j(t)}{\partial E_j(t)} = \infty$.[21] Household j's demands for M- and S-sector output at time t corresponding to the direct utility function in (1.3) are

$$Y_{jM}(t) = \frac{E_j(t)}{1 + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p(t)^{1-\varepsilon}} \quad \text{and} \quad Y_{jS}(t) = \frac{E_j(t)}{p(t) + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p(t)^{\varepsilon}}. \tag{1.5}$$

Household j holds assets at time t, $a_j\left(t\right)$, in the form of ownership claims on I-sector firms (see Chapter 1.4.2) or as loans, with negative loans representing debts. These two forms of assets are assumed to be perfect substitutes as stores of value and, as such, bear the same real interest rate, $r\left(t\right)$, at time t. These assets allow for the transfer of funds between households, who want to smooth consumption, and I-sector firms, who require investment to develop new varieties of intermediate inputs. The model in this chapter represents a closed economy, so while households can borrow from and lend to each other, there are zero net loans in the economy for all t.

[21] (1.4) and (1.5) are dervied in Appendix B.

The total income received by household j at time t is the sum of labor income, $w_j(t)l_j(t)$, where $w_j(t)$ is average wage of members of household j at time t, asset income, $r(t)\,a_j(t)$, where $r(t)$ is the interest rate at time t, and income from ownership of the fixed factor in regions a and b, $w_a^F(t)f_{ja} + w_b^F(t)f_{jb}$, where $w_k^F(t)$ is the rent to the fixed factor in region k, $k = a, b$, at time t, and $f_{jk} \geq 0$ is the amount of the fixed factor in region k owned by household j (the fixed factor is used in production by firms in the S-sector and is immobile between regions; see Chapter 1.4.2).[22] Given these sources of income, household j's assets evolve over time according to

$$\dot{a}_j(t) \equiv \frac{da_j(t)}{dt} = w_j(t)l_j(t) + w_a^F(t)f_{ja} + w_b^F(t)f_{jb} + r(t)\,a_j(t) - E_j(t). \tag{1.6}$$

The objective of household j is to find an expenditure path that maximizes the discounted sum of its future instantaneous utilities,

$$\max_{[E_j(t)]} \int_0^\infty l_j(0) \ln \left\{ \frac{E_j(t)}{l_j(t)} \left[\mu^\varepsilon + (1-\mu)^\varepsilon p(t)^{1-\varepsilon}\right]^{\frac{1}{\varepsilon-1}} \right\} e^{-(\delta-n)t} dt, \tag{1.7}$$

subject to (1.6), its initial size, $l_j(0)$, asset holdings, $a_j(0)$, and ownership of the fixed factor, f_{ja} and f_{jb}.[23]

The current-value Hamiltonian corresponding to household j's intertemporal pro-

[22] The average wage for household j is used because household members are free to choose their sector and region of employment to maximize their wage, so it may be the case that not all household members have the same wage at time t.

[23] A sufficient condition for the improper intergral in (1.7) to converge is for $\delta - n > 0$ (which is assumed) and for household j's attainable utility to be bounded for all $t \in [0, \infty)$, i.e., for $\left| l_j(0) \ln \left\{ \frac{E_j(t)}{l_j(t)} \left[\mu^\varepsilon + (1-\mu)^\varepsilon p(t)^{1-\varepsilon}\right]^{\frac{1}{\varepsilon-1}} \right\} \right| \leq B \in \mathbb{R}_{++}$ for all $t \in [0, \infty)$. When these two conditions hold, which is assumed in this chapter, the improper integral in (1.7) will converge and not exceed $\frac{B}{\delta - n}$ (Caputo 2005, p.384). The assumption that household j's attainable utility is bounded for all $t \in [0, \infty)$ is equivalent to assuming that household j's resources are constrained so that it cannot obtain an arbitrarily large level utility (i.e., an arbitrarily large level of consumption) at any time t.

gram is

$$H^c\left(a_j\left(t\right), E_j\left(t\right), v_j(t)\right) = l_j\left(0\right)\left\{\ln E_j(t) - \ln l_j\left(t\right) + \ln\left[\mu^{\varepsilon} + \left(1-\mu\right)^{\varepsilon} p(t)^{1-\varepsilon}\right]^{\frac{1}{\varepsilon-1}}\right\} \tag{1.8}$$

$$+ v_j(t)\left(w_j(t)l_j\left(t\right) + w_a^F(t)f_{ja} + w_a^F(t)f_{ja} + r\left(t\right)a_j\left(t\right) - E_j(t)\right),$$

where $v_j(t)$ is the co-state variable. Applying Theorem 14.3 from Caputo (2005), if $\left\{a_j^*\left(t\right), E_j^*\left(t\right)\right\}$ is an optimal solution, then it is necessary that there exists a piecewise smooth function $v_j(t)$ such that for all $t \in [0,\infty)$,[24]

$$\frac{\partial H^c\left(a_j^*\left(t\right), E_j^*\left(t\right), v_j^*(t)\right)}{\partial E_j(t)} = \frac{l_j\left(0\right)}{E_j^*\left(t\right)} - v_j^*(t) = 0 \tag{1.9}$$

$$-\frac{\partial H^c\left(a_j^*\left(t\right), E_j^*\left(t\right), v_j^*(t)\right)}{\partial a_j(t)} = -v_j^*(t)r(t) = \dot{v}_j(t) - \left(\delta - n\right)v_j^*(t) \tag{1.10}$$

and

$$\frac{\partial H^c\left(a_j^*\left(t\right), E_j^*\left(t\right), v_j^*(t)\right)}{\partial v_j(t)} = \dot{a}_j(t) = w_j(t)l_j\left(t\right)+w_a^F(t)f_{ja}+w_b^F(t)f_{jb}+r\left(t\right)a_j^*\left(t\right)-E_j^*\left(t\right). \tag{1.11}$$

Applying Theorem 14.4 and Lemma 14.1 from Caputo (2005), given that $H^c\left(a_j\left(t\right), E_j\left(t\right), v_j(t)\right)$ is a concave function of $a_j\left(t\right)$ and $E_j\left(t\right)$ for all $t \in [0,\infty)$ over the open convex set containing all the admissible values of $a_j\left(t\right)$ and $E_j\left(t\right)$, i.e., over the set $\left\{\left(a_j\left(t\right), E_j\left(t\right)\right) \mid a_j\left(t\right) \in \mathbb{R}, E_j\left(t\right) > 0\right\}$, the necessary conditions in (1.9), (1.10), and (1.11) are sufficient to identify the global maximum to household j's intertemporal program if the transversality condition

$$\lim_{t\to\infty} a_j^*\left(t\right)v_j(t)e^{-(\delta-n)t} = 0 \tag{1.12}$$

[24]The function $v_j(t)$ is piecewise smooth on the interval $t \in [0,\infty)$ if its deriative function $\dot{v}_j(t)$ is piecewise continous on the interval $t \in [0,\infty)$ (Caputo 2005, Definition 1.2). The function $\dot{v}_j(t)$ is piecewise continous on on the interval $t \in [0,\infty)$ if the inteval $t \in [0,\infty)$ can be partitioned into a finite number of points $0 = t_0 < t_1 < ... < t_k < \infty$ such that (i) $\dot{v}_j(t)$ is continous on each open subinterval $t_{k-1} < t < t_k$ and (ii) $\dot{v}_j(t)$ approaches a finite limit as the end points of each subinteval are approached from within the subinterval (Caputo 2005, Definition 1.1).

is satisfied.[25]

Combining (1.9) and (1.10), the optimal expenditure path for household j is characterized by the Euler equation

$$\hat{E}_j(t) \equiv \dot{E}_j(t)/E_j(t) = r(t) - (\delta - n)\,, \tag{1.13}$$

together with the initial conditions, $l_j(0)$, $a_j(0)$, f_{ja}, and f_{jb}, the state equation (1.6), and the transversality condition (1.12).

Household j's transversality condition in (1.12) can be re-expressed in a more intuitive form. Integrating (1.10) with respect to time yields

$$v_j(t) = v_j(0)\exp\left(-\int_0^t r(\tau)d\tau\right)\exp\left[(\delta - n)\,t\right]. \tag{1.14}$$

The term $v_j(0)$ equals the marginal utility of expenditure at time 0, i.e., $v_j(0) = \frac{\partial \mathbb{V}_j(0)}{\partial E_j(0)} = \frac{l_j(0)}{E_j(0)}$, which is positive and finite as $E_j(0)$ is positive and finite. Substituting this expression for $v_j(t)$ from (1.14) into the transversality condition from (1.12) yields

$$\lim_{t\to\infty} a_j(t)\exp\left(-\int_0^t r(\tau)d\tau\right) = 0. \tag{1.15}$$

(1.15) implies that when the transversality condition is satisfied, an optimal expenditure path entails household j's assets, $a_j(t)$, growing asymptotically at a rate lower than $r(t)$. When the transversality condition holds for every household in the economy, it rules out chain-letter debt financing, because for a given household to borrow on a perpetual basis – which would imply that its debt was growing at a rate higher than $r(t)$ – there would have to be at least one lender in the economy willing to violate their own transversality condition by holding positive assets that grow at a rate higher than $r(t)$. This would be suboptimal for the lender, and, hence, would not occur in equilibrium.

[25] Theorems 14.3 and 14.4 and Lemma 14.1 in Caputo (2005) are presented for the present-value Hamiltonian. Adjustments have been made for the use of the current-value Hamiltonian in this chapter.

Household / Preference Parameters	
$H > 0$	**Total number of households in the economy.** H is fixed over time.
$n \in [0, \delta)$	**Rate of growth in household size.** n is constant over time and identical for all households in the economy.
$\delta > 0$	**Household rate of time preference.** δ is constant over time and identical for all households in the economy.
$\varepsilon \in (0, \infty)$	**Elasticity of substitution between M- and S-sector output in household j's utility function.**
$\mu \in (0, 1)$	**Parameter of household j's utility function that determines its expenditure shares for M- and S-sector output.**
$f_k(j) \geq 0$	**The amount of the fixed factor in region k, $k = a, b$, owned by household j.**

Table 1.1: Household / Preference Parameters

Household / Preference Variables	
$l_j(t)$	**Household j's size at time t.** Household members supply their labor inelastically, so $l_j(t)$ also denotes household j's labor supply at time t.
$L(t) = \sum_{j=1}^{H} l_j(t)$	**Total population (i.e., total labor supply) at time t.**
$U(Y_{jM}(t), Y_{jS}(t), l_j(t))$	**Household j's utility at time t.**
$Y_{jM}(t)$ & $Y_{jS}(t)$	**Household j's consumption of M- and S-sector output at time t.**
$E_j(t)$	**Household j's expenditure at time t.**
$E(t) = \sum_{j=1}^{H} E_j(t)$	**Societal expenditure at time t.**
$a_j(t)$	**Household j's asset holdings at time t.**

Table 1.2: Household / Preference Variables

Technology Parameters	
$\alpha \in (0,1)$	**Technology parameter in the M-sector.** α determines the output elasticities of $I_k(t)$ and $L_{kM}(t)$ in M-sector production. In addition, $(1-\alpha)^{-1}$ is the elasticity of substitution between any two intermediate inputs in M-sector production.
$\Gamma \geq 1$	**"Iceberg" transportation costs for intermediate inputs.** The transport of an intermediate input within the region where it is produced is costless, but when the intermediate input is transported between regions, only a fraction, $1/\Gamma \leq 1$, of the input arrives.
$\eta \in [0,1]$	**Technology parameter in I-sector innovation.** η is proportional to the ease of transferring knowledge between workers in I-sector innovation in region a and region b.
$\beta \in (0,1)$	**Technology parameter in the S-sector.** β determines the share of total cost devoted to labor and to the fixed factor in the S-sector.
$F_k > 0$	**The quantity of the fixed factor in region k, $k = a, b$.**

Table 1.3: Technology Parameters

1.4.2 Technology

In this section, the technology in the M-, I-, and S-sectors is described.

M-Sector

M-sector output is produced competitively using a constant returns to scale Cobb-Douglas production technology that combines labor and a number of differentiated intermediate inputs to produce a final output,

$$Y_{kM}(t) = I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha}, \; I_k(t) = \left(\sum_{i \in \mathbb{M}(t)} q_{ik}(t)^{\alpha} \right)^{\frac{1}{\alpha}}, \tag{1.16}$$

where $Y_{kM}(t)$ is the quantity of M-sector output produced in region k at time t, $L_{kM}(t) \geq 0$ is the quantity of labor employed in the M-sector in region k at time t, and $I_k(t)$ is an index of intermediate inputs used in M-sector production in region k at time t, where $q_{ik}(t) \geq 0$ is the quantity of I-sector firm i's intermediate input used in M-sector production in region k at time t and $\mathbb{M}(t)$ is the set of I-sector

Prices	
$r(t)$	**The interest rate at time t.**
$p(t)$	**Price of S-sector output at time t** (in units of M-sector output).
$p_{i_k k}(t)$	**Price in region k, $k = a, b$, at time t of the intermediate input produced by the I-sector firm i_k that is located in region k** (in units of M-sector output).
$p_{i_k l}(t)$	**Price in region l, $l \neq k$, at time t of the intermediate input produced by the I-sector firm i_k that is located in region k** (in units of M-sector output). Transportation costs in the I-sector imply that $p_{i_k l}(t) = \Gamma p_{i_k k}(t)$ for all $i_k \in \mathbb{M}(t)$ and for all t.
$P_k(t)$	**Price index that gives the minimum cost for M-sector firms in region k, $k = a, b$, of purchasing a unit of $I_k(t)$ at time t** (in units of M-sector output).
$w_j(t)$	**Average wage of members of household j at time t** (in units of M-sector output).
$w_k(t)$	**Wage in region k, $k = a, b$, at time t** (in units of M-sector output).
$w_k^F(t)$	**Rent to the fixed factor in region k, $k = a, b$, at time t** (in units of M-sector output).

Table 1.4: Prices

Technology Variables	
$L(t) = \sum_{j=1}^{H} l_j(t)$	**Total labor supply (i.e., total population) at time t.**
$L_{kP}(t)$	**Labor employed in the P-sector, $P = M, I, S$, in region k, $k = a, b$, at time t.** $L_P(t) = L_{aP}(t) + L_{bP}(t)$ is total labor employed in the P-sector at time t.
$\lambda_P(t) \equiv \frac{L_{aP}(t)}{L_P(t)}$	**Proportion of P-sector employment, $P = M, I, S$, in region a at time t.** $\lambda_N(t) \equiv N_a(t)/N(t)$ is the proportion of I-sector firms in region a.
$Y_{kP}(t)$	**Quantity of P-sector output, $P = M, I, S$, produced in region k, $k = a, b$, at time t.** $Y_P(t) = Y_{aP}(t) + Y_{bP}(t)$ is the total quantity of P-sector output produced in the economy at time t.
$\Pi_{kP}(t)$	**P-sector profits, $P = M, I, S$, in region k, $k = a, b$, at time t.** Because constant returns to scale technology is assumed in the M- and S-sectors and in I-sector innovation, $\Pi_{kP}(t) = 0$ in equilibrium for $P = M, I, S$.
$\mathbb{M}_k(t)$	**The set of I-sector firms operating in region k, $k = a, b$, at time t.** $M(t) = M_a(t) \cup M_b(t)$ is the set of all I-sector firms operating at time t.
$N(t)$	**The number of I-sector firms and the number of varieties of intermediate inputs in the economy at time t.** It is assumed that each I-sector firm is the monopoly supplier of a single variety of intermediate input.
$N_k(t)$	**Number of I-sector firms in region k, $k = a, b$, at time t.** It is assumed that each I-sector firm produces in only one region, so that $N_a(t) + N_b(t) = N(t)$.
$q_{i_k k}(t)$ and $q_{i_k l}(t)$	**The quantity of I-sector firm i_k that is located in region k's intermediate input demanded by M-sector firms in regions k and l, respectively, at time t.**
$I_k(t)$	**Index of intermediate inputs used in M-sector production in region k, $k = a, b$, at time t.**
$Q_{i_k}(t)$	**The quantity of intermediate input produced by I-sector firm i_k that is located in region k, $k = a, b$, at time t.**
$Y_{i_k M}(t)$	**The quantity of M-sector output used in production by I-sector firm i_k that is located in region k, $k = a, b$, at time t.**
$\pi_{i_k}(t)$	**The profits of I-sector firm i_k that is located in region k, $k = a, b$, at time t.**
$\Omega_k(t)$	**The marginal product of labor in I-sector innovation in region k, $k = a, b$, at time t.** $\Omega_k(t) \equiv N(t)\left[\frac{L_{kI}(t)}{L_I(t)} + \eta \frac{L_{lI}(t)}{L_I(t)}\right]$
$\dot{N}_k(t)$	**The number of new varieties of intermediate input developed in region k at time t.** $\dot{N}(t) = \dot{N}_a(t) + \dot{N}_b(t)$ is the total number of new varieties of intermediate inputs developed in the economy at time t.
$V(t)$	**Value of an I-sector firm at time t.** $V(t)$ is the present value of future profits from the production of a unique variety of intermediate input at time t. As I-sector firms are free to move between regions to maximize profits, $V(t)$ does not depend on the region in which the firm is located at time t.

Table 1.5: Technology Variables

firms operating at time t.[26] The fixed parameter $\alpha \in (0,1)$ determines the output elasticities of $I_k(t)$ and $L_{kM}(t)$, where α is the output elasticity of $I_k(t)$ and $(1-\alpha)$ is the output elasticity of $L_{kM}(t)$. In addition, $(1-\alpha)^{-1}$ is the elasticity of substitution between any two intermediate inputs in M-sector production.

The assumption that α determines both the output elasticities of $I_k(t)$ and $L_{kM}(t)$ and the elasticity of substitution between intermediate inputs imposes certain restrictions on M-sector production. In particular, the assumption that $\alpha \in (0,1)$ implies that the output elasticities of $I_k(t)$ and $L_{kM}(t)$ are less than one and that the elasticity of substitution between any two intermediate inputs is greater than one, i.e., $(1-\alpha)^{-1} > 1$. In addition, $\alpha < 1$ implies that $I_k(t)$ is a concave function, and $\alpha > 0$ implies that no intermediate input is essential to M-sector production (see Dixit and Stiglitz 1977).

I-sector output is subject to transportation costs. In particular, the transport of an intermediate input within the region where it is produced is costless, but when the intermediate input is transported across regions, only a fraction, $1/\Gamma \leq 1$, of the input arrives. As I-sector firms cannot price discriminate across regions, transportation costs in the I-sector imply that $p_{i_k l}(t) = p_{i_k k}(t)\Gamma$, where $p_{i_k k}(t)$ is the price of firm i_k's intermediate input in region k at time t and $p_{i_k l}(t)$ is the price of firm i_k's intermediate input in region l at time t (the "k" subscript implies that I-sector firm i_k is located in region k at time t).

The cost minimization problem for M-sector firms in region k at time t is

$$\min_{[q_i(t)]_{i\in\mathbb{M}(t)}, L_{kM}(t)} \sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t) + w_k(t)\, L_{kM}(t)$$

[26] Constant returns to scale technology means that only regional levels of production and input use in the M-sector need be considered (this also holds for the S-sector and for innovative activity in the I-sector, where constant returns to scale are also assumed). See Appendix C for a formal proof that when all firms in a sector operate with identical constant returns to scale production technologies, sectoral production can be represented by an aggregate production function.

$$s.t.\ I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM} \geq 0,\ q_i(t) \geq 0,\ i = 1, ..., N(t),\ \text{ and } L_{kM}(t) \geq 0,$$

where $w_k(t)$ is the wage in region k at time t and $\mathbb{M}(t)$ is the set of all I-sector firms operating at time t ($\mathbb{M}(t) = \mathbb{M}_a(t) \cup \mathbb{M}_b(t)$). Appendix D demonstrates that the cost-minimizing input choices for M-sector firms in region k at time t are derived in two stages. This two-stage procedure yields the following expressions for the cost-minimizing input demands:

$$\begin{aligned} I_k(t) &= Y_{kM}\left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha}\left(\frac{w_k(t)}{P_k(t)}\right)^{1-\alpha} \\ L_{kM}(t) &= Y_{kM}\left(\frac{\alpha}{1-\alpha}\right)^{-\alpha}\left(\frac{w_k(t)}{P_k(t)}\right)^{-\alpha}, \end{aligned} \tag{1.17}$$

and for the cost function for the M-sector in region k at time t:

$$C_{kM}(t) = Y_{kM}\left(\frac{P_k(t)}{\alpha}\right)^{\alpha}\left(\frac{w_k(t)}{1-\alpha}\right)^{1-\alpha}, \tag{1.18}$$

where $P_k(t) \equiv \left(\sum_{i\in\mathbb{M}(t)} p_{ik}(t)^{\frac{\alpha}{\alpha-1}}\right)^{\frac{\alpha-1}{\alpha}}$ is a price index that gives the minimum cost for M-sector firms in region k of purchasing a unit of $I_k(t)$ at time t. (1.18) implies that the average cost of production for the M-sector in region k at time t is constant (i.e., independent of output). That the average cost of production in the M-sector is constant follows directly from the assumptions that the M-sector operates with constant returns to scale technology and no fixed costs.

Factors of production will move into the M-sector in region k at time t until the factor prices, $P_k(t)$ and $w_k(t)$, have adjusted so that the price of M-sector output (which is normalized to 1) is equal to the constant average cost of production and profits are driven to zero. From 1.18, the condition that the price of M-sector output equals the average cost of production for the M-sector in region k at time t implies that

$$1 = \left(\frac{P_k(t)}{\alpha}\right)^{\alpha}\left(\frac{w_k(t)}{1-\alpha}\right)^{1-\alpha}. \tag{1.19}$$

(1.19) defines the relationship between $w_k(t)$ and $P_k(t)$ in region k at time t as

$$w_k(t) = (1-\alpha)\,\alpha^{\frac{\alpha}{1-\alpha}} P_k(t)^{\frac{-\alpha}{1-\alpha}}. \tag{1.20}$$

In addition, the two-stage cost minimization procedure described in Appendix D implies that the demands for firm i_k's variety of intermediate input by M-sector firms in region k and region l at time t are

$$q_{i_k k}(t) = p_{i_k k}(t)^{\frac{1}{\alpha-1}} \left(\frac{\alpha}{1-\alpha}\right) w_k(t)\, L_{kM}(t)\, P_k(t)^{\frac{-\alpha}{a-1}} \tag{1.21}$$

and

$$q_{i_k l}(t) = p_{i_k l}(t)^{\frac{1}{\alpha-1}} \left(\frac{\alpha}{1-\alpha}\right) w_l(t)\, L_{lM}(t)\, P_l(t)^{\frac{-\alpha}{a-1}}.$$

where $p_{i_k k}(t)$ is the price of I-sector firm i_k's intermediate input in region k at time t and $p_{i_k l}(t)$ is the price of firm i_k's intermediate input in region l at time t. Using these expressions and the fact that $p_{i_k l}(t) = p_{i_k k}(t)\Gamma$ for all t because I-sector firms cannot price discriminate across regions, the total demand for firm i_k's variety of intermediate input at time t is given by

$$q_{i_k k}(t) + \Gamma q_{i_k l}(t) = p_{i_k k}(t)^{\frac{1}{\alpha-1}} \Delta_k(t), \tag{1.22}$$

where

$$\Delta_k(t) = \left(\frac{\alpha}{1-\alpha}\right)\left(w_k(t)\, L_{kM}(t)\, P_k(t)^{\frac{-\alpha}{a-1}} + \Gamma^{\frac{\alpha}{a-1}} w_l(t)\, L_{lM}(t)\, P_l(t)^{\frac{-\alpha}{a-1}}\right).$$

From (1.22), firm i_k faces a constant elasticity demand curve for its variety of intermediate input, where $(\alpha-1)^{-1}$ is the price elasticity of demand.

S-Sector

S-sector output is produced competitively using a constant returns to scale Cobb-Douglas production technology that combines labor and the fixed factor to produce a final output

$$Y_{kS}(t) = L_{kS}(t)^{\beta} F_k{}^{1-\beta}, \tag{1.23}$$

where $L_{kS}(t) \geq 0$ is the quantity of labor employed in the S-sector in region k, $F_k > 0$ is the quantity of the fixed factor in region k, and $\beta \in (0,1)$ is a parameter that determines the share of total cost devoted to labor and to the fixed factor. Regions a and b are identical except for their endowments of the fixed factor. The fixed factor represents all of a region's innate geographic characteristics that influence either firm production. As is detailed in Chapter 1.2, these innate geographic characteristics, or first-nature features, include the amount of available land in a region and its proximity to deposits of natural resources.

The profit maximization problem for the S-sector in region k at time t is

$$\max_{L_{kS}(t)\geq 0} p(t) L_{kS}(t)^{\beta} F_k^{1-\beta} - w_k(t) L_{kS}(t) - w_k^F(t) F_k. \tag{1.24}$$

$L_{kS}(t)$ must be greater than zero at an optimum because $\lim\limits_{L_{kM}(t)\to 0} dY_{kS}(t)/dL_{kS}(t) = \infty$ for $\beta \in (0,1)$. The first-order condition for (1.24) is

$$\beta p(t) L_{kS}(t)^{\beta-1} F_k^{1-\beta} - w_k(t) = 0. \tag{1.25}$$

(1.25) gives the following expression for S-sector labor demand in region k at time t:

$$L_{kS}(t) = \left(\frac{\beta p(t)}{w_k(t)}\right)^{\frac{1}{1-\beta}} F_k. \tag{1.26}$$

The second-order condition for (1.24) is

$$-\beta(1-\beta) p(t) L_{kS}(t)^{\beta-2} F_k^{1-\beta} < 0,$$

which establishes that $p(t) L_{kS}(t)^{\beta} F_k^{1-\beta} - w_k(t) L_{kS}(t) - w_k^F(t) F_k$ is strictly concave on $\mathbb{R}_+$ and that (1.26) is the unique maximum of (1.24).

I-Sector Production

The development of new varieties of intermediate inputs takes place in the I-sector. The I-sector is characterized by monopolistic competition. The monopoly position of each I-sector firm allows it to charge a price over average cost and make positive profits in equilibrium; however, as in a competitive economy, the demand for each I-sector firm's variety of intermediate input is influenced by the production decisions of all other firms in the I-sector. It is assumed that each I-sector firm is the monopoly supplier of a single variety of intermediate input, so that $N(t)$ is both the number of I-sector firms at time t and the number of varieties of intermediate inputs at time t.[27]

Each I–sector firm produces with a linear production technology that uses one unit of M-sector output to produce one unit of its unique variety of intermediate input. Given this technology, the term $Y_{i_k}(t) \geq 0$ denotes the quantity of intermediate input produced by I-sector firm i_k at time t and $Y_{i_k M}(t)$ denotes the quantity of M-sector output used in production by I-sector firm i_k with $Y_{i_k}(t) = Y_{i_k M}(t)$. It is assumed that the development of an intermediate input involves the construction of indivisible, firm-specific physical capital, so that each intermediate input is produced in only one region. This means that at time t there are $N_a(t)$ I-sector firms in region a and $N_b(t)$ I-sector firms in region b, with $N_a(t) + N_b(t) = N(t)$.

As firm i_k is the monopoly supplier of its unique variety of intermediate input, it maximizes profits at time t by setting marginal revenue equal to marginal cost. That firm i_k faces a constant elasticity demand curve for its variety of intermediate input implies that its marginal revenue function is a constant fraction of its inverse demand function. In particular, from (1.22), $p_{i_k k}(t) = \Delta_k(t)^{1-\alpha} q_{i_k}(t)^{\alpha-1}$ is firm i_k's

[27] The assumption that each I-sector firm is a monopoly supplier of a single variety of intermediate input could be relaxed so that a single I-sector firm is the monopoly supplier of several varieties of intermediate inputs; however, given that there are no economies of scope in I-sector production, this change would not affect the analysis. Conversely, if this assumption were relaxed so that multiple firms could produce the same variety of intermediate input (i.e., so that there is no legal institution protecting each I-sector firm's monopoly position), there would stop being an incentive for firms to invest in developing new varieties of intermediate inputs, which would remove the driver of economic growth from the model.

inverse demand function at time t and $MR_{i_k}(t) = \alpha \left(\Delta_k (t)^{1-\alpha} q_{i_k}(t)^{\alpha-1}\right) = \alpha p_{i_k k}(t)$ is its marginal revenue function at time t. As the marginal cost of production for firm i_k is equal to one for all t, the condition that marginal revenue equals marginal cost implies that firm i_k maximizes profits by setting its price at a constant mark-up over its marginal costs, $p_{i_k k}(t) = \alpha^{-1}$, regardless of the level of demand it faces.

I-Sector Innovation

Firms entering the I-sector develop new varieties of intermediate inputs using a constant returns to scale production technology where labor is the only input. The productivity of labor in developing new varieties of intermediate inputs – i.e., in I-sector "innovation" – in region k at time t is

$$\Omega_k (t) \equiv N (t) \left[\frac{L_{kI}(t)}{L_I (t)} + \eta \left(1 - \frac{L_{kI}(t)}{L_I (t)}\right)\right], \tag{1.27}$$

where $\eta \in [0, 1]$ is a parameter proportional to the ease of transferring knowledge between workers in I-sector innovation in region k and region l. Hence, each firm entering the I-sector in region k at time t requires $(\Omega_k (t))^{-1}$ units of labor to produce its unique variety of intermediate input. It is assumed that there is a large number of firms entering the I-sector at any time t, so that each firm's location choice (region k or l) has a negligible impact on $\Omega_k (t)$. As such, firms entering the I-sector do not take into account the impact of their location choice on $\Omega_k (t)$.

(1.27) implies that there are positive knowledge spillovers in I-sector innovation, i.e., that developing a new variety of intermediate input becomes easier as the economy becomes more technologically advanced (as $N (t)$ increases). (1.27) also implies that when $\eta < 1$, I-sector knowledge spillovers are a partially local public good so that I-sector innovation in region k is more productive when a larger share of total I-sector employment is located in region k.

The number of new varieties of intermediate input developed in region k at time

t is

$$Y_{kI}(t) = \dot{N}_k(t) = \Omega_k(t) L_{kI}(t), \tag{1.28}$$

where $L_{kI}(t) \geq 0$ is the total quantity of labor employed in developing new varieties of intermediate inputs in region k at time t. From (1.27) and (1.28), the total number of new varieties of intermediate inputs developed at time t is

$$Y_I(t) = \dot{N}(t) = \dot{N}_a(t) + \dot{N}_b(t) = N(t) L_I(t) \left[1 - 2(1-\eta)\frac{L_{aI}(t)}{L_I(t)}\left(1 - \frac{L_{aI}(t)}{L_I(t)}\right)\right]. \tag{1.29}$$

Examining (1.29), when $\eta < 1$, $\dot{N}(t)$ is increasing as I-sector innovation becomes concentrated in either region a or region b.

Entrepreneurs will develop new varieties of intermediate inputs and enter the I-sector until the value of an I-sector firm at time t, $V(t)$, is equal to the cost of developing a new variety. (As I-sector firms are free to move between regions to maximize profits, $V(t)$ does not depend on the region in which the firm is located.) Using the expression from (1.28), this means that the number of new varieties of intermediate inputs developed in region k at time t, $\dot{N}_k(t)$, will increase until

$$V(t)\dot{N}_k(t) - w_k(t) L_{kI}(t) = V(t)\Omega_k(t) L_{kI}(t) - w_k(t) L_{kI}(t) = 0. \tag{1.30}$$

(1.30) must hold, because if $V(t)$ were greater than the cost of development of a new variety of intermediate input at time t, there would be more innovation, which would drive down $V(t)$ until (1.30) held. Conversely, if the cost of development of a new variety of intermediate input at time t were greater than $V(t)$, there would be no innovation at time t. (1.30) implies that the value of an I-sector firm at time t is $V(t) = w_k(t)/\Omega_k(t)$ for $k = a, b$.

Market-Clearing

The market-clearing condition in the labor market at time t is that total labor in M- and S-sector production and in I-sector innovation is equal to the total labor supply in the economy. The market-clearing condition in the labor market at time t can be written as

$$L_{aM}(t) + L_{bM}(t) + L_{aS}(t) + L_{bS}(t) + L_{aI}(t) + L_{bI}(t) = L(t), \tag{1.31}$$

where $L(t) = \sum_{j=1}^{H} l_j(t)$ is the total labor supply in the economy at time t.

The market-clearing condition for M-sector output at time t is that the total consumption of M-sector output by households and total use of M-sector output by I-sector firms is equal to the total supply of M-sector output in the economy. The market-clearing condition for M-sector output at time t can be written as

$$\sum_{j=1}^{H} Y_{jM}(t) = Y_{aM}(t) + Y_{bM}(t) - \sum_{i_a \in \mathbb{M}_a(t)} Y_{i_a}(t) - \sum_{i_b \in \mathbb{M}_b(t)} Y_{i_b}(t), \tag{1.32}$$

where $\sum_{j=1}^{H} Y_{jM}(t)$ is total demand for M-sector output at time t and $\sum_{i_k \in \mathbb{M}_k(t)} Y_{i_k}(t)$ is the total quantity of M-sector output used by I-sector firms in region k at time t.

The market-clearing condition for S-sector output at time t is that the total consumption of S-sector output by households is equal to the total supply of S-sector output in the economy. The market-clearing condition for S-sector output at time t can be written as

$$\sum_{j=1}^{H} Y_{jS}(t) = Y_{aS}(t) + Y_{bS}(t), \tag{1.33}$$

where $\sum_{j=1}^{H} Y_{jS}(t)$ is total demand for S-sector output at time t.

Finally, the market-clearing condition for I-sector firm i_k's variety of intermediate input at time t is that its total use by M-sector firms is equal to its total supply in the economy. The market-clearing condition for I-sector firm i_k's variety of intermediate input at time t can be written as

$$Y_{i_k}(t) = q_{i_k k}(t) + \Gamma q_{i_k l}(t), \tag{1.34}$$

where $q_{i_k k}(t) \geq 0$ is the demand for firm i_k's variety of intermediate input from M-sector firms in region k and $q_{i_k l}(t) \geq 0$ is the demand from M-sector firms in region l.

1.5 Static Equilibrium

The characterization of the equilibrium for the economy is divided into two parts: the static equilibrium and the dynamic equilibrium. The static equilibrium, which is described and analyzed in this section, defines an allocation of factors of production and a set of prices such that firms maximize profits, works maximize their wage, households maximize utility from consumption, and the markets for labor and for M-, S-, and I-sector output clear at time t and for given levels of technology, $N(t)$, societal expenditure, $E(t) = \sum_{j=1}^{H} E_j(t)$, and population, $L(t)$. A formal definition for the static equilibrium is presented below. The dynamic equilibrium, which is described and analyzed in Chapter 1.6, is a series of static equilibria for the economy over time where the evolution of $N(t)$ and $E(t)$ is consistent with the first-order conditions for each household's intertemporal program.

Definition 1 *The static equilibrium is a set of prices – the relative price of S-sector output, $p(t)$, the wage in regions a and b, $w_a(t)$ and $w_b(t)$, the rent to the fixed factor in regions a and b, $w_a^F(t)$ and $w_b^F(t)$, prices of intermediate inputs, $\{p_{i_a a}(t)\}_{i_a \in \mathbb{M}_a(t)}$, and $\{p_{i_b b}(t)\}_{i_b \in \mathbb{M}_b(t)}$, and the value of I-sector firms, $V(t)$ – labor allocations, $L_{aM}(t)$, $L_{bM}(t)$, $L_{aS}(t)$, $L_{bS}(t)$, $L_{aI}(t)$, and $L_{bI}(t)$, and levels of production $Y_{aM}(t)$, $Y_{bM}(t)$, $Y_{aS}(t)$, $Y_{bS}(t)$, $Y_{aI}(t)$, $Y_{bI}(t)$, $\{Y_{i_a}(t)\}_{i_a \in \mathbb{M}_a(t)}$, and $\{Y_{i_b}(t)\}_{i_b \in \mathbb{M}_b(t)}$, at a given time t and for given levels of technology, $N(t)$, societal expenditure, $E(t)$, and population, $L(t)$, such that each household maximizes their utility from consumption, firms in the M-, I-, and S-sectors maximize profits, workers maximize their wage, and the market-clearing conditions hold for the labor market, for M- and S-sector output,* *and for each I-sector firm's variety of intermediate input.*

The static equilibrium in Definition 1 defines a Walrasian equilibrium for the economy at time t. In a Walrasian equilibrium for the economy: (i) Each firm in the M-, I-, and S-sectors maximizes profits given the equilibrium prices; (ii) Each household maximizes their utility from consumption given equilibrium prices and total household expenditure; and (iii) All markets clear at equilibrium prices (i.e., all households and firms are able to achieved their desired trades at the equilibrium prices). A formal definition of a Walrasian equilibrium (with formal descriptions of the production sets, consumption sets, and preference relations in the economy) and a proof that a Walrasian equilibrium exists for the economy in this chapter are both given in Appendix E.

The static equilibrium for the economy is characterized in two stages. In the first stage, expressions for prices, sectoral employment, and sectoral output when the economy is in static equilibrium are derived. These expressions are summarized in Proposition 1. In the second-stage, the expressions for prices, sectoral employment, and sectoral output from the first stage are used to define the distributions of factors of production across regions a and b at time t such that workers cannot increase their wage and firms cannot increase their profit by moving between regions, i.e., such that the economy is in a *spatial equilibrium.* The static equilibrium is analyzed via comparative static analysis in Chapter 1.5.3.

1.5.1 Static Equilibrium: Prices, Sectoral Employment, and Sectoral Output

Proposition 1 *When the economy is in static equilibrium at time t, the prices for a given I-sector firm in region k's intermediate input in regions k and l are ($k, l = a, b$,*

$k \neq l$)

$$p_{i_k k}(t) = p_{kk}(t) = 1/\alpha \text{ and } p_{i_k l}(t) = p_{kl}(t) = \Gamma/\alpha \text{ for all } i_k \in \mathbb{M}_k(t), \tag{1.35}$$

the price index that gives the minimum cost for M-sector firms in region k of purchasing a unit of $I_k(t)$ at time t is

$$P_k(t) = \alpha^{-1}\left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t)\right)^{\frac{\alpha-1}{\alpha}}, \tag{1.36}$$

the wage in region k is

$$w_k(t) = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}\left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t)\right), \tag{1.37}$$

the rental rate for the fixed factor in region k is

$$w_k^F(t) = (1-\beta)\,p(t)\left(\frac{\beta p(t)}{w_k(t)}\right)^{\frac{\beta}{1-\beta}}, \tag{1.38}$$

the relative price of S-sector output, $p(t)$, is defined implicitly in terms of $E(t)$, $w_a(t)$, and $w_b(t)$ by the S-sector market-clearing condition,

$$\frac{E(t)}{p(t) + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p(t)^{\varepsilon}} = \left(\frac{\beta p(t)}{w_a(t)}\right)^{\frac{\beta}{1-\beta}} F_a + \left(\frac{\beta p(t)}{w_b(t)}\right)^{\frac{\beta}{1-\beta}} F_b, \tag{1.39}$$

and the value of an I-sector firm in region k is

$$V(t) = \frac{w_k(t)}{\Omega_k(t)} = \frac{(1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}\left[\frac{N_k(t)}{N(t)} + \Gamma^{\frac{\alpha}{a-1}}\left(1 - \frac{N_k(t)}{N(t)}\right)\right]}{\left[\frac{L_{kI}(t)}{L_I(t)} + \eta\left(1 - \frac{L_{kI}(t)}{L_I(t)}\right)\right]}; \tag{1.40}$$

All I-sector firms in region k produce

$$Y_{i_k}(t) = Y_{ki}(t) = \alpha^{\frac{2}{1-\alpha}}\left(L_{kM}(t) + \Gamma^{\frac{\alpha}{a-1}} L_{lM}(t)\right) \text{ for all } i_k \in \mathbb{M}_k(t), \tag{1.41}$$

intermediate inputs at time t, and make profits equal to

$$\pi_{i_k}(t) = \pi_{ki}(t) = (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}}\left(L_{kM}(t) + \Gamma^{\frac{\alpha}{a-1}} L_{lM}(t)\right) \text{ for all } i_k \in \mathbb{M}_k(t); \tag{1.42}$$

Employment in M-, S-, and I-sectors is

$$L_M(t) = \frac{1}{(1+\alpha)} \frac{1}{w_a(t)\left(\frac{L_{aM}(t)}{L_M(t)}\right) + w_b(t)\left(1 - \frac{L_{aM}(t)}{L_M(t)}\right)} \frac{E(t)}{1 + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p(t)^{1-\varepsilon}}, \tag{1.43}$$

$$L_S(t) = L_{aS}(t) + L_{bS}(t) = \left(\frac{\beta p(t)}{w_a(t)}\right)^{\frac{1}{1-\beta}} F_a + \left(\frac{\beta p_a(t)}{w^{SC}(t)}\right)^{\frac{1}{1-\beta}} F_b, \tag{1.44}$$

and

$$L_I^{SC}(t) = L(t) - L_M^{SC}(t) - L_S^{SC}(t); \tag{1.45}$$

And output in M- and S-sectors and in I-sector innovation is

$$Y_{kM}(t) = \left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t)\right) q_{kk}(t)^{\alpha} L_{kM}(t)^{1-\alpha} = \alpha^{\frac{2\alpha}{1-\alpha}} L_{kM}(t) \left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t)\right), \tag{1.46}$$

$$Y_S(t) = Y_{aS}(t) + Y_{bS}(t) = \left(\frac{\beta p(t)}{w_a(t)}\right)^{\frac{1}{1-\beta}} F_a + \left(\frac{\beta p(t)}{w_b(t)}\right)^{\frac{1}{1-\beta}} F_b, \tag{1.47}$$

and

$$Y_I(t) = \dot{N}(t) = \dot{N}_a(t) + \dot{N}_b(t) = N(t) L_I(t) \left[1 - 2(1-\eta)\frac{L_{aI}(t)}{L_I(t)}\left(1 - \frac{L_{aI}(t)}{L_I(t)}\right)\right]. \tag{1.48}$$

Proof. See Appendix F. ■

In Proposition 1, the expressions for prices, sectoral employment, and sectoral output when the economy is in static equilibrium are functions of $N(t)$, $E(t)$, and $L(t)$, of the spatial distribution of labor in the M-, S-, and I-sectors, and of the spatial distribution of I-sector firms. In the next section, these expressions are used in two theorems to characterize the spatial distribution of labor in the M-, S-, and I-sectors and the spatial distribution of I-sector firms so that no worker can increase their wage and no I-sector firm can increase its profits by moving between regions. Together, Proposition 1 and the two theorems presented in the next section characterize the static equilibrium for the economy.

Several of the expressions in Proposition 1 reflect the influence of agglomeration externalities on productivity in M-sector production and I-sector innovation, and on the profits of I-sector firms. For example, (1.37) implies that when transportation costs for intermediate inputs are positive ($\Gamma > 1$), the zero-profit wage in the M-sector in region k, $w_k(t)$, is increasing with the total number of I-sector firms in region k,

$N_k(t)$. Similarly, (1.42) implies when $\Gamma > 1$ that that the profit of each I-sector firm in region k is increasing in the level of employment in the M-sector in region k, $L_{kM}(t)$. These two results reflect the influence of agglomeration externalities related to forward and backward linkages between the M- and I-sectors (see Theorem 1 for further discussion).

In addition, from (1.40), the value of an I-sector firm, $V(t)$, is increasing as I-sector firms become concentrated in either region a or b when $\Gamma > 1$, which reflects the influence of agglomeration externalities on I-sector profits, and, hence, on $V(t)$. In addition, when $\eta < 1$, $V(t)$ is decreasing as the I-sector innovation become concentrated in either region a or b.[28] This latter result reflects the fact that when I-sector innovation becomes concentrated in a region, knowledge spillovers reduce the costs of developing a new variety of intermediate input. This reduction in development cost leads to more new varieties of intermediate input being developed at time t, which lowers $V(t)$. The next section will demonstrate the role of these agglomeration externalities in determining the spatial distribution of labor in the M-, S-, and I-sectors and in the spatial distribution of I-sector firms when the economy is in static equilibrium.

The expression for M-sector production in region k at time t in (1.46) from Proposition 1 illustrates how an increase in the number of varieties of intermediate input – i.e., technological advance – leads to Hicks-neutral technical change in the M-sector. To see why this is the case, (1.46) can be re-expressed as

$$Y_{kM}(t) = \left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t)\right)^{1-\alpha} \left[q_{kk}(t)\left(N_k(t) + \Gamma^{\frac{1}{a-1}} N_k(t)\right)\right]^{\alpha} L_{kM}(t)^{1-\alpha}, \tag{1.49}$$

[28] Recall that $\eta \in [0, 1]$ is a parameter proportional to the ease of transferring knowledge between workers in I-sector innovation in region a and region b, so that when $\eta < 1$ I-sector knowledge is a partially local public good.

where $N_k(t)\, q_{kk}(t) + N_l(t)\, q_{lk}(t) = q_{kk}(t)\left(N_k(t) + \Gamma^{\frac{1}{a-1}} N_l(t)\right)$ is the total quantity of intermediate inputs used in the M-sector in region k at time t. (1.49) implies that for a given $L_{kM}(t)$ and a given total quantity of intermediate inputs, $q_{kk}(t)\left(N_k(t) + \Gamma^{\frac{1}{a-1}} N_l(t)\right)$, $Y_{kM}(t)$ increases in accordance with the term $\left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t)\right)^{1-\alpha}$ This reflects the benefit of spreading a given number of intermediate inputs over a wider range of $N(t)$. These benefits exist because there are diminishing returns for each $q_{ik}(t)$ individually. Moreover, (1.49) implies that increases in $N(t)$ lead to Hicks-neutral technical change in the M-sector, as the marginal rate of substitution between labor and the total quantity of intermediate inputs does not change with changes in $N(t)$.

1.5.2 Static Equilibrium: Regional Factor Distributions

This section presents two theorems that together characterize the distributions of M-, S-, and I-sector labor and I-sector firms across regions a and b at time t such that workers cannot increase their wage and firms cannot increase their profit by moving between regions. The first theorem uses the expressions for the wage in region k at time t (1.37) and I-sector profits in region k at time t (1.42) from Proposition 1 to characterize the distributions of M-sector labor and I-sector firms across regions a and b such no worker in the M-sector can increase their wage and no firm in the I-sector can increase its profits by moving between regions. The second theorem uses the expression for the value of an I-sector firm in region k at time t (1.40) to characterize the distribution of workers in I-sector innovation across regions a and b such no worker in I-sector innovation can increase his wage by moving between regions. Together with Proposition 1, these two theorems complete the characterization of the static equilibrium.

Before presenting the two theorems, it is necessary to introduce some new notation. To begin, let $\lambda_P(t) \equiv L_{aP}(t)/L_P(t)$, $P = S, M, I$, be a spatial distribution of P-

sector labor, and let $N_a(t)/N(t)$ be a spatial distribution of I-sector firms. $\lambda_P^*(t) \in [0,1]$ is a spatial equilibrium in the P-sector at time t if no worker in the P-sector can increase their wage at time t by moving between regions. Similarly, $N_a^*(t)/N(t) \in [0,1]$ is a spatial equilibrium for I-sector firms at time t if no I-sector firm can increase its profits at time t by moving between regions.

Next, a spatial equilibrium of P-sector labor, $\lambda_P^*(t)$, is stable if there exists a bounded, convex set Λ_P containing $\lambda_P^*(t)$ defined by $\Lambda_P = (\lambda_P^*(t) - \rho, \lambda_P^*(t) + \rho) \cap [0,1]$, $\rho > 0$, such that from any point $\lambda_P(t) \in \Lambda_P$ workers in the P-sector will increase his wage by moving between regions until $\lambda_P^*(t)$ is restored as the spatial distribution of P-sector labor. Similarly, a spatial equilibrium of I-sector firms, $N_a^*(t)/N(t)$, is stable if there exists a bounded, convex set Λ_N containing $N_a^*(t)/N(t)$ defined by $\Lambda_N = (N_a^*(t)/N(t) - \rho, N_a^*(t)/N(t) + \rho) \cap [0,1]$, $\rho > 0$, such that from any point $N_a(t)/N(t) \in \Lambda_N$ firms in the I-sector will increase their profits by moving between regions until $N_a^*(t)/N(t)$ is restored as the spatial distribution of I-sector firms.

Theorem 1 *When there are positive transportation costs for intermediate inputs,* $\Gamma > 1$, *there are three spatial equilibria for M-sector labor and I-sector firms:*

1. $\lambda_M(t) = 1$ $(L_{aM}(t) = L_M(t), L_{bM}(t) = 0)$ *and* $\frac{N_a(t)}{N(t)} = 1$ $(N_a(t) = N(t), N_b(t) = 0)$;
2. $\lambda_M(t) = 0$ $(L_{aM}(t) = 0, L_{bM}(t) = L_M(t))$ *and* $\frac{N_a(t)}{N(t)} = 0$ $(N_a(t) = 0, N_b(t) = N(t))$;
3. $\lambda_M(t) = 0.5$ $\left(L_{aM}(t) = L_{bM}(t) = \frac{1}{2}L_M(t)\right)$ *and* $\frac{N_a(t)}{N(t)} = 0.5$ $\left(N_a(t) = N_b(t) = \frac{1}{2}N(t)\right)$.

The first two spatial equilibria are stable; the third, unstable. When there are no transportation costs for intermediate inputs, $\Gamma = 1$, *any spatial distribution of M- and I-sector firms is an unstable spatial equilibrium.*

Proof. See Appendix F. ■

The intuition for Theorem 1 follows from how agglomeration externalities influence production in the M- and I-sectors. The combination of increasing returns to scale in the I-sector and transportation costs for intermediate inputs provides an incentive for the demanders and suppliers of the intermediate inputs to locate in the same region. M-sector firms locate in the region with the larger concentration of I-sector firms because doing so lowers their cost of intermediate inputs through reduced transportation costs. This allows M-sector firms to use more intermediate inputs, which raises M-sector labor productivity and the zero-profit wage in the region, i.e.,

$$w_k(t) = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}} \left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t) \right) > (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}} \left(\Gamma^{\frac{\alpha}{a-1}} N_k(t) + N_l(t) \right) = w_l(t) \tag{1.50}$$

when $N_k(t) > N_l(t)$. This, in turn, causes M-sector production in the region with a smaller number of I-sector firms to not be an equilibrium, and leads to the spatial concentration of the M-sector in the region with a larger concentration of I-sector firms.

I-sector firms locate in the region with the large concentration of M-sector firms because doing so lowers the cost of their variety of intermediate inputs to M-sector firms, which increases the demand for their variety and, hence, their profits. For example, when $L_{kM}(t) > L_{lM}(t)$,

$$\pi_{ki}(t) = (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}} \left(L_{kM}(t) + \Gamma^{\frac{\alpha}{a-1}} L_{lM}(t) \right) > (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}} \left(\Gamma^{\frac{\alpha}{a-1}} L_{kM}(t) + L_{lM}(t) \right) = \pi_{li}(t)\,. \tag{1.51}$$

This will cause I-sector firms to concentrate themselves in the region with the larger concentration of M-sector firms.

Theorem 1 also implies that when $\Gamma > 1$, M- and I-sector production will only take place in both regions if $N_a(t) = N_b(t)$ and $L_{aM}(t) = L_{bM}(t)$. When there are no transportation costs ($\Gamma = 1$), $w_a(t) = w_b(t)$ and $\pi_{ai}(t) = \pi_{bi}(t)$ always hold, and there are no incentives for spatial concentration in the M- and I-sectors. In this case, all spatial distributions of M- and I-sector firms are unstable spatial equilibrium.

From (1.26), the spatial distribution of labor in the S-sector is

$$\lambda_S(t) \equiv \frac{L_{aS}(t)}{L_S(t)} = \frac{F_a}{F_a + F_b\left(\frac{w_a(t)}{w_b(t)}\right)^{\frac{1}{1-\beta}}}. \tag{1.52}$$

Theorem 1 demonstrates that when the M-sector and I-sector production are in spatial equilibrium, the wage is equalized across regions (e.g., $w(t) = w_a(t) = w_b(t)$), so that (1.52) simplifies to $\lambda_S(t) = F_a/(F_a + F_b)$. That is, when M- and I-sector production are in spatial equilibria, the spatial distribution of labor in the S-sector is determined solely by the distribution of the fixed factor across regions a and b.

Theorem 2 *When I-sector knowledge is a partially local public good, $\eta < 1$, there are three spatial equilibria for labor in I-sector innovation:*

1. $\lambda_I(t) = 1 \; (L_{aI}(t) = L_I(t), L_{bI}(t) = 0)$;

2. $\lambda_I(t) = 0 \; (L_{aI}(t) = 0, L_{bI}(t) = L_I(t))$;

3. $\lambda_I(t) = 0.5 \; \left(L_{aI}(t) = L_{bI}(t) = \frac{1}{2}L_I(t)\right)$.

The first two spatial equilibria are stable; the third, unstable. When I-sector knowledge is a global public good, $\eta = 1$, any spatial distribution of labor in I-sector innovation is an unstable spatial equilibrium.

Proof. See Appendix F. ■

The intuition for Theorem 2 follows from how knowledge spillover externalities influence productivity in I-sector innovation. Labor employed in I-sector innovation benefits from the exchange of knowledge. When I-sector knowledge is a partially local public good, $\eta < 1$, I-sector workers in the region with the larger concentration of labor in I-sector innovation are made more productive from the exchange of knowledge than are I-sector workers in the region with the smaller concentration of labor in I-sector innovation. As free-entry drives profits from I-sector innovation to zero, interregional differences in productivity in I-sector innovation cannot be sustained in equilibrium. This means that I-sector innovation is either equally split between the two regions or concentrated in the region where I-sector innovation is most productive.

The analysis in the remainder of this chapter focuses on two *spatial configurations* (SC) for the economy, where a spatial configuration for the economy at time t is defined as $\{\lambda_M^*(t), N_a^*(t)/N(t), \lambda_I^*(t), \lambda_S^*(t)\}$.[29] The two spatial configurations considered in the remainder of this chapter are: *full agglomeration* (FA), where M- and I-sector labor and I-sector innovation are concentrated in one region (region a), i.e., $\{\lambda_M^*(t), N_a^*(t)/N(t), \lambda_I^*(t), \lambda_S^*(t)\} = \{1, 1, 1, F_a/(F_a+F_b)\}$, and *full dispersion* (FD), where the M- and I-sectors are evenly dispersed across the two regions, i.e., $\{\lambda_M^*(t), N_a^*(t)/N(t), \lambda_I^*(t), \lambda_S^*(t)\} = \left\{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, F_a/(F_a+F_b)\right\}$.[30]These definitions imply that region a and b's populations are

$$L_a^{FA}(t) = L_M^{FA}(t) + L_{aS}^{FA}(t) + L_I^{FA}(t) \text{ and } L_b^{FA}(t) = L_{bS}^{FA}(t) \tag{1.53}$$

under full agglomeration, and

$$L_a^{FD}(t) = \frac{1}{2}L_M^{FD}(t) + L_{aS}^{FD}(t) + \frac{1}{2}L_I^{FD}(t) \text{ and } L_b^{FD}(t) = \frac{1}{2}L_M^{FD}(t) + L_{bS}^{FD}(t) + \frac{1}{2}L_I^{FD}(t) \tag{1.54}$$

under full dispersion. Focusing on full agglomeration and full dispersion allows for four relevant cases of regional heterogeneity in economic characteristics to be considered: 1. Regions differ in their endowment of the fixed factor and in the agglomeration externalities associated with the concentration of the M- and I-sectors (full agglomeration with $F_a \neq F_b$); 2. Regions differ in the fixed factor but not in agglomeration externalities (full dispersion with $F_a \neq F_b$); 3. Regions differ in agglomeration exter-

[29]Theorems 1 and 2 give five spatial configurations (nine when the symmetry between region a and region b is ignored): $\nu(t) = \{1, 1, 1, F_a/(F_a+F_b)\}$, $\nu(t) = \left\{1, 1, \frac{1}{2}, F_a/(F_a+F_b)\right\}$, $\nu(t) = \{1, 1, 0, F_a/(F_a+F_b)\}$, $\nu(t) = \left\{\frac{1}{2}, \frac{1}{2}, 1, F_a/(F_a+F_b)\right\}$, and $\nu(t) = \left\{\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, F_a/(F_a+F_b)\right\}$.

[30]Recall that $\lambda_S(t) = F_a/(F_a+F_b)$ is the unique spatial equilibrium for S-sector labor.

nalities and not the fixed factor (full agglomeration with $F_a = F_b$); and Regions are identical (full dispersion with $F_a = F_b$).

An immediate corollary of Theorems 1 and 2 and (1.40) is that when the economy is in static equilibrium, the value of an I-sector firm, $V(t)$, is constant for all t. That $V(t)$ is constant over time follows from the fact that $N(t)$ determines productivity in both M-sector production (and, hence, the wage at time t) and productivity in I-sector innovation (1.27). As such, when the economy is in static equilibrium, increases in the wage that occur as the economy evolves are matched by increases in labor productivity in I-sector innovation, leaving the marginal cost of developing a new variety of intermediate input, $w(t)/\Omega_k(t)$, $k = a, b$, and, hence, $V(t)$, constant.

That $V(t)$ is constant over time when the economy is in static equilibrium allows for a closed-form expression for the interest rate to be derived. To see how this is the case, recall from Theorem 1 that when the economy is in static equilibrium, the profits from I-sector firms are equal in both regions (i.e., $\pi_{ai}(t) = \pi_{bi}(t) = \pi_i(t)$). This means that in static equilibrium, the value of an I-sector firm at time t can be expressed as

$$V(t) = \int_t^\infty \pi_i(v) e^{-R(v)} dv \quad \text{where} \quad R(v) = \int_t^v r(\tau) d\tau. \tag{1.55}$$

Differentiating (1.55) with respect to t and, using the fact that $V(t)$ is constant over time, setting the resulting expression equal to zero yields

$$\frac{dV(t)}{dt} = -\pi_i(t) + r(t)V(t) = 0 \Rightarrow \; r(t) = \frac{\pi_i(t)}{V(t)}. \tag{1.56}$$

Substituting the expressions for $\pi_i(t)$ and $V(t)$ from (1.42) and (1.40) gives the following expression for the interest rate,

$$r(t) = \frac{\left[\frac{L_{kM}(t)}{L_M(t)} + \Gamma^{\frac{\alpha}{a-1}}\left(1 - \frac{L_{kM}(t)}{L_M(t)}\right)\right]\left[\frac{L_{kI}(t)}{L_I(t)} + \eta\left(1 - \frac{L_{kI}(t)}{L_I(t)}\right)\right]}{\left[\frac{N_k(t)}{N(t)} + \Gamma^{\frac{\alpha}{a-1}}\left(1 - \frac{N_k(t)}{N(t)}\right)\right]} \alpha L_M(t), \tag{1.57}$$

$k = a, b$. From (1.57), $r(t)$ is increasing with total household expenditure at time t because, from (1.43), $L_M(t)$ is increasing with total household expenditure in the economy at time t. The intuition for this result is that higher household expenditure at time t lowers the supply of household savings available to entrepreneurs in the I-sector, leading to higher $r(t)$. (1.57) also implies that $r(t)$ is larger when labor in I-sector innovation is concentrated in one region. This is because I-sector innovation is more productive when is it concentrated in a single region. Higher productivity in I-sector innovation increases the demand for household savings by entrepreneurs in the I-sector, which leads to higher $r(t)$.

1.5.3 Comparative Static Analysis

This section analyzes the spatial consequences of supply-side NBG when the economy is in a static equilibrium. The comparative static results presented in this section focus on the elasticity of several of the key variables in the model with respect to societal expenditure, $E(t)$, technology, $N(t)$, and population, $L(t)$, all of which are fixed in the static equilibrium. Comparative static results for the case of demand-driven NBG are presented in Chapter 1.7.1 and for the case of NBG in an open economy in Chapter 1.7.2. Following the arguments presented in Chapter 1.5.2, results are presented for two spatial configurations of the economy: full agglomeration, where the M- and I-sector labor and I-sector innovation are concentrated in region a, and full dispersion, where the M- and I-sectors are evenly dispersed across the two regions.

Proposition 2 *In equilibrium, 1.* $\frac{d\ln p^{SC}(t)}{d\ln N(t)} > 0$*; 2.* $\frac{d\ln p(t)}{d\ln E(t)} > 0$*; and 3.* $\frac{d\ln p(t)}{d\ln L(t)} = 0$*, where* $SC = FA,\ FD$*.*

Proof. See Appendix G. ■

There are three implications of Proposition 2 that are important for understanding the other comparative statics presented in this section. First, part 1 of Proposition 2 establishes that the elasticity of $p(t)$ with respect to $N(t)$ is always positive. This result follows from the fact that technological advance – i.e., increases in $N(t)$ –

increases productivity in the M-sector but not the S-sector. The increase in productivity in the M-sector reduces the cost of producing a given amount of M-sector output, thereby increasing the relative price of S-sector output.

Second, part 2 of Proposition 2 establishes that the elasticity of $p(t)$ with respect to $E(t)$ is also always positive. This result follows from the fact that an increase in $E(t)$ increases the demand for both M- and S-sector output (see (1.5)). Meeting this increase in demand increases the marginal cost of production in the S-sector, where there is decreasing returns to scale, but not in the M-sector, where, from (1.18), there is a constant marginal cost that is invariant to the level of production. Given that zero profits must hold in both sectors, $p(t)$ must increase in response to an increase in S-sector output to compensate for the increase in the marginal cost of production.

Third, from part 3 of Proposition 2, the relative price of S-sector output, $p(t)$, is invariant to change in population, $L(t)$. The intuition for this result, and how it relates to the assumption of homothetic household preferences, is presented in the discussion of Proposition 3 below.

Proposition 3 *In equilibrium,*

1. $\frac{d\ln L_M^{SC}(t)}{d\ln N(t)} < 0$, $\frac{d\ln L_S^{SC}(t)}{d\ln N(t)} < 0$, *and* $\frac{d\ln L_I^{SC}(t)}{d\ln N(t)} > 0$;
2. $\frac{d\ln L_M^{SC}(t)}{d\ln E(t)} > 0$, $\frac{d\ln L_S^{SC}(t)}{d\ln E(t)} > 0$, *and* $\frac{d\ln L_I^{SC}(t)}{d\ln E(t)} < 0$;
3. $\frac{d\ln L_M^{SC}(t)}{d\ln L(t)} = \frac{d\ln L_S^{SC}(t)}{d\ln L(t)} = 0$, *and* $\frac{d\ln L_I^{SC}(t)}{d\ln L(t)} = 1 > 0$;

where $SC = FA, FD$.

Proof. See Appendix G. ■

Part 1 of Proposition 3 implies that an increase in technology, $N(t)$, will reduce employment in both the M- and S-sectors – i.e., will reduce employment dedicated to producing consumption goods – and increase employment in I-sector innovation, which increases the rate of development of new varieties of intermediate inputs (i.e., the rate of technical progress) in the economy. The intuition for part 1 of Proposition 3 follows from the fact that an increase in $N(t)$ increases productivity in the M- and I-sector sectors (see (1.27) and (1.49)), but not in the S-sector. Proposition 2 establishes that by increasing productivity in the M-sector but not the S-sector, an increase in $N(t)$ increases the relative price of S-sector output, $p(t)$. For fixed $E(t)$, this increase in $p(t)$ reduces the demand for S-sector output, (1.5), and thereby reduces S-sector employment, $L_S^{SC}(t)$. Furthermore, from (1.5), an increase in $p(t)$ also decrease the demand of M-sector output, and, hence, employment, when $\varepsilon < 1$ (i.e., when M- and S-sector output are gross complements). Meanwhile, while an increase in $p(t)$ will increase the demand for M-sector output when $\varepsilon > 1$, the increase in $N(t)$ reduces the amount of M-sector labor required to meet a given level of M-sector demand. Proposition 3 establishes that this latter effect dominates, and $L_M^{SC}(t)$ decreases with an increase in $N(t)$ even when $\varepsilon > 1$.

It follows from the market-clearing condition for labor, (1.31), that $L_I^{SC}(t)$ must increase with $N(t)$ because both $L_M^{SC}(t)$ and $L_S^{SC}(t)$ decrease with $N(t)$. The intuition for this result is that an increase in $N(t)$ increases the productivity of labor in the I-sector, which attracts labor to the I-sector, and, by increasing productivity in the M-sector, reduces the amount of labor required to satisfy consumption of M- and S-sector output for a given level of $E(t)$. Given this result, a noteworthy implication of part 1 of Proposition 3 is that, all else equal, a more technically advanced economy will deploy a larger portion of its labor force towards innovation and have a higher rate of technological progress.

Part 2 of Proposition 3 establishes that both $L_M^{SC}(t)$ and $L_S^{SC}(t)$ increase, and $L_I^{SC}(t)$ decreases, with an increase in $E(t)$. The intuition for Proposition 3 is that, from (1.5), an increase in $E(t)$ increases the demand for both M- and S-sector output, which leads to higher levels of employment in both sectors. The increase in M- and

S-sector employment necessarily decreases employment in I-sector innovation. In this way, part 2 of Proposition 3 illustrates the trade-off in the model between expenditure on current consumption, $E(t)$, and investment in I-sector innovation, which is the source of productivity growth.

Part 3 of Proposition 3 demonstrates that changes in population, $L(t)$, do not influence $L_M^{SC}(t)$ or $L_S^{SC}(t)$. Indeed, both $L_M^{SC}(t)$ and $L_S^{SC}(t)$ are determined in the static equilibrium by $N(t)$ and $E(t)$ alone. In Section 1.7.1, it is shown that when household preferences are non-homothetic, population growth does indeed change output and employment in the M- and S-sectors. This is because when household preferences are non-homothetic, output demand in the M- and S-sectors, and, hence, employment in either sector, is determined by per capita expenditure, which is influenced by total population. In the baseline model of supply-side NBG analyzed here, household preferences are homothetic, and it is total societal expenditure, rather than per capita expenditure, that influences output demand in the M- and S-sectors. Further, part 3 of Proposition 3 demonstrates that, all else equal, the rate of technological progress in the economy is increasing in the size of the population (i.e., is increasing in $L(t)$).

Taken as a whole, Proposition 3 demonstrates that $N(t)$, $E(t)$, and $L(t)$ have countervailing influences employment in the M-, S-, and I-sectors. As such, analyzing the rates of employment growth in each sector (and, hence, the direction of change in sectoral employment shares) entails analyzing changes in $N(t)$, $E(t)$, and $L(t)$ simultaneously, which can only be accomplished in the analysis of the dynamic equilibrium in the next section.

Proposition 4 *In equilibrium,*

1. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln N(t)} > 0 \Leftrightarrow \varepsilon > 1$;
2. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln E(t)} > 0 \Leftrightarrow \varepsilon > \left[1 + \frac{\beta}{(1-\beta)}\left(\frac{d\ln p^{SC}(t)}{d\ln E(t)}\right)^{-1}\right] > 1$;

where $SC = FA,\ FD$.

Proof. See Appendix G. ■

Proposition 4 illustrates how changes in technology, $N(t)$, and societal expenditure, $E(t)$, influence employment in the M-sector, $L_M^{SC}(t)$, relative to the S-sector, $L_S^{SC}(t)$. (Recall from Proposition 3 that change in population, $L(t)$, do not influence employment in the M- and S-sectors.) Proposition 4 implies that the direction of NBG in the economy – i.e., whether $L_M^{SC}(t)$ increases relative to $L_S^{SC}(t)$, or vice versa, as the economy grows – is determined by the elasticity of substitution between M- and S-sector output, ε. Notice that in contrast to Proposition 3, where increases in $N(t)$ and $E(t)$ have opposing effects on the absolute levels of $L_M^{SC}(t)$ and $L_S^{SC}(t)$, Proposition 4 establishes that increases in $N(t)$ and $E(t)$ have symmetric effects on $L_M^{SC}(t)/L_S^{SC}(t)$ for most values of ε.[31]

The intuition for Proposition 4 is that an increase in either $N(t)$ or $E(t)$ has two countervailing effects on $L_M^{SC}(t)/L_S^{SC}(t)$. First, increases in $N(t)$ or $E(t)$ increase the marginal product of labor in the M-sector relative to the S-sector, which draws labor to the M-sector.[32] Second, from Proposition 3, increases in both $N(t)$ and $E(t)$ raise the relative price of S-sector output, $p^{SC}(t)$. In the case of an increase in $N(t)$, when $\varepsilon < 1$, and M- and S-sector output are gross complements, the increase in $p^{SC}(t)$ dominates, and $L_M^{SC}(t)/L_S^{SC}(t)$ declines with $N(t)$. Conversely, when $\varepsilon > 1$,

[31] From Proposition 4, increases in $N(t)$ and $E(t)$ have opposing effects on $L_M^{SC}(t)/L_S^{SC}(t)$ for $\varepsilon \in \left(1, 1 + \frac{\beta}{(1-\beta)}\left(\frac{d\ln p^{SC}(t)}{d\ln E(t)}\right)^{-1}\right)$.

[32] From Proposition 1, the marginal product of labor in the M-sector is $(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}\left[N_k(t) + \Gamma^{\frac{\alpha}{\alpha-1}}N_l(t)\right]$ and the marginal product of labor in the S-sector is $\beta L_{kS}(t)^{\beta-1}F_k{}^{1-\beta}$, $k = a, b$. An increase in $N(t)$ increases the marginal product of labor in the M-sector but not the S-sector for fixed employment in two sectors. Similarly, as is established in Proposition 3 an increase in $E(t)$ increases both $L_M^{SC}(t)$ and $L_S^{SC}(t)$. This increases do not change the marginal product of labor in the M-sector, but decrease the marginal product of labor in the S-sector.

and M- and S-sector output are gross substitutes, the relative decline in the marginal product of labor in the S-sector dominates, and $L_M^{SC}(t)/L_S^{SC}(t)$ expands with $N(t)$. When $\varepsilon = 1$, the two effects cancel, and $L_M^{SC}(t)/L_S^{SC}(t)$ is unchanged by change in $N(t)$. The same intuition holds for an increase in $E(t)$, with the only difference being that $L_M^{SC}(t)/L_S^{SC}(t)$ increases with $E(t)$ if and only if $\varepsilon > 1 + \frac{\beta}{(1-\beta)}\left(\frac{d\ln p^{SC}(t)}{d\ln E(t)}\right)^{-1} > 1$.

The role of ε in determining the direction of the change in $L_M^{SC}(t)/L_S^{SC}(t)$ as a result of increases in $N(t)$ and $E(t)$ that is illustrated in Proposition 4 is consistent with previous studies that have emphasized supply-side explanations for non-balanced growth (e.g., Baumol 1967; Ngai and Pissarides 2007; and Acemoglu and Guerrieri 2008). As is discussed in Chapter 1.2.1, these previous studies posit that non-balanced growth occurs primarily as the result of differential rates of productivity growth across sectors. To explain deindustrialization, these studies contend that productivity growth has proceeded more slowly in services, but, because the elasticity of substitution between service-sector output and output from goods-producing sectors is low, the increase in the relative price of service-sector output that results from slower sectoral productivity growth more than compensates for the decrease in the relative return to factors of production in the service sector, causing the sector's share of total employment to increase as the economy grows.

Proposition 5 *In equilibrium,*

1. $\frac{d\ln L_a^{FA}(t)}{d\ln N(t)} > 0$, $\frac{d\ln L_a^{FA}(t)}{d\ln E(t)} < 0$, *and* $\frac{d\ln L_a^{FA}(t)}{d\ln L(t)} = \frac{L(t)}{L_a^{FA}(t)} > 1$;

2. $\frac{d\ln L_a^{FD}(t)}{d\ln N(t)} > 0 \Longleftrightarrow F_a - F_b < 0$, $\frac{d\ln L_a^{FD}(t)}{d\ln E(t)} > 0 \Longleftrightarrow F_a - F_b > 0$, *and* $\frac{d\ln L_a^{FD}(t)}{d\ln L(t)} = \frac{1}{2}\frac{L(t)}{L_a^{FA}(t)} > 1 \Longleftrightarrow F_a - F_b < 0$.

Proof. See Appendix G. ∎

The intuition for Proposition 5 follows directly from Proposition 3. In particular, from Proposition 3, an increase in $N(t)$ increases employment in the I-sector at the expense of both the M- and S-sectors. Under full agglomeration, this increase in I-sector employment will occur entirely in region a, where the I-sector is concentrated. Under full dispersion, as employment in the I-sector is evenly distributed between the two regions, the increase in I-sector employment does not influence the interregional distribution of employment in the I-sector. On the other hand, the share of S-sector employment in region a relative to region b is determined by each region's endowment of the fixed factor. The decline in S-sector employment as a result of an increase in $N(t)$ is greater in the region with the larger endowment of the fixed factor. This is the intuition for the first term in part 2 of Proposition 3, where the region with smaller endowment of the fixed fact or grows in population as a consequence of an increase in $N(t)$.

Similarly, Proposition 3 establishes that an increase in $E(t)$ increases employment in the M- and S-sectors at the expense of the I-sector. Under full agglomeration, when the I-sector is concentrated in region a, the decline in I-sector employment occurs entirely in region a, thereby lowering region a's population. Under full dispersion, an increase in $E(t)$ does not influence the interregional distribution of employment in either the M- or I-sectors; however, the region with the larger endowment of the fixed factor will experience a larger increase in S-sector employment as a result of an increase in $E(t)$, and, as a consequence, will increase in population.

Finally, again from Proposition 3, an increase in population, $L(t)$, increases employment in the I-sector, and does not affect employment in either the M- or S-sectors. As such, under full agglomeration, an increase in $L(t)$ increases population in region a, where the I-sector is concentrated. Conversely, under full dispersion, the increases in I-sector employment occurs evenly across regions a and b, thereby increasing the relative employment share of the region with the smaller endowment of the fixed factor, and, hence, the smaller share of S-sector employment.

Proposition 5 is similar to Proposition 3 in that it illustrates the limits to comparative static analysis in this chapter. In particular, Proposition 5 demonstrates that $N(t)$, $E(t)$, and $L(t)$ have countervailing influences on the distribution of population between regions, and that understanding the influence of NBG on regional population movements will require analyzing changes in $N(t)$, $E(t)$, and $L(t)$ simultaneously, which can only be accomplished in the analysis of the dynamic equilibrium.

1.6 Dynamic Equilibrium

Definition 2 *The dynamic equilibrium is a set of dynamic trajectories for prices – the relative price of S-sector output, $p(t)$, the wage in regions a and b, $w_a(t)$ and $w_b(t)$, the rent to the fixed factor in regions a and b, $w_a^F(t)$ and $w_b^F(t)$, prices of intermediate inputs, $\{p_{i_a a}(t)\}_{i_a \in \mathbb{M}_a(t)}$, and $\{p_{i_b b}(t)\}_{i_b \in \mathbb{M}_b(t)}$, the value of I-sector firms, $V(t)$, and the interest rate, $r(t)$ – labor allocations, $L_{aM}(t)$, $L_{bM}(t)$, $L_{aS}(t)$, $L_{bS}(t)$, $L_{aI}(t)$, and $L_{bI}(t)$, and levels of production, $Y_{aM}(t)$, $Y_{bM}(t)$, $Y_{aS}(t)$, $Y_{bS}(t)$, $Y_{aI}(t)$, $Y_{bI}(t)$, $\{Y_{i_a}(t)\}_{i_a \in \mathbb{M}_a(t)}$, and $\{Y_{i_b}(t)\}_{i_b \in \mathbb{M}_b(t)}$ such that the economy is in static equilibrium at every time t and that the first-order conditions for household j's intertemporal program hold for every household $j = 1, ..., H$ in the economy (i.e., that every household j's expenditure path is characterized by its Euler equation, (1.13), budget-flow constraint, (1.6), initial conditions $l_j(0)$, $a_j(0)$, f_{ja}, and f_{jb}, and transversality condition, (1.12)).*

This section characterizes and analyzes the dynamic equilibrium for the economy. This section begins with a proposition that defines the necessary conditions for the economy to be in dynamic equilibrium. These necessary conditions are used to characterize the optimal dynamic trajectory for the economy. Following this, a long-run growth path for the economy is defined that features a constant real interest rate and constant growth in total societal expenditure. It is shown that along this long-run growth path, economic growth is non-balanced, with output and employment growing at different rates across sectors. This section ends by demonstrating that there is a unique optimal dynamic trajectory for the economy, and that this trajectory converges with the long-run growth path. Analyzing the behavior of the economy along this optimal trajectory yields many of the key results in this chapter concerning the implications of NBG for the spatial development of the economy.

Proposition 6 *A dynamic equilibrium satisfies the following three differential equations,*

$$\begin{aligned} \hat{N}^{FA}(t) &= L(t) - \frac{1}{(1+\alpha)} \frac{E(t)}{w^{FA}(t)\left[1 + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FA}(t)^{1-\varepsilon}\right]} - \left(\frac{\beta p^{FA}(t)}{w^{FA}(t)}\right)^{\frac{1}{1-\beta}} (F_a + F_b) \\ \text{or } \hat{N}^{FD}(t) &= \frac{1}{2}(1+\eta) \left\{ \begin{array}{c} L(t) - \frac{1}{(1+\alpha)} \frac{E(t)}{w^{FD}(t)\left[1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}\right]} \\ - \left(\frac{\beta p^{FD}(t)}{w^{FD}(t)}\right)^{\frac{1}{1-\beta}} (F_a + F_b) \end{array} \right\}, \end{aligned} \tag{1.58}$$

$$\begin{aligned} \hat{E}^{FA}(t) &= \frac{\alpha}{(1+\alpha)} \frac{E(t)}{w^{FA}(t)\left[1 + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FA}(t)^{1-\varepsilon}\right]} - (\delta - n) \\ \text{or } \hat{E}^{FD}(t) &= \frac{1}{2}[1+\eta] \frac{\alpha}{(1+\alpha)} \frac{E(t)}{w^{FD}(t)\left[1 + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}\right]} - (\delta - n), \end{aligned} \tag{1.59}$$

and

$$\hat{L}(t) = n, \tag{1.60}$$

where, from (1.37), $w^{FA}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}} N(t)$ *and* $w^{FD}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}} \frac{1}{2}\left[1 + \Gamma^{\frac{\alpha}{\alpha-1}}\right] N(t)$, *and* $p^{SC}(t)$, $SC = FA, FD$, *is defined implicitly in terms of* $E(t)$ *and* $N(t)$ *(as well as the parameters* ε, μ, α, Γ, *and* β*) by the* S*-sector market-clearing condition, (1.39), along with the initial values of technology,* $N(0)$, *and population,* $L(0)$, *and the transversality condition*

$$\lim_{t \to \infty} \exp \left[\int_0^t \left(\hat{N}^{SC}(\tau) - r^{SC}(\tau) \right) d\tau \right] = 0, \tag{1.61}$$

where $\hat{N}^{SC}(\tau)$ *is given by (1.58), and, from (1.57),* $r^{FA}(t) = \alpha L_M^{FA}(t)$ *and* $r^{FD}(t) = \alpha \frac{1}{2}(1+\eta) L_M^{FD}(t)$ *with* $L_M^{SC}(t)$ *defined in (1.43).*

Proof. See Appendix H. ■

Proposition 6 defines a set of necessary conditions for the economy to be in dynamic equilibrium. In particular, (1.59) - (1.61) are necessary conditions for the first-order conditions for household j's intertemporal program to hold for every household in the economy.[33] The necessary conditions in Proposition 6 are used in the remainder of this chapter to characterize the dynamic trajectories for sectoral output, sectoral employment, and regional population when the economy is in dynamic equilibrium.

1.6.1 Equilibrium Growth Path

This section defines an *equilibrium growth path* (EGP) as a dynamic equilibrium for the economy that features a constant real interest rate and constant growth in societal expenditure.[34] Theorem 3 (below) shows that when $\varepsilon < 1$, there exists a unique EGP for the economy and that along this EGP the S-sector dominates the asymptotic distribution of employment. Theorem 3 provides closed-form solutions for the growth rates of $N(t)$ and $E(t)$ along this EGP, as well as for growth rates of sectoral output, sectoral employment, and regional population. In addition, Theorem 3 establishes that there does not exist an EGP when $\varepsilon \geq 1$.[35]

---

[33]It was established in Chapter 1.4.1 that these first-order conditions are both necessary and sufficient to identify a unique maximum for household j's intertemporal program.

[34]Kongsamut, Rebelo, and Xie (2001) refer to a long-run growth path for the economy undergoing NBG that features a constant real interest rate as a "generalized balance growth path".

[35]That an EGP for the economy only exists when $\varepsilon < 1$ is consonant with Acemoglu and Guerrieri (2008), who impose a set parameter restrictions on their model of NBG that ensure that the less-progressive, labor-intensive sector will be the "asymptotically dominant sector", where they define the asymptotically dominant sector as the sector that determines the long-run growth rate of the economy. These restrictions are necessary for Acemoglu and Guerrieri (2008)'s model to generate patterns of NBG in sectoral output and employment, while also exhibiting long-run growth consistent with the "Kaldor Facts" (i.e., constancy of the rate of consumption growth, real interest rate, capital-

Before presenting Theorem 3, it is necessary to introduce new notation for the asymptotic growth rates in the economy:

$$g_E^{SC} = \lim_{t\to\infty} \hat{E}^{SC}(t) \text{ and } g_N^{SC} = \lim_{t\to\infty} \hat{N}^{SC}(t);$$

$$g_P^{SC} = \lim_{t\to\infty} \hat{Y}_P^{SC}(t), \; n_P^{SC} = \lim_{t\to\infty} \hat{L}_P^{SC}(t), \text{ and } \bar{L}_P^{SC} = \lim_{t\to\infty} L_P^{SC}(t), \; P = M, S, I;$$

$$n_k^{SC} = \lim_{t\to\infty} \hat{L}_k^{SC}(t) \text{ and } \overline{\frac{L_a^{SC}(t)}{L_b^{SC}(t)}} = \lim_{t\to\infty} \frac{L_a^{SC}(t)}{L_b^{SC}(t)};$$

and

$$g_p^{SC} = \lim_{t\to\infty} \hat{p}^{SC}(t), \; g_w^{SC} = \lim_{t\to\infty} \hat{w}^{SC}(t), \; g_{w^F}^{SC} = \lim_{t\to\infty} \hat{w}^{F,SC}(t), \; g_P^{SC} = \lim_{t\to\infty} \hat{P}^{SC}(t), \; \bar{r}^{SC} = \lim_{t\to\infty} r^{SC}(t);$$

$SC = FA, FD$ and $k = a, b$.

Theorem 3 *When $\varepsilon < 1$, there exists a unique EGP for the economy where expenditure, technology, and M-sector output grow at the constant rate*

$$g_E^{SC} = g_N^{SC} = g_M^{SC} = \left[\frac{(1-\varepsilon)\beta + \varepsilon}{(1-\varepsilon)}\right] n = \left(\beta + \frac{\varepsilon}{1-\varepsilon}\right) n, \tag{1.62}$$

$SC = FA, FD$, S-sector output and employment grow at the constant rates

$$g_S^{SC} = \beta n \text{ and } n_S^{SC} = n, \tag{1.63}$$

the relative price of S-sector output, the wage, the rental rate of the fixed factor, and the price index that gives the minimum cost for M-sector firms of purchasing a unit of $I^{SC}(t)$, grow at the constant rates

$$g_p^{SC} = (1-\varepsilon)^{-1} n, \; g_w^{SC} = g_E^{SC}, \; g_{w^F}^{SC} = \left\{\frac{1 - \beta\left[(1-\beta)\varepsilon + \beta\right]}{1 - \left[(1-\beta)\varepsilon + \beta\right]}\right\} n, \tag{1.64}$$

and

$$g_P^{SC} = -\left(\frac{1-\alpha}{\alpha}\right)\left[\frac{(1-\varepsilon)\beta + \varepsilon}{(1-\varepsilon)}\right],$$

to-output ratio, and the share of capital income in total output).

M-sector employment is constant and equal to

$$\bar{L}_M^{FA} = \frac{1}{\alpha}\left\{\left[\frac{(1-\varepsilon)\beta+\varepsilon}{(1-\varepsilon)}\right]n + (\delta - n)\right\} \text{ and } \bar{L}_M^{FD} = \frac{2}{\alpha(1+\eta)}\left\{\left[\frac{(1-\varepsilon)\beta+\varepsilon}{(1-\varepsilon)}\right]n + (\delta - n)\right\}, \tag{1.65}$$

I-sector employment is constant and equal to

$$\bar{L}_I^{FA} = \left[\frac{(1-\varepsilon)\beta+\varepsilon}{(1-\varepsilon)}\right]n \text{ and } \bar{L}_I^{FD} = \frac{2}{1+\eta}\left[\frac{(1-\varepsilon)\beta+\varepsilon}{(1-\varepsilon)}\right]n. \tag{1.66}$$

each regions' population grows at the constant rate,

$$n_a^{FA} = n_b^{FA} = n_a^{FD} = n_b^{FD} = n, \tag{1.67}$$

and regional population shares are constant and equal to

$$\overline{\frac{L_a^{FA}(t)}{L_b^{FA}(t)}} = \overline{\frac{L_a^{FD}(t)}{L_b^{FD}(t)}} = \frac{F_a}{F_b}. \tag{1.68}$$

In addition, there does not exist an EGP when $\varepsilon \geq 1$.

Proof. See Appendix H. ■

Theorem 3 implies that along the EGP, economic growth is non-balanced, with employment and output in the M- and S-sectors growing at different asymptotic rates. (Along the EGP, output grows at the same rate in the M- and I-sectors, and employment in the M- and I-sectors is constant.) The intuition for the mechanics of NBG in the model follow from the fact that productivity growth occurs in the M-sector but not in the S-sector. All else equal, faster productivity growth in the M-sector would cause employment to grow more rapidly in the M-sector than in the S-sector. However, productivity growth in the M-sector also causes output to grow faster in the M-sector than in the S-sector, which increases the relative price of S-sector output, $p(t)$. The increase in $p(t)$ induces the reallocation of labor to the S-sector. When $\varepsilon < 1$, the increase in relative price of S-sector output is more than proportional to the increase in the marginal product of labor in the M-sector, and the S-sector gains in employment relative to the M-sector along the EGP. In particular, along the EGP, the increase in the relative price of S-sector output is exactly such that no new labor is allocated towards or taken away from either the M-sector (or the I-sector) as the economy grows.

Notice that the reallocation of labor to the S-sector along the EGP cannot offset the increase in output in the M-sector relative to the S-sector that results from M-sector productivity growth because, if it did, the change in relative prices would not take place, and there would be no force in the economy attracting labor to the S-sector. For this reason, even though employment in the S-sector is growing along the EGP while employment in the M-sector is constant, output in the M-sector is growing faster than output in the S-sector along the EGP, i.e., $g_M^{SC} > g_S^{SC}$ along the EGP.

The fact that the S-sector dominates the asymptotic distribution of employment along the EGP implies that the rate of population growth, n, determines the asymptotic growth rates for the economy. This is the case because output in the S-sector can only increase as a result of an increase in employment in the S-sector. For this reason, the asymptotic growth rate of output in the S-sector, and, as a consequence, the asymptotic growth rates for the other aggregates in the economy, is determined by n. The asymptotic growth rate of S-sector output along the EGP is $g_S^{SC} = \beta n$, where β is the output elasticity of labor in S-sector production. The asymptotic growth rate of S-sector output is increasing in β because a larger value of β implies a larger increase in S-sector output for a given increase in S-sector employment. The asymptotic growth rates of $N(t)$, $E(t)$, and $Y_M^{SC}(t)$ are also determined by n (i.e., $g_N^{SC} = g_E^{SC} = g_M^{SC} = \left(\beta + \frac{\varepsilon}{(1-\varepsilon)}\right) n$). The asymptotic growth rates of these aggregates along the EGP are closer to the asymptotic growth rate of S-sector output for lower values of ε. Lower values of ε imply that M- and S-sector output are more complementary. As a result, lower values of ε lead to a larger portion of total labor being devoted to S-sector production at each point along the EGP, with less labor devoted to I-sector innovation, which depresses the growth rates of $N(t)$, $E(t)$, and $Y_M^{SC}(t)$.

The fact that the S-sector determines the asymptotic growth rates for the economy explains the fact that the spatial distribution of the economy, and, hence, agglomeration externalities, do not influence the asymptotic growth rates of $E(t)$ or $N(t)$. The S-sector, after all, does not benefit from agglomeration externalities. Agglomeration externalities do, however, influence the constant levels of employment along the EGP in the M- and I-sectors. Indeed, from (1.65) and (1.66), the constant levels of employment in the M- and I-sectors are higher under full dispersion than under full agglomeration. This reflects the role of knowledge spillovers, captured by the parameter η, in I-sector innovation, which make it so that more I-sector labor is required under full dispersion to maintain a given growth rate of $N(t)$ along the EGP. In addition, agglomeration externalities related to forward and backward linkages between the M- and I-sectors imply that, all else equal, the profits of I-sector firms are lower under full dispersion than under full agglomeration. Lower I-sector profits reduce the incentive for firms to develop new varieties of intermediate input and enter the I-sector, which reduces the growth rate of $N(t)$. For $N(t)$ to grow at the same constant rate under both full agglomeration and full dispersion, M-sector production, and, hence, M-sector employment and demand for intermediate inputs, must be higher under full dispersion. Higher M-sector demand for intermediate inputs increases the profits of I-sector firms, which compensates for the reduction in profits due to the decreased benefit from agglomeration externalities to I-sector firms under full dispersion.

That the S-sector dominates the asymptotic distribution of employment along the EGP also implies that the S-sector determines the asymptotic population growth rates of regions a and b, which are equal to the rate of population growth, n, under both spatial configurations of the economy, as well as the asymptotic regional population shares, which are determined by the distribution of S-sector employment across regions a and b (which, in turn, are determined by the distribution of the fixed factor across regions a and b, i.e., F_a and F_b).

Regional employment shares are constant along the EGP because employment in the M- and I-sectors along the EGP becomes, to use the terminology of Acemoglu and Guerrieri (2008), "vanishingly small" relative to S-sector employment.[36] That employment in the M- and I-sectors becomes vanishingly small means that analyzing the EGP does not address the dynamic behavior of the economy when there are comparable (i.e., non-trivial) levels of employment in the M-, S-, and I-sectors. The next section demonstrates that there is a unique dynamic trajectory for the economy that is consistent with the necessary conditions for a dynamic equilibrium from Proposition 6, and analyzes the dynamics of sectoral employment, sectoral output, regional population along this dynamic trajectory when there are non-trivial levels of employment in the M-, S-, and I-sectors.

1.6.2 Optimal Dynamic Trajectory

In this section, phase diagram analysis is used to characterize the dynamic equilibrium outside of the EGP. Phase diagram analysis is used to establish that there is a unique dynamic trajectory for the economy that is consistent with the necessary conditions for a dynamic equilibrium from Proposition 6, and that along this optimal dynamic trajectory $N(t)$ and $E(t)$ grow at the same rate. Furthermore, it is shown that this optimal dynamic trajectory approaches the EGP. The dynamic behavior of sectoral employment, sectoral output, and regional population along the optimal dynamic trajectory defined in this section is analyzed in Chapter 1.6.3.

[36] Along the EGP, the share of labor allocated to the S-sector tends to one. Despite this, output in the M- and I-sectors continue to grow at positive rates, and both sectors maintain positive levels of employment. This implies that this limit point, i.e., $\lim_{t\to\infty} L_S^{SC}(t)/L(t) = 1$, $SC = FA, FD$, is never reached.

Proposition 6 demonstrates that the dynamic equilibrium for the economy is represented by a boundary value system of three ordinary differential equations in time, the initial values of technology, $N(0)$, and population, $L(0)$, and the transversality condition in (1.61). By incorporating the expression for $\hat{L}(t)$ from (1.60) into the expression for $\hat{N}(t)$ from (1.58), this system can be re-expressed as a system of one autonomous differential equation for $E(t)$, (1.59), and one non-autonomous differential equation for $N(t)$:

$$\begin{aligned}\hat{N}^{FA}(t) &= L(0)e^{rt} - \frac{1}{(1+\alpha)}\frac{E(t)}{w^{FA}(t)\left[1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{FA}(t)^{1-\varepsilon}\right]} \\ &\quad - \left(\frac{\beta p^{FA}(t)}{w^{FA}(t)}\right)^{\frac{1}{1-\beta}}(F_a+F_b) \\ \text{or } \hat{N}^{FD}(t) &= \frac{1}{2}(1+\eta)\left\{\begin{array}{c} L(0)e^{rt} - \frac{1}{(1+\alpha)}\frac{E(t)}{w^{FD}(t)\left[1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{FD}(t)^{1-\varepsilon}\right]} \\ -\left(\frac{\beta p^{FD}(t)}{w^{FD}(t)}\right)^{\frac{1}{1-\beta}}(F_a+F_b)\end{array}\right\},\end{aligned} \tag{1.69}$$

where, from (1.37), $w^{FA}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}N(t)$ and $w^{FD}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}\frac{1}{2}\left[1+\Gamma^{\frac{\alpha}{\alpha-1}}\right]N(t)$, and $p^{SC}(t)$, $SC = FA, FD$, is defined implicitly in terms of $E(t)$ and $N(t)$ by the S-sector market-clearing condition, (1.39). Standard phase diagram analysis cannot be used for non-autonomous systems. For this reason, phase diagram analysis with moving isoclines is used in this section.

It is established that the $\hat{N}^{SC}(t) = 0$, $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, and $\hat{E}^{SC}(t) = 0$ isoclines are as depicted in Figure 1.1 for all t in three steps. First, that the three isoclines are upward sloping and pass through the origin follows from the expressions for $\hat{N}^{SC}(t)$ and $\hat{E}^{SC}(t)$, $SC = FA, FD$, from (1.69) and (1.59). In particular, using (1.69) and the S-sector market-clearing condition from (1.39), under full agglomeration, $\hat{N}^{FA}(t) = 0$ implies

$$N(t) = \frac{1}{L(0)e^{rt}}\frac{1}{(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}}\left[\frac{(1+\alpha)^{-1}}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{FA}(t)^{1-\varepsilon}} + \frac{\beta}{1+\left[\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{FA}(t)^{1-\varepsilon}\right]^{-1}}\right]E(t), \tag{1.70}$$

and, under full dispersion, $\hat{N}^{FD}(t) = 0$ implies

$$N(t) = \frac{1}{L(0)e^{rt}} \frac{1}{(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}} \frac{1}{\frac{1}{2}\left(1+\Gamma^{\frac{\alpha}{a-1}}\right)} \left[\begin{array}{c} \frac{(1+\alpha)^{-1}}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}} \\ +\frac{\beta}{1+\left[\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}\right]^{-1}} \end{array} \right] E(t). \quad (1.71)$$

Next, using (1.69), (1.59), and the S-sector market-clearing condition from (1.39), under full agglomeration, $\hat{N}^{FA}(t) = \hat{E}^{FA}(t)$ implies

$$N(t) = \frac{1}{L(0)e^{rt} + (\delta - n)} \frac{1}{(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}} \left[\frac{1}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}} + \frac{\beta}{1+\left[\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FA}(t)^{1-\varepsilon}\right]^{-1}} \right] E(t), \quad (1.72)$$

and, under full dispersion, $\hat{N}^{FD}(t) = \hat{E}^{FD}(t)$ implies

$$N(t) = \frac{1}{L(0)e^{rt} + \left[\frac{1}{2}(1+\eta)\right]^{-1}(\delta - n)} \frac{1}{(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}} \frac{1}{\frac{1}{2}\left(1+\Gamma^{\frac{\alpha}{a-1}}\right)} \left[\begin{array}{c} \frac{1}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}} \\ +\frac{\beta}{1+\left[\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}\right]^{-1}} \end{array} \right] E(t). \quad (1.73)$$

Finally, using (1.59), under full agglomeration, $\hat{E}^{FA}(t) = 0$ implies

$$N(t) = \frac{1}{(\delta - n)} \frac{1}{(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}} \left[\frac{\alpha(1+\alpha)^{-1}}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FA}(t)^{1-\varepsilon}} \right] E(t), \quad (1.74)$$

and, under full dispersion, $\hat{E}^{FD}(t) = 0$ implies

$$N(t) = \frac{(1+\eta)}{\left(1+\Gamma^{\frac{\alpha}{a-1}}\right)} \frac{1}{(\delta - n)} \frac{1}{(1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}} \left[\frac{\alpha(1+\alpha)^{-1}}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{FD}(t)^{1-\varepsilon}} \right] E(t). \quad (1.75)$$

It is clear from (1.70)-(1.75) that the $\hat{N}^{SC}(t) = 0$, $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, and $\hat{E}^{SC}(t) = 0$ isoclines are all upward sloping and pass through the origin.

Second, that the $\hat{N}^{SC}(t) = 0$ and $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$ isoclines move to the southeast as t gets larger (as is depicted in Figure 1.1) follows from the fact that the term $L(0)e^{rt}$ appears in the denominator of the expressions for the $\hat{N}^{SC}(t) = 0$ and $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$ isoclines.

Third, to establishes that the $\hat{N}^{SC}(t) = 0, \hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, and $\hat{E}^{SC}(t) = 0$ isoclines are always in the same relation to one another as is depicted in Figure 1.1, it is sufficient to demonstrate that $E_{\hat{E}=0}(t) < E_{\hat{E}=\hat{N}}(t) < E_{\hat{N}=0}(t)$ for all $N(t) > 0$ and for all t, where

$$\begin{aligned} E_{\hat{E}=0}(t) &= \left\{E(t) \mid \hat{E}^{SC}(t) = 0, N(t)\right\} \\ E_{\hat{E}=\hat{N}}(t) &= \left\{E(t) \mid \hat{E}^{SC}(t) = \hat{N}^{SC}(t), N(t)\right\} \\ E_{\hat{N}=0}(t) &= \left\{E(t) \mid \hat{N}^{SC}(t) = 0, N(t)\right\}. \end{aligned} \tag{1.76}$$

That $E_{\hat{E}=0}(t) < E_{\hat{E}=\hat{N}}(t) < E_{\hat{N}=0}(t)$ for all $N(t) > 0$ and for all t follows from the facts that

$$\hat{E}^{FA}(t)\Big|_{E(t)=0} = -(\delta - n) < 0 \text{ and } \hat{E}^{FD}(t)\Big|_{E(t)=0} = -\frac{1}{2}(1+\eta)(\delta - n) < 0,$$

$$\frac{\partial \hat{E}^{FA}(t)}{\partial E(t)} = \alpha \frac{d \ln L_M^{FA}(t)}{d \ln E(t)} \frac{L_M^{FA}(t)}{E(t)} > 0 \text{ and } \frac{\partial \hat{E}^{FD}(t)}{\partial E(t)} = \alpha \frac{1}{2}(1+\eta) \frac{d \ln L_M^{FD}(t)}{d \ln E(t)} \frac{L_M^{FD}(t)}{E(t)} > 0,$$

$$\hat{N}^{FA}(t)\Big|_{E(t)=0} = L(0) e^{rt} > 0 \text{ and } \hat{N}^{FD}(t)\Big|_{E(t)=0} = \frac{1}{2}(1+\eta) L(0) e^{rt} > 0,$$

and

$$\frac{\partial \hat{N}^{FA}(t)}{\partial E(t)} = \frac{d \ln L_I^{FA}(t)}{d \ln E(t)} \frac{L_I^{FA}(t)}{E(t)} < 0 \text{ and } \frac{\partial \hat{N}^{FA}(t)}{\partial E(t)} = \frac{1}{2}(1+\eta) \frac{d \ln L_I^{FD}(t)}{d \ln E(t)} \frac{L_I^{FD}(t)}{E(t)} < 0.$$

That $\frac{d \ln L_M^{SC}(t)}{d \ln E(t)} > 0$ and $\frac{d \ln L_I^{SC}(t)}{d \ln E(t)} < 0$ is established in Proposition (3).

Next, the phase diagram in Figure 1.2 is used to establish that the saddle-path defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ is the only dynamic trajectory for the economy that is consistent with the dynamic equilibrium conditions from Proposition 6. As the discussion above establishes that the $\hat{N}^{SC}(t) = 0, \hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, and $\hat{E}^{SC}(t) = 0$ isoclines are as depicted in Figure 1.1 for all t, the phase diagram in Figure 1.2 indicates the directions of movement for $N(t)$ and $E(t)$ for any time t. Also, Figures 1.1 and 1.2 depict the $\hat{N}^{SC}(t) = 0, \hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, and $\hat{E}^{SC}(t) = 0$ isoclines as linear, despite it being clear from (1.70)-(1.75) that they are not linear. The

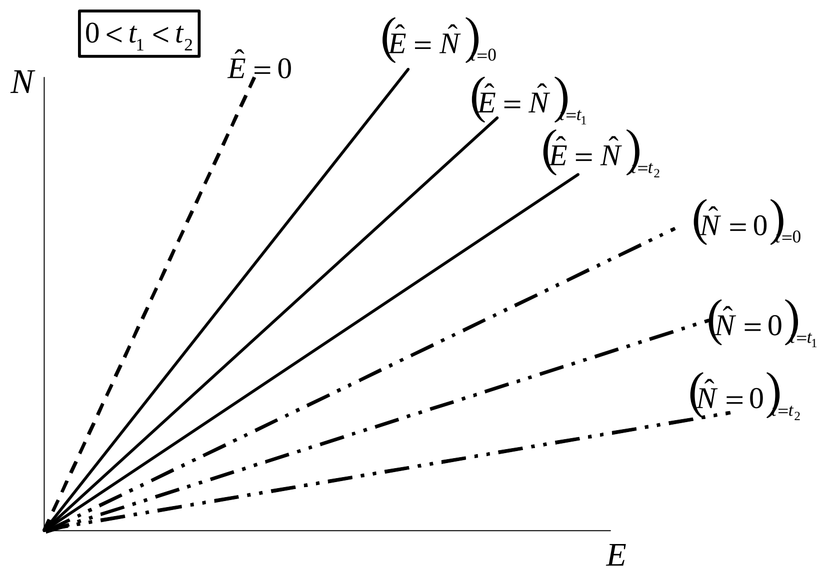

Figure 1.1: The $\hat{N}^{SC}(t) = 0, \hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, and $\hat{E}^{SC}(t) = 0$ isoclines for $t = 0$, $t = t_1$, and $t = t_2 > t_1$. See text for further detail.

isoquants are depicted as linear because doing so does not affect the demonstration of the key result that the saddle-path defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ is the only dynamic trajectory for the economy that is consistent with the necessary conditions for a dynamic equilibrium from Proposition 6.

The optimality of the saddle-path defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ is established in two steps. First, consider a dynamic trajectory of the economy to the northwest of the $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ isocline, such as trajectory "A" in Figure 1.2. Along such trajectories, $N(t)$ is growing faster than $E(t)$. This will cause the trajectory to eventually cross the $\hat{E}^{SC}(t) = 0$ isocline and results in $E(t) = 0$ in finite time. From (1.57) and (1.58), that $E(t) = 0$ in finite time along trajectory "A" implies that in the long-run, $r^{SC}(t) = 0$ and $\hat{N}^{SC}(t) > 0$, which violates the necessary condition for households' transversality conditions to hold in (1.61) from Proposition 6. Hence trajectory "A" and any trajectory to the northwest of the $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ isocline will not be optimal because it will violate the transversality condition for at least one household in the economy. Along such trajectories, households would realize that they were over-investing in the develop of new varieties of intermediate inputs, and would choose to move to an expenditure path that entailed less investment.

Second, consider a dynamic trajectory for the economy to the southeast of the $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ isocline, such as trajectory "B" in Figure 1.2. Along such trajectories, where $E^{SC}(t)$ is growing at a faster rate than $N^{SC}(t)$, (1.59) from Proposition 6 is violated. (1.59) implies that for all points to the southeast of the $\hat{E}^{SC}(t) = 0$ isocline, $\hat{E}^{SC}(t) > 0$. However, along trajectory "B", $\hat{N}^{SC}(t) = 0$ in finite time, which means that $\hat{E}^{SC}(t)$ jumps to zero in finite time, violating (1.59). Along trajectory "B" and any trajectory to the southeast of the $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ isocline, households would realize that they were under-investing in the development of new varieties of intermediate inputs and would choose to move to an expenditure path that entailed more investment. Taken together, these two steps establish that the saddle-path

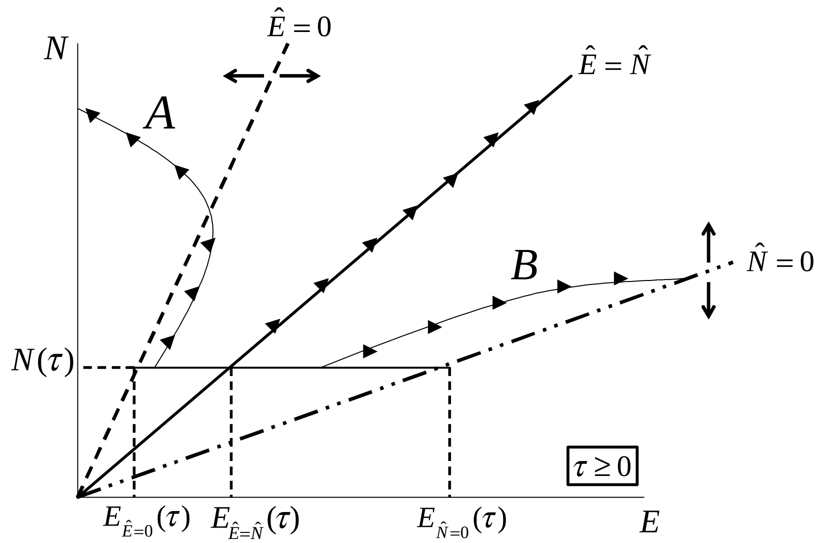

Figure 1.2: That the saddle-path defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ is the only dynamic trajectory for the economy that satisfies the necessary conditions for a dynamic equilibrium described in Proposition 6 is established in two steps: (1) Along trajectory A, (1.61) from Proposition 6 (the necessary condition for each household's transversality condition to hold) is violated; (2) Along trajectory B, (1.59) from Proposition 6 (the necessary condition for each households Euler equation to hold) is violated. See text for further detail.

defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ is the only dynamic trajectory for the economy that satisfies the necessary conditions for a dynamic equilibrium from Proposition 6.

Proposition 7 *A sufficient condition for the optimal dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$, $SC = FA, FD$, to approach the EGP is for $\hat{L}_S^{SC}(t)$ to approach n from above as $t \to \infty$.*

Proof. See Appendix H. ■

Proposition 7 states that a sufficient condition for the dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$, $SC = FA, FD$, to approach the EGP is that the S-sector's share of total employment is expanding relative to the M- and I-sector's (i.e., that $\hat{L}_S^{SC}(t) > n$) as the economy grows along the dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$.[37] This condition agrees with the description of the EGP in Theorem 3 in that when $\hat{L}_S^{SC}(t) > n$ along the dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ then the S-sector will dominate the asymptotic distribution of employment, as is the case along the EGP, and both M-sector employment (and, from (1.57), the interest rate) will tend towards a constant, as is also the case along the EGP. The results presented in Propositions 8 and 9 (below) discuss the implications of the condition that $\hat{L}_S^{FA}(t) > n$ for the dynamic behavior of sectoral output, sectoral employment, and regional population as the economy grows along the dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$.

1.6.3 Optimal Dynamic Trajectory: Analysis

This subsection analyzes how sectoral output, sectoral employment, and regional population evolve as the economy grows along the optimal dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$. The comparative static results in Chapter 1.5.3 demonstrate that increases $N(t)$, $E(t)$, and $L(t)$ have countervailing influences on sectoral output, sectoral employment, and regional population. For this reason, in order to analyze how these aggregates change as the economy evolves it is necessary to consider simultaneous growth in $N(t)$, $E(t)$, and $L(t)$, as is done in this section.

Many of the key results concerning the implications of supply-side NBG for the spatial development of the economy are presented in this section. In particular, it is shown that supply-side NBG can undermine the self-reinforcing relationship between agglomeration and regional population growth, and the positive relationship between

[37] If $\hat{L}_S^{SC}(t) > n$, then the rate of growth of $L_M^{SC}(t) + L_I^{SC}(t)$ must be less than n in order for total population to grow at rate n.

regional productivity growth and regional population growth predicted in previous models of urban and regional development. In addition, in the case of deindustrialization, it is shown that the positive relationship between regional productivity growth and regional population growth emphasized in previous models may be systematically reversed. Further, the results presented in this section confirm Theorem 3 by demonstrating that the long-run spatial distribution of the economy is dominated by the region whose economic characteristics most advantage the sector that dominates the long-run distribution of employment.

Proposition 8 *Along the optimal dynamic trajectory defined by* $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$*: 1.* $\hat{L}_M^{SC}(t) > 0 \Leftrightarrow \varepsilon > 1$*; 2.* $\hat{L}_S^{SC}(t) > 0 \Leftrightarrow \varepsilon < 1$*; 3.* $\hat{Y}_M^{SC}(t) > 0$*; 4.* $\hat{Y}_S^{SC}(t) > 0 \Leftrightarrow \varepsilon < 1$*; 5.* $\hat{L}_I^{SC}(t) = \hat{Y}_I^{SC}(t) = \left[nL(t) - \left(\dot{L}_M^{SC}(t) + \dot{L}_S^{SC}(t)\right)\right] L_I^{SC}(t)^{-1}$*,* $SC = FA, FD$*.*

Proof. See Appendix H. ■

Parts 1 and 2 of Proposition 8 extend the comparative static results presented in Proposition 4. Proposition 4 establishes that the direction of NBG in the economy (whether the progressive M-sector is gaining employment relative to the S-sector, or vice-versa) will be determined by the elasticity of substitution between M- and S-sector output, ε. Proposition 8 extends Proposition 4 by establishing that when the economy is growing along the optimal trajectory defined by $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$ that whether $L_M^{SC}(t)$ and $L_S^{SC}(t)$ grow in absolute terms is also determined by ε. In particular, when $\varepsilon > 1$, $L_M^{SC}(t)$ expands and $L_S^{SC}(t)$ contracts as the economy grows. The converse holds when $\varepsilon < 1$. This implies that when $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$ and $\varepsilon > 1$, the increase in the marginal product of labor in the M-sector is more than proportional to the increase in the relative price of S-sector output, $p(t)$, so that S-sector employment contracts and M-sector employment expands as the economy grows. Conversely, when $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$ and $\varepsilon < 1$, the increase in $p(t)$ is more than proportional to the increase in the marginal product of labor in the M-sector, and the opposite result holds.

Parts 3 and 4 of Proposition 8 relate to the growth in M- and S-sector output. Notice that M-sector output, $Y_M^{SC}(t)$, is growing for all parameter values. This implies that when $\varepsilon < 1$ and the economy is growing along an optimal trajectory defined by $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, the increase in M-sector productivity that results from increases in $N(t)$ is always great enough to offset the decline in $L_M^{SC}(t)$. Part 4 of Proposition 8 confirms that S-sector output, $Y_S^{SC}(t)$, is growing if and only if $\varepsilon < 1$. This result follows from part 1 of Proposition 8 in that the output growth in the S-sector can only be achieved by an increase in employment, and S-sector employment is growing if and only if $\varepsilon < 1$.

Part 5 of Proposition 8 implies that whether I-sector employment, $L_I^{SC}(t)$, and, hence, $Y_I^{SC}(t) = \hat{N}^{SC}(t)$, is increasing or decreasing as the economy grows depends on whether the combined increase in employment in the M- and S-sectors is greater than or less than the increase in total population, $nL(t)$. The intuition for this result is that if the combined increase in $L_M^{SC}(t)$ and $L_S^{SC}(t)$ is greater than $nL(t)$, then $L_I^{SC}(t)$ must decline as the economy grows. The reverse is true if the combined increase in $L_M^{SC}(t) + L_S^{SC}(t)$ is less than $nL(t)$. Proposition 9 (below) demonstrates that the population growth rate, n, through its influence on the growth of $L_I^{SC}(t)$, plays a key role in determining how NBG influences regional population dynamics.

Proposition 9 *Along the optimal dynamic trajectory defined by* $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$*:* *1.* $\hat{L}_a^{FA}(t) - \hat{L}_b^{FA}(t) > 0 \Leftrightarrow n > \hat{L}_S^{FA}(t)$*; 2.* $\hat{L}_a^{FD}(t) - \hat{L}_b^{FD}(t) > 0 \Leftrightarrow (F_a - F_b)\left(\hat{L}_S^{FD}(t) - n\right) > 0$.

Proof. See Appendix H. ■

Proposition 9 describes the influence of NBG on regional population shares when the economy is growing along the optimal dynamic trajectory defined by $\hat{N}^{SC}(t) =$

$\hat{E}^{SC}(t)$. The intuition for parts 1 and 2 of Proposition 9 is that NBG will increase the population share of the region whose economic characteristics give it a comparative advantage in the sector(s) whose share of total employment is growing. Under full agglomeration, agglomeration externalities cause the M- and I-sectors to concentrate themselves in region a. This implies that under full agglomeration, region a has a comparative advantage in M- and I-sector production, and region b has a comparative advantage in S-sector production (under full agglomeration, if $F_a > F_b$ then region a has an absolute advantage over region b in M-, I-, and S-sector production). As such, from part 1 of Proposition 9, when the S-sector's share of total employment is expanding as the economy grows ($\hat{L}_S^{FA}(t) > n$), region a's population declines relative to region b's, and when the S-sector's share of total employment is declining ($\hat{L}_S^{FA}(t) < n$), region a's population increases relative to region b's.

Under full dispersion, the even distribution of the M- and I-sectors across the two regions implies that the region with the larger endowment of the fixed factor has a comparative (and absolute) advantage in S-sector production, while the region with the smaller endowment of the fixed factor has comparative advantage in M- and I-sector production. As such, from part 2 of Proposition 9, when the S-sector's share of total employment is expanding as the economy grows ($\hat{L}_S^{FA}(t) > n$), the share of total population in the region with the larger endowment of the fixed factor increases. Conversely, when the S-sector's share of total employment is declining ($\hat{L}_S^{FA}(t) < n$), the share of total population of the region with the smaller endowment of the fixed factor increases.

A corollary to Proposition 9 is that regions need only differ in their economic characteristics on one dimension for NBG to influence regional population shares. To see that this is the case, in part 1 of Proposition 9, under full agglomeration where the benefits of agglomeration externalities in the M- and I-sectors differ between regions, NBG results in changes in regional population shares when $F_a = F_b$. Similarly, in part 2 of Proposition 9, under full dispersion where the benefits of agglomeration externalities are the same in both regions, NBG results in changes in regional population shares provided that $F_a \neq F_b$. Part 2 of Proposition 9 also demonstrates that when regions are identical in their economic characteristics (i.e., under full dispersion and $F_a = F_b$) NBG does not influence regional population shares.

The main results concerning the implications of supply-side non-balanced growth for the relationship between agglomeration and regional population growth and the relationship between regional productivity growth and regional population growth are established in part 1 of Proposition 9. First, part 1 of Proposition 9 implies that NBG can undermine the self-reinforcing relationship between agglomeration and regional population growth that is traditionally emphasized in models of urban and regional development. When $\hat{L}_S^{FA}(t) > n$, NBG increases the employment share of the sector that benefits less from agglomeration, i.e., the S-sector. The increase in the S-sector's share of total employment causes the fixed factor, which is of relatively greater importance to the S-sector, to become salient relative to agglomeration externalities, and causes the agglomerated region (region a) to lose population relative to the less agglomerated region (region b). Conversely, when $\hat{L}_S^{FA}(t) < n$ and the M- and I-sectors share of total employment is expanding, NBG reinforces the advantage of the agglomerated region, region a, causing it to gain population relative to region b. It is noteworthy that the model can generate a negative relationship between agglomeration and regional population growth without including congestion costs or disamenities related to the spatial agglomeration of population and production, such as air and noise population, as is done in studies such as Lange and Quaas (2007).

Part 1 of Proposition 9 also illustrates how NBG attenuates the positive relationship between regional productivity and population growth emphasized in previous models of urban and regional development. When $\hat{L}_S^{FA}(t) > n$ and the S-sector's share of total employment is expanding, region a, which is concentrated both in the innovative I-sector and the progressive M-sector, loses population as a consequence of economic growth. This result suggests that NBG may cause a divergence between the determinants of national economic growth – factors that influence productivity growth, such as human and physical capital accumulation, innovation, etc. – and the determinants of regional population growth. Furthermore, this result suggests that in a deindustrializing economy with supply-side non-balanced growth, where relative employment is expanding in the less progressive service sector, this divergence between the determinants of regional productivity and population growth may be systematic, with regions whose employment is concentrated in low productivity growth, service sector industries experiencing stronger population growth. Alternatively, when $\hat{L}_S^{FA}(t) < n$ and the M- and I-sectors share of total employment is expanding, region a gains in population as a consequence of economic growth, and there is a positive relationship between regional productivity growth and regional population growth.

That the presence of supply-side non-balanced growth in the economy can undermine, and, in the case of deindustrialization, reverse, the positive relationship between regional productivity and population growth agrees with recent patterns of regional development in the United States and other advanced countries. For example, many of the fastest growing regions in the United States in the period of deindustrialization have been sun-belt cities, such as Phoenix and Las Vegas, whose economies are concentrated, by and large, in low productivity growth, service sector industries. Indeed, Glaeser and Tobio (2007) have documented that since 1980, several of the fastest growing sun-belt cities have sustained strong population growth without correspondingly strong regional productivity growth.

That there may be a systematic divergence between the determinants of regional productivity and population growth in a deindustrializing economy has implications for previous empirical studies, such as Glaeser, Scheinkman, and Shleifer (1995) and

Glaeser and Shapiro (2003), that have used population growth as a proxy for regional productivity growth in the United States over the period of deindustrialization. These previous studies regress regional population growth on a number of regional economic characteristics. The coefficients on the regressors are interpreted as capturing the influence of each economic characteristic on regional productivity growth, which is assumed to spur regional population growth. The analysis in this chapter suggests that in addition to capturing the influence of the regional economic characteristics on regional productivity growth, these regressions may also be capturing the change in the economy-wide importance of these regional economic characteristics over time that takes place as a consequence of non-balanced growth.

Parts 1 and 2 of Proposition 9 imply that regional comparative advantage determines how regional population shares change as the economy grows. The long-run distribution of population across regions, however, is determined by the interregional distribution of the economic characteristic(s) that advantage the sector that dominates the long-run distribution of employment. For example, Theorem 3 establishes that along the EGP, where the S-sector dominates the long-run distribution of employment, each region's employment share is proportional to its endowment of the fixed factor, which is the regional economic characteristic that is most important to S-sector production. Together, Theorem 3 and Proposition 9 imply that it is regional comparative advantage determines the direction of change in regional population shares in an economy undergoing NBG, while it is a regions' absolute advantage in the sector that dominates the long-run distribution of employment that determines long-run regional population shares.

1.7 Extensions

This section extends the baseline model of supply-side NBG to consider the two alternative explanations for NBG: demand-side NBG, where differential income elasticities of demand for sectoral output lead to NBG, and NBG resulting from international trade, where changes in the terms of trade between countries influence sectoral output and employment shares. Along with supply-side NBG, demand-side NBG and NBG resulting from international trade are the explanations for NBG that are most emphasized in the literature.

It is shown in this section that these two alternative explanations for NBG have different implications for the economic forces driving spatial change in the economy relative to the baseline model of supply-side NBG. Moreover, because supply-side NBG and these two alternative explanations for NBG are neither mutually exclusive nor contradictory, all three explanations for NBG may be simultaneously influencing the spatial development of an economy undergoing NBG. It is also shown that the mechanism through which NBG influences the spatial development of the economy discussed in the previous section in the context of supply-side NBG – i.e., that when there are differences in the relative importance of regional economic characteristics between the major sectors of the economy, NBG will cause economic activity and population to shift towards regions whose economic characteristics give them a comparative advantage the expanding sector(s) – also holds for the two alternative explanations for NBG considered in this section.

This section defines the static equilibrium and performs comparative static analysis for the models of the spatial consequences of demand-side NBG and NBG in an open economy, but, in the interest of space, does not define or analyze the dynamic equilibrium for either model. In addition, because the models presented in this section are extensions of the baseline model of supply-side NBG, model development is limited to areas where the new models differ from the baseline specification.

1.7.1 Demand-Side Non-Balanced Growth

This section extends the baseline model to consider the spatial consequences of demand-side NBG, where demand-side NBG results from non-homothetic household preferences. Non-homothetic household preferences imply differences in the income elasticities of demand for sectoral outputs and lead to uneven patterns of output and employment growth between sectors. It is shown that in the case of demand-side NBG, NBG results from changes in per capita expenditure that accompany economic growth, rather than differences in productivity growth between sectors. Per capita expenditure growth drives NBG because differences in the income elasticities of demand for sectoral output cause output and employment growth to be biased towards certain sectors as expenditure increases on a per capita basis (i.e., as the economy becomes wealthier). That growth in per capita expenditure drives NBG in this section is consistent with previous models of demand-side NBG, including Murphy, Shleifer, and Vishny (1989), Caselli and Coleman (2001), and Gollin, Parente, and Rogerson (2002).

The Model

The technology in the M-, S-, and I-sectors in this section is identical to the baseline model of supply-side NBG. For this reason, the model development in this section focuses on how the assumption of non-homothetic household preferences changes households' optimizing behavior. As in Chapter 1.4.1, it is assumed that household j's utility at time t is the summation of the utilities of all household members at time t, and that household j maximizes utility by spreading consumption evenly across all of its members.[38] Household j's utility at time t is

$$U\left(Y_{jM}(t), Y_{jS}(t), l_j\left(t\right)\right) = l_j\left(t\right) u\left(\frac{Y_{jM}(t)}{l_j\left(t\right)}, \frac{Y_{jS}(t)}{l_j\left(t\right)}\right), \tag{1.77}$$

where

$$u\left(\frac{Y_{jM}(t)}{l_j\left(t\right)}, \frac{Y_{jS}(t)}{l_j\left(t\right)}\right) = \ln\left[\left(\frac{Y_{jM}(t)}{l_j\left(t\right)} - \gamma\right)^{\mu}\left(\frac{Y_{jS}(t)}{l_j\left(t\right)}\right)^{1-\mu}\right]$$

is utility per household member at time t. γ is a parameter through which non-homothetic preferences are included in the model and is the only parameter that appears in this section that does not appear in the baseline model. In keeping with the baseline model, $Y_{jM}(t)$ and $Y_{jS}(t)$ are household j's total consumption of M- and S-sector output at time t (hence, $Y_{jM}(t)/l_j\left(t\right)$ and $Y_{jS}(t)/l_j\left(t\right)$ are consumption of M- and S-sector output per household member at time t), and $\mu \in (0,1)$ is a parameter that determines household j's expenditure shares for M- and S-sector output. (1.77) implies that the elasticity of substitution between M- and S-sector output is a fixed parameter equal to one.

The household j's indirect utility function at time t corresponding to the direct utility function in (1.77) is

$$\mathbb{V}_j\left(t\right) = l_j\left(t\right)\ln\left[\mu^{\mu}\left(1-\mu\right)^{(1-\mu)}\left(E_j(t) - \gamma l_j\left(t\right)\right)p(t)^{\mu-1}\right]. \tag{1.78}$$

where, as in Chapter 1.4, $E_j(t)$ is household j's expenditure at time t, $p\left(t\right)$ is the price of S-sector output at time t, and the price of M-sector output normalized to 1.[39] Household j's demands for M- and S-sector output at time t corresponding to the direct utility function in (1.78) are

$$Y_{jM}(t) = \mu E_j(t) + (1-\mu)\gamma l_j\left(t\right) \quad \text{and} \quad Y_{jS}(t) = \left(E_j(t) - \gamma l_j\left(t\right)\right)(1-\mu)p(t)^{-1}. \tag{1.79}$$

[38] Appendix A gives a formal proof that household j maximizes utility at time t by spreading consumption evenly across all of its members for the case of non-homethetic preferences considered in this section.

[39] (1.78) and (1.79) are dervied in Appendix B.

In the remainder of this section it is assumed that $-\frac{\mu}{1-\mu}\frac{E_j(t)}{l_j(t)} < \gamma < \frac{E_j(t)}{l_j(t)}$. This condition guarantees that $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$ for all t.[40]

To see how non-homothetic preferences result in NBG, consider the expressions

[40]The specification for utility per household member in (1.77) does not guarantee that household j will consume positive quantities of M- and S-sector output. In particular, the Inada conditions that guaranteed that household j will consume positive quantities of M- and S-sector output in the baseline model do not hold for (1.77) when $\gamma \neq 0$. As the analysis focuses on NBG, where both the M- and S-sectors are in operation, the corner solutions where household j's demands for M- or S-sector output are equal to zero are not of interest. For this reason, it is necessary to place two assumption on $E_j(t)$ and $l_j(t)$ in the remainder of this chapter that guarantee that $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$ for all t.

The first condition is that $E_j(t)/l_j(t) > \gamma$. To see why this condition guarantees $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$, notice that when $\gamma > 0$, (1.77) is only defined for $Y_{jM}(t)/l_j(t) > \gamma$. A sufficient condition for $Y_{jM}(t)/l_j(t) > \gamma$ is for household expenditure per member, $E_j(t)/l_j(t)$, to be greater than γ. When this condition holds, the Inada conditions apply and household j will demand positive quantities of both M- and S-sector output.

The second condition is that $\gamma > -\frac{\mu}{1-\mu}\frac{E_j(t)}{l_j(t)}$. When $\gamma < 0$, a sufficient condition for $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$ is that the marginal rate of substitution between S-sector output and M-sector output to be greater than the relative price of M-sector output when $Y_{jM}(t) = 0$ (note that when $\gamma < 0$ the Inada condition ensures that $Y_{jS}(t) > 0$ for all t). The marginal rate of substitution between S-sector output and M-sector output for household j at time t is

$$-\frac{dY_{jS}(t)}{dY_{jM}(t)} = \frac{\partial U\left(Y_{jM}(t), Y_{jS}(t), l_j(t)\right)/\partial Y_{jM}(t)}{\partial U\left(Y_{jM}(t), Y_{jS}(t), l_j(t)\right)/\partial Y_{jS}(t)} = \frac{\mu}{1-\mu}\frac{Y_{jS}(t)/l_j(t)}{Y_{jM}(t)/l_j(t)-\gamma}$$

Given this expression, the condition for $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$ at time t when $\gamma < 0$ is

$$-\frac{\mu}{1-\mu}\frac{Y_{jS}(t)/l_j(t)}{\left(Y_{jM}(t)/l_j(t)\right)-\gamma}\bigg|_{Y_{jM}(t)=0} = \frac{\mu}{1-\mu}\frac{E_j(t)}{l_j(t)\,p(t)}\frac{1}{-\gamma} > \frac{1}{p(t)}. \tag{1.80}$$

Rearranging terms, Equation 1.80 implies that when $\gamma < 0$, $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$ if $\gamma > -\frac{\mu}{1-\mu}\frac{E_j(t)}{l_j(t)}$. When this condition holds, the amount that household j's consumption of S-sector output would have to increase to compensate it for the loss of a unit of M-sector output when $Y_{jM}(t) = 0$ would be greater than the relative price of M-sector output, $p(t)^{-1}$. This means that when $0 > \gamma > -\frac{\mu}{1-\mu}\frac{E_j(t)}{l_j(t)}$, $Y_{jM}(t) = 0$ would be inconsistent with household j's preferences.

for the income elasticities of demand for M- and S-sector output implied by (1.79):

$$\frac{d\ln Y_{jM}(t)}{d\ln E_j(t)} = \frac{\mu E_j(t)}{\mu E_j(t) + (1-\mu)\gamma l_j(t)} > 1 \Leftrightarrow \gamma < 0 \tag{1.81}$$

$$\frac{d\ln Y_{jS}(t)}{d\ln E_j(t)} = \frac{E_j(t)}{E_j(t) - \gamma l_j(t)} > 1 \Leftrightarrow \gamma > 0 \quad . \tag{1.82}$$

(1.81) and (1.82) show that non-homothetic preferences lead to different income elasticities of demand for sectoral output. When $\gamma > 0$, an increase in household j's expenditures leads to a less than proportional increase in the consumption of $Y_{jM}(t)$ and a greater than proportional increase in the consumption of $Y_{jS}(t)$; when $\gamma < 0$, the converse holds. When $\gamma = 0$ and household preferences are homothetic, household j's consumption of $Y_{jM}(t)$ and $Y_{jS}(t)$ increase proportionately with its expenditure.

These differences in income elasticities of demand imply that when preferences are non-homothetic ($\gamma \neq 0$), an increase in household j's per capita expenditure (e.g., an increase in $E_j(t)$ for fixed $l_j(t)$) will change the composition of its demand:

$$\frac{d(Y_{jM}(t)/Y_{jS}(t))}{dE_j(t)} = \frac{-\gamma l_j(t)}{(1-\mu)\, p(t)^{-1} \left(E_j(t) - \gamma l_j(t)\right)^2} > 0 \Leftrightarrow \gamma < 0 \ . \tag{1.83}$$

From (1.83), when $\gamma > 0$, an increase in per capita expenditure shifts the composition of household j's demand in favor of $Y_{jS}(t)$; when $\gamma < 0$, an increase in per capita expenditure shifts the composition of household j's demand in favor of $Y_{jM}(t)$; and when household j's preferences are homothetic and $\gamma = 0$, an increase in per capita expenditure does not change the composition of household j's demand. It is the influence of increases in per capita expenditure on the composition of household demand, and hence on the composition of aggregate demand, that determines the direction of NBG in this section.

Static Equilibrium

The definition of the static equilibrium in this section is the same as the definition in Chapter 1.5. In addition, the static equilibrium conditions in this section are identical to the static equilibrium conditions from Proposition 1 and Theorems 1 and 2 in Chapter 1.5 with the exception of the expressions for M- and S-sector employment and for the relative price of S-sector output, which are different because the new expressions for the aggregate demands for M- and S-sector output implied by (1.79) change the market-clearing conditions in the M- and S-sectors. In particular, the new expressions for the aggregate demands for M- and S-sector output are

$$\sum_{j=1}^{H} Y_{jM}(t) = \mu \sum_{j=1}^{H} E_j(t) + (1-\mu)\gamma \sum_{j=1}^{H} l_j(t) = \mu E(t) + (1-\mu)\gamma L(t) \tag{1.84}$$

and

$$\sum_{j=1}^{H} Y_{jS}(t) = \left(\sum_{j=1}^{H} E_j(t) - \gamma \sum_{j=1}^{H} l_j(t) \right) (1-\mu) p(t)^{-1} = (E(t) - \gamma L(t))(1-\mu) p^{SC}(t)^{-1},$$

for $SC = FA, FD$, where, as in Chapter 1.4, $E(t)$ is societal expenditure and $L(t)$ is total labor supply.[41] Note that the conditions on $E_j(t)$ and $l_j(t)$ that ensure that $Y_{jM}(t) > 0$ and $Y_{jM}(t) > 0$ hold for every household in the economy for all t imply that aggregate demands for M- and S-sector output in (1.84) are positive for all t.

Using the expression for aggregate demand for M-sector output from (1.84) and the M-sector market-clearing condition from (1.32), M-sector employment at time t is

$$L_M^{SC}(t) = \frac{\mu E(t) + (1-\mu)\gamma L(t)}{(1+\alpha) w^{SC}(t)}, \tag{1.85}$$

where, from (1.37), $w^{FA}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}} N(t)$ and $w^{FD}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}} \frac{1}{2}\left(1 + \Gamma^{\frac{\alpha}{a-1}}\right) N(t)$. Similarly, using the expression for aggregate demand for S-sector output from (1.84), and the S-sector market-clearing condition from (1.39), the relative price of S-sector

[41] Household demand functions can be aggregated to form societal demand functions because all households have the same linear Engel curves ($dY_{jM}(t)/dE_j(t) = \mu$ and $dY_{jS}(t)/dE_j(t) = (1-\mu)p(t)^{-1}$ for all $j = 1, ..., H$ and for all t). Linear Engel curves imply that changes in household demand that would result from a redistribution of a given level of societal expenditure between households would cancel out, leaving aggregate demands for M- and S-sector output unaffected.

output at time t is

$$p^{SC}(t) = \left[\frac{(1-\mu)(E(t) - \gamma L(t)) w^{SC}(t)^{\frac{\beta}{1-\beta}}}{\beta^{\frac{\beta}{1-\beta}} (F_a + F_b)} \right]^{1-\beta}, \tag{1.86}$$

and, using the expression for S-sector employment from (1.26), S-sector employment at time t is

$$\begin{aligned} L_S^{SC}(t) &= L_{aS}^{SC}(t) + L_{bS}^{SC}(t) \qquad (1.87) \\ &= \left[\frac{\beta(1-\mu)(E(t) - \gamma L(t))}{w^{SC}(t)(F_a + F_b)} \right] F_a + \left[\frac{\beta(1-\mu)(E(t) - \gamma L(t))}{w(t)(F_a + F_b)} \right] F_b \\ &= \frac{\beta(1-\mu)(E(t) - \gamma L(t))}{w^{SC}(t)}. \end{aligned}$$

The expressions for $L_M^{SC}(t)$ and $L_S^{SC}(t)$ from (1.85) and (1.87) change the static equilibrium expressions for M- and S-sector output from Proposition 1. The expression for M- and S-sector output under the assumption of non-homothetic preferences are not presented here because they are not used in the analysis and discussion in the remainder of this section.

Comparative Static Analysis

This section presents comparative static results for the influence of demand-side NBG on relative employment on the M- and S-sectors and on regional population movements. As in Chapter 1.5.3, comparative static results are presented for two spatial configurations of the economy: full agglomeration, where the M- and I-sector labor and I-sector innovation are concentrated in region a, and full dispersion, where the M- and I-sectors are evenly dispersed across the two regions.

Proposition 10 *In equilibrium, when preferences are non-homothetic, 1.* $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln N(t)} = 0$*; 2.* $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln E(t)} > 0 \Leftrightarrow \gamma < 0$*; and 3.* $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln L(t)} > 0 \Leftrightarrow \gamma > 0$, $SC = FA, FD$.

Proof. Proposition 10 follows directly from the expressions for $L_M^{SC}(t)$ and $L_S^{SC}(t)$ from (1.85) and (1.87). ■

Part 1 of Proposition 10 implies that technological advance (i.e., growth in $N(t)$) does not influence M-sector employment relative to S-sector employment in this section. The intuition for why this is the case is given in the discussion of Proposition 4 in Chapter 1.5.3. In particular, Proposition 4 demonstrates that when the elasticity of substitution between M- and S-sector output is equal to 1 (which is the case in (1.77)), the increase in the marginal product of labor in the M-sector resulting from an increase in $N(t)$, which attracts labor to the M-sector, is precisely proportional to the increase in the relative price of S-sector output, $p(t)$, which attracts labor to the S-sector, leaving $L_M^{SC}(t)/L_S^{SC}(t)$ unchanged.

Parts 2 and 3 of Proposition 10 demonstrate that sectoral employment growth as a result of an increase in per capita expenditure will be biased towards the sector whose output has a higher income elasticity of demand. This result holds regardless of whether the increases in per capita expenditure occur as a result of increases in $E(t)$ for fixed $L(t)$ (part 2 of Proposition 10) or a decrease in $L(t)$ for fixed $E(t)$ (part 3 of Proposition 10). When $\gamma > 0$, an increase in per capita expenditure leads to a greater than proportional increase in consumption of S-sector output and a less than proportional increase in consumption of M-sector output (see (1.81) and (1.82)), which causes the S-sector to gain employment relative to the M-sector. The converse result holds when $\gamma < 0$. When $\gamma = 0$ and consumption of both M- and S-sector output increases proportionately with expenditure, $L_M^{SC}(t)/L_S^{SC}(t)$ does not change with increases in $E(t)$ or $L(t)$.

Re-expressing $L_M^{SC}(t)$ and $L_S^{SC}(t)$ from (1.85) and (1.87) in terms of per capita expenditure, $e(t) = E(t)/L(t)$, yields

$$L_M^{SC}(t) = \frac{[\mu e(t) + (1-\mu)\gamma]\,L(t)}{(1+\alpha)\,w^{SC}(t)} \text{ and } L_S^{SC}(t) = \frac{\beta(1-\mu)(e(t)-\gamma)\,L(t)}{w^{SC}(t)}. \quad (1.88)$$

From (1.88), $L_M^{SC}(t)$ and $L_S^{SC}(t)$ change proportionately with $L(t)$ for constant $e(t)$. This implies that holding $e(t)$ constant, changes in $L(t)$ do not affect $L_M^{SC}(t)/L_S^{SC}(t)$.

Parts 2 and 3 of Proposition 10 demonstrate when household preferences are non-homothetic, changes in per capita expenditure lead to changes in the composition of household demand for M- and S-sector output, and that these changes in the composition of household demand lead to NBG. It is the case, of course, that sustained growth in per capita expenditure is only possible in a growing economy. This means that demand-side NBG in this section is ultimately driven by growth in $N(t)$, which is the source of productivity growth in the model. The direction of NBG in this section, however, is determined by the income elasticities of demand for sectoral output, and not by differences in productivity growth between sectors, as was the case in the baseline model of supply-side NBG.

Proposition 11 *In equilibrium, when preferences are non-homothetic,*

1. $\frac{d\ln L_a^{FA}(t)}{d\ln N(t)} > 0$, $\frac{d\ln L_a^{FA}(t)}{d\ln e(t)} < 0$, *and* $\frac{d\ln L_a^{FA}(t)}{d\ln L(t)} = 1$*;*

2. $\frac{d\ln L_a^{FD}(t)}{d\ln N(t)} > 0 \Leftrightarrow F_a - F_b < 0$, $\frac{d\ln L_a^{FD}(t)}{d\ln e(t)} > 0 \Leftrightarrow F_a - F_b > 0$, *and* $\frac{d\ln L_a^{FD}(t)}{d\ln L(t)} = 1$*;*

$SC = FA, FD$.

Proof. Proposition 11 follows directly from the expression for $L_{kS}^{SC}(t)$, $k = a, b$, from (1.88) and the expressions for region a's population under full agglomeration and full dispersion from (1.53) and (1.54). ■

Proposition 11 demonstrates that, as is the case with supply-side NBG, demand-side NBG will cause employment, and, hence, population, to move into the region whose economic characteristics give it a comparative advantage in the expanding sector(s). In particular, Proposition 11 demonstrates that an increase in $N(t)$ influences the distribution of population between regions a and b because, from (1.85) and (1.87), an increase in $N(t)$ reduces $L_M^{SC}(t)$ and $L_S^{SC}(t)$, and, thus, increases $L_I^{SC}(t)$.[42] Under full agglomeration, an increase in $N(t)$ causes region a's population share to increase because agglomeration externalities cause the expanding I-sector to concentrated itself in region a (i.e., because region a has a comparative advantage in I-sector production). Under full dispersion, an increase in $N(t)$ expands the population share of the region with the smaller endowment of the fixed factor. The region with the smaller endowment of the fixed factor has a comparative advantage in M- and I-sector production relative to S-sector production because the benefits from agglomeration externalities to the M- and I-sectors are the same in regions a and b under full dispersion.

Similarly, Proposition 11 demonstrates that an increase in per capita expenditure, $e(t)$, changes the distribution of population between regions a and b because, from (1.88), an increase in $e(t)$ increases $L_M^{SC}(t)$ and $L_S^{SC}(t)$, and, thus, reduces $L_I^{SC}(t)$. Under full agglomeration, an increase in $e(t)$ increases S-sector employment in region b, and, thus, region b's population, because region b has a comparative advantage in S-sector production. Under full dispersion, where the benefits from agglomeration externalities in the M- and I-sectors are the same in both regions, the region with the larger endowment of the fixed factor has a comparative advantage in S-sector production and will experience the larger increase in S-sector employment – and will thus expand its population – as a result of an increase in $e(t)$.

Proposition 11 also demonstrates that increases in $L(t)$ do not influence regional population shares. This is because, from (1.88), an increase in $L(t)$ leads to proportional increases in employment in the M-, S- and I-sectors for fixed $N(t)$ and $e(t)$,

[42] As in Proposition 3 in Chapter 1.5.3, an increase in $N(t)$ reduces $L_M^{SC}(t)$ and $L_S^{SC}(t)$ becuase the improvement in M-sector productivity resulting from the increase in $N(t)$ reduces the total labor required to meet the quantity of M- and S-sector output demanded by consumers for a given level of societal expenditure, $E(t)$.

and thus does not influence regional population shares.

Proposition 11 illustrates the limits to comparative static analysis with regards to the interregional population distribution. In particular, Proposition 11 demonstrates that $N(t)$ and $e(t)$ have countervailing influences on regional population movements, so that the influence of NBG on regional population movements will depend on the growth rates of both $N(t)$ and $e(t)$. Despite this limitation, the comparative static results presented in Propositions 10 and 11 suggest that many of the key results concerning the implications of supply-side NBG for regional population movements developed in the dynamic analysis in Chapter 1.6.3 will hold in the case of demand-side NBG. In particular, Proposition 10 indicates that when $\gamma > 1$, an increase in per capita expenditure will cause the employment share of the less progressive S-sector to expand. Proposition 11 demonstrates that this will cause the population share of the region with a comparative advantage in S-sector production to grow, and will thus undermine the positive relationship between regional productivity and regional population growth emphasized in the previous literature and discussed in Chapter 1.6.3. In addition, the sector that is expanding in employment benefits less from agglomeration externalities, so that demand-side NBG will undermine the self-reinforcing relationship between agglomeration and regional population growth emphasized in the previous literature and discussed in Chapter 1.6.3.

1.7.2 Non-Balanced Growth in an Open Economy

This section extends the baseline model to consider the spatial consequences of NBG in an open economy, where NBG results from changes in the terms of trade between countries. International trade will cause the process of NBG to unfold differently across countries because, as has been pointed out in Matsuyama (2007), faster productivity growth in a given sector will shift a country's comparative advantage towards that sector. In this case of deindustrialization, this means that while faster productivity growth in goods production relative to services in a given country may lead to deindustrialization globally, it may not cause employment in goods-producing sectors in the country to decline. International trade provides an explanation for why deindustrialization has been uneven across advanced economies, while occurring across advanced economies as a whole. For example, Rowthorn and Ramaswamy (1999) find, when looking at changes in service sector employment relative to manufacturing employment, that between 1960 and 1994 the total number of workers engaged in manufacturing across advanced economies as a whole remained roughly constant, while employment in the service sector grew at an average annual rate of 2.2 percent. Against this backdrop of a declining employment share for manufacturing across advanced economies, Obstfeld and Rogoff (1996) show that Germany and Japan have experienced smaller declines in manufacturing's employment share than the United States and that certain advanced Pacific Rim economies, such as Hong Kong, Taiwan, and South Korea, have had their manufacturing sector's share of employment continue to rise through the early 1990s.

The analysis in this section considers the spatial consequences of NBG in a small open economy. This means that production and investment decisions in the domestic economy are assumed not to influence world prices, as would be the case in a large open economy. How the results presented in this section would change if the analysis were extended to consider a large open economy is discussed below.

The Model

There are two differences between the model used in this section to analyze the spatial consequences of NBG in an open economy and the baseline model of supply-side NBG: (i) both M- and S-sector output are traded internationally and (ii) households and firms have access to international capital markets. These assumptions imply that both $p(t)$ and $r(t)$ are set internationally (henceforth $p^{int}(t)$ and $r^{int}(t)$). Intermediate inputs are not traded internationally in this chapter, though they are traded between regions. This preserves the mechanism that creates the incentive for agglomeration in the M- and I-sectors from the baseline model. Furthermore, I-sector innovation can only develop new varieties of intermediate inputs for the domestic economy, and new varieties of intermediate inputs developed in foreign economies cannot be imported.

The exogenously determined international prices, $p^{int}(t)$ and $r^{int}(t)$, imply new expressions for M-, S-, and I-sector employment when the economy is in static equilibrium. To begin, using (1.26), the expression for S-sector employment in an open economy at time t is

$$L_S^{SC}(t) = L_{aS}^{SC}(t) + L_{bS}^{SC}(t) = \left(\frac{\beta p^{int}(t)}{w^{SC}(t)}\right)^{\frac{1}{1-\beta}} F_a + \left(\frac{\beta p^{int}(t)}{w^{SC}(t)}\right)^{\frac{1}{1-\beta}} F_b, \qquad (1.89)$$

for $SC = FA, FD$, where, from (1.37), $w^{FA}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}} N(t)$ and $w^{FD}(t) = (1-\alpha)\alpha^{\frac{2\alpha}{1-\alpha}}\frac{1}{2}\left(1+\Gamma^{\frac{\alpha}{\alpha-1}}\right)N(t)$. Next, using (1.43) and (1.57), the expression for M-sector employment in an open economy at time t is

$$L_M^{FA}(t) = \alpha^{-1} r^{int}(t) \text{ and } L_M^{FD}(t) = \left(\frac{2}{1+\eta}\right)\alpha^{-1} r^{int}(t). \qquad (1.90)$$

Finally, given the expressions for $L_S^{SC}(t)$ and $L_M^{SC}(t)$ from (1.89) and (1.90), $L_I^{SC}(t)$ is determined by market-clearing condition in the labor market, (1.31).[43]

(1.90) implies that in an open economy, the distribution of non-S-sector employment (i.e., $L(t) - L_S^{SC}(t)$) between M-sector production and I-sector innovation is

[43] The analysis in this section focuses on the case where M- and S-sector production and I-sector innovation in operation. The analysis focuses on this case becuase NBG requires that both the M- and S-sectors are in operation and that there is economic growth in the economy, i.e., that I-sector innovation produces new varieties of intermediate inputs. Two conditions on $p^{int}(t)$ and $r^{int}(t)$ are necessary for M-sector production and I-sector innovation to be in operation for all t. Notice that no additional conditions are necessary to guarantee that S-sector is in operation for all t becuase the presence of the fixed factor in each region means that the marginal product of labor in the S-sector will go to infinite in a given region – and, hence, in the domestic economy as a whole – as S-sector employment in the region goes to zero.

First, it is necessary to place a restriction on the world price of S-sector output, $p^{int}(t)$, to determined by the interest rate, $r^{int}(t)$. The intuition for why $L_M^{SC}(t)$ is increasing in $r^{int}(t)$ is that, all else equal, an increase in $r^{int}(t)$ will cause employment to move from I-sector innovation into the M-sector production until demand for intermediate inputs from the M-sector, and, hence, the profits of I-sector firms, has increased to the point that there is once again an incentive to invest in the development of new varieties of intermediate inputs (i.e., an incentive for positive employment in the I-sector). Higher values of $r^{int}(t)$ imply that more employment will move from the I-sector to the M-sector before this point is reached (and, hence, higher total levels of M-sector employment).

Comparative Statics

Proposition 12 *In equilibrium, in a small open economy,*

1. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln N(t)} > 0$, $\frac{d\ln L_a^{FA}(t)}{d\ln N(t)} > 0$, *and* $\frac{d\ln L_a^{FD}(t)}{d\ln N(t)} > 0 \Leftrightarrow F_a - F_b < 0$;
2. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln E(t)} = \frac{d\ln L_a^{FA}(t)}{d\ln E(t)} = \frac{d\ln L_a^{FD}(t)}{d\ln E(t)} = 0$; *and*

---

guarantee that the M-sector is in operation at time t. Notice that if the M-sector is not in operation at time t, then I-sector production, which furnishes intermediate inputs to the M-sector, will also not be in operation at time t. The restriction is that the marginal revenue product of labor in the S-sector is lower than the marginal revenue product of labor in the M-sector (which is equal to $w^{SC}(t)$) when all labor is devoted to S-sector production, i.e., $p^{int}(t)\,\beta\left(\frac{F_a+F_b}{L(t)}\right)^{1-\beta} < w^{SC}(t)$. This restriction guarantees that $p^{int}(t)$ will not be so high as to perclude domestic production in the M-sector.

Second, from (1.42) and (1.56), a sufficient condition for I-sector innovation to be in operation at time t (i.e., $L_I^{SC}(t) > 0$ at time t) is $\pi_i^{SC}(t) = (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}}\left[\lambda_M^{SC}(t) + \Gamma^{\frac{\alpha}{a-1}}\left(1-\lambda_M^{SC}(t)\right)\right]\left(L(t) - L_S^{SC}(t)\right) > V^{SC} r^{int}(t)$. This condition implies that when all labor not used in S-sector production is directed to the M-sector, that the profits for I-sector firms are sufficiently high that there is an incentive to invest in the development of new varieties of intermediate inputs. Indeed, when this ineqaulity holds, labor will be allocated from the M-sector to the I-sector until $V^{SC} r^{int}(t) = \pi_i^{SC}(t)$ holds with equality.

3. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln L(t)} = 0, \quad \frac{d\ln L_a^{FA}(t)}{d\ln L(t)} = \frac{L(t)}{L_a^{FA}(t)} > 0,$ *and* $\frac{d\ln L_a^{FD}(t)}{d\ln L(t)} = \frac{1}{2}\frac{L(t)}{L_a^{FA}(t)} > 1 \Leftrightarrow$
$F_a - F_b > 0$*;*

where $SC = FA, FD$*.*

Proof. Proposition 12 follows directly from the expressions for $L_S^{SC}(t)$ and $L_M^{SC}(t)$ from (1.89) and (1.90), and the expressions for population in regions a and b under full agglomeration and full dispersion from (1.53) and (1.54). ■

Part 1 of Proposition 12 shows that in an open economy, productivity growth in the M-sector caused by technological advance (change in $N(t)$) shifts a country's comparative advantage towards the M-sector. This causes M-sector employment to increase and causes the population of the region whose economic characteristics give it a comparative advantage in M-sector production to increase, i.e., region a when the economy is in full agglomeration and the region with less of the fixed factor when the economy is in full dispersion. Technological advance unambiguously leads to an increase in M-sector employment in a small open economy because it is assumed that domestic technological advance does not affect world prices, so that there is not the countervailing influence of an increase in $p(t)$ drawing labor into the S-sector that accompanies increases in $N(t)$ in the baseline model of a closed economy.

Parts 2 and 3 of Proposition 12 follow from the fact that, from (1.90) and (1.89), changes in $E(t)$ and $L(t)$ do not influence employment in the M- and S-sectors, and, as such, do not influence either relative employment in the M- and S-sectors or relative population in regions a and b.

Proposition 12 illustrates the key insight in this section that the direction of NBG in an open economy is determined by how a country's comparative advantage, rather than its absolute advantage, changes over the process of development. In the case of deindustrialization, this means that the extent of deindustrialization within a country is determined by the performance of the domestic goods-producing sector compared to its global competitors relative to the performance of the domestic service sector compared to its global competitors. Put differently, in an open economy it is the differential impact of global competition on each sector that determines the direction and extent of NBG. This means that even if deindustrialization is occurring globally as a result of faster productivity growth in the goods-producing sector relative to services, the extent of deindustrialization, and, hence, the extent of its spatial consequences, will be different across countries.

Proposition 13 *In equilibrium, in an open economy,*

1. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln p^{int}(t)} < 0,\ \frac{d\ln L_a^{FA}(t)}{d\ln p^{int}(t)} < 0,$ *and* $\frac{d\ln L_a^{FD}(t)}{d\ln p^{int}(t)} > 0 \Leftrightarrow F_a - F_b > 0;$
2. $\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln r^{int}(t)} > 0,\ \frac{d\ln L_a^{FA}(t)}{d\ln r^{int}(t)} > 0,$ *and* $\frac{d\ln L_a^{FD}(t)}{d\ln r^{int}(t)} > 0 \Leftrightarrow F_a - F_b < 0;$

where $SC = FA, FD$.

Proof. Proposition 13 follows directly from the expressions for $L_S^{SC}(t)$ and $L_M^{SC}(t)$ from (1.89) and (1.90), and the expressions for population in regions a and b under full agglomeration and full dispersion from (1.53) and (1.54). ■

Proposition 13 demonstrates that in an open economy, a change in $p^{int}(t)$ will cause employment to expand in the sector whose relative output price is increasing and that this will lead to the region whose regional economic characteristics give it a comparative advantage the expanding sector's production to increase in population. Similarly, an increase in $r^{int}(t)$ will reduce investment in I-sector innovation and cause M-sector employment to expand. This will cause the population of the region whose economic characteristics give it a comparative advantage in M-sector production to increase, i.e., region a when the economy is in full agglomeration and the region with less of the fixed factor when the economy is in full dispersion.

The importance Proposition 13 is that it holds regardless of the cause of the change in world prices. In particular, the change in world prices could result from differences in productivity growth between sectors (supply-side non-balanced growth) or differences in the income elasticities for sectoral demand (demand-side non-balanced growth) occurring in either the domestic economy or in any of the domestic economy's trading partners. In this way, Proposition 13 demonstrates that economic activity in one country can "export" NBG and its spatial consequences to other countries. In addition, Proposition 13 implies that NBG can occur in an open economy without there being differences in productivity growth between sectors or differences in the income elasticities of demand for sectoral output in the domestic economy provided that these differences exist in one or more of the domestic economy's trading partners.

It is straightforward to extend the comparative static results presented in Proposition 12 and 13 to a large open economy. The most important implication that would follow from expanding the analysis to consider a large open economy is that in a large open economy changes in the domestic economy will cause NBG to occur asymmetrically across countries. For example, in a large open economy, an increase in $N(t)$ would change the domestic economy's comparative advantage, as is shown in Proposition 12, and would also influence $p^{int}(t)$ and $r^{int}(t)$. As is shown in Proposition 13, changes in $p^{int}(t)$ and $r^{int}(t)$ will influence sectoral and regional employment in both the domestic economy and in foreign economies; however, the changes in sectoral and regional employment in the foreign economies will occur without the offsetting increase in $N(t)$. This means that the consequences of an increase in $N(t)$ for sectoral and regional employment will differ between the domestic economy and foreign economies, i.e., that in a large open economy an increase in $N(t)$ will cause NBG to occur asymmetrically across countries.

1.8 Conclusions and Directions for Further Research

This essay analyzed the consequences of NBG at the national level for the regional distribution of population and production. To analyze this issue, this essay developed a multi-sector, multi-region, dynamic general-equilibrium model of regional development in the presence of non-balanced growth. This model extends the previous literature by allowing for endogenous labor allocations across both sectors and regions within an analytically tractable growth model and in the presence of agglomeration externalities. It is argued that these innovations are essential for analyzing the spatial consequences of non-balanced growth, and, in particular, of deindustrialization. The model is used to consider the implications for regional development of the three explanations for NBG most emphasized in the literature: supply-side NBG (Chapters 1.4 to 1.6), demand-side NBG (Chapter 1.7.1), and NBG resulting from international trade (Chapter 1.7.2).

The theoretical analysis in this essay suggests a simple mechanism for how NBG may influence the spatial distribution of the economy. According to this mechanism, when there are differences in the relative importance of regional economic characteristics between the sectors of the economy, NBG will cause population to shift towards regions whose economic characteristics give them a comparative advantage in the sector(s) whose share of total employment is expanding. The major theoretical results in this essay follow from the economic forces driving NBG also influencing regional economic development and interregional population movements through this mechanism. In addition, the model implies that while it is regions' comparative advantage that determines the direction of change in regional population shares in an economy undergoing NBG, it is regions' absolute advantage in the sector(s) that dominates the long-run distribution of employment that determines long-run distribution of population between regions.

The theoretical results demonstrate that the presence of NBG at the national level can influence several of the relationships between regional economic activity and regional population growth that are emphasized in the previous literature. In particular, it is shown that both supply-side and demand-side NBG can undermine the self-reinforcing relationship between agglomeration and regional population growth and the positive relationship between regional productivity growth and regional population growth that have been predicted in previous studies. In addition, in the case where deindustrialization results from slower productivity growth in the service sector (i.e., from supply-side NBG), it is shown that the positive relationship between regional productivity growth and regional population growth may be systematically reversed. It is argued that this result agrees with the recent experience of the United States and other advanced economies undergoing deindustrialization, where many of the regions that have sustained strong population growth have regional economies that are concentrated in relatively low-productivity growth industries in the service sector.

The analysis in this essay suggests several directions for further research. For example, the analysis demonstrates that NBG can lead to systematic differences between regional productivity growth and regional population growth, with regions concentrated in high productivity growth sector(s) experiencing relatively weak population growth and vice versa. This result suggests that further research into how NBG can lead to interregional differences in the process of economic development may be merited. For example, if the model in this essay were extended to included physical capital, NBG may also lead to systematic differences in the rates of population growth and capital accumulation between regions. These differences would express themselves in interregional variation in capital-to-labor ratios.

The analysis in this essay suggests that international trade can cause the NBG to occur asymmetrically across countries, and that this may lead to the process of regional development unfolding asymmetrically across countries as well. This result suggests that future research into how differences in patterns of regional development between countries are related to differences in the process of NBG between countries is merited. In addition to cross-country differences in the process of NBG resulting from international trade, differences in the process of NBG between countries due to cross-country differences in factors such as human and physical capital accumulation, the distribution of income, and natural resource endowments could also be considered.

Finally, the analysis in this essay considers how regional economic characteristics influence firm productivity, and, hence, firm location decisions, but does not consider how regional economic characteristics influence household utility or location decisions. Numerous studies, beginning with Roback (1982), have shown that households will accept a lower wage and a higher rent in order to live in a region with desirable economic characteristics (e.g., desirable climate amenities, high environmental quality, low congestion). Given this previous work, a potential topic for further research is whether NBG changes the economy-wide importance of the regional economic characteristics that influence household utility relative to those that influence firm productivity. This research could shed light on the findings in previous empirical studies, such as Cheshire and Magrini (2006), that in recent decades climate amenities (low rainfall, warm winters, etc.) have been important determinants of not only regional population levels but also of regional population growth in advanced economies. This empirical finding may be explained by deindustrialization if it has increased the overall importance of climate amenities in household and firm location decisions relative to the regional economic characteristics that originally informed the location decisions of firms in declining industries in goods-producing sectors (e.g., proximity to inexpensive sources of energy or low transportation costs for physical output).

Chapter 2

Non-Balanced Growth in the United States: Evaluating Supply-Side versus Demand-Side Explanations

2.1 Introduction

The term *non-balanced growth* (NBG) refers to systematic changes in the employment and output shares of the major sectors of the economy over time in the process of economic development.[1] In reference to advanced economies, the term NBG generally refers to what Kongsamut, Rebelo, and Xie (2001) term the "Kuznets facts": the decline in the relative importance of the agricultural sector in terms of employment and output shares in the early stages of development, and the successive rise in the

[1] Certain studies have used the term "structural change" rather than NBG to refer to changes in sectoral employment and output shares that occur in the process of economic development (e.g., Ngai and Pissarides 2007). This essay uses the term NBG rather than structural change because, as has been pointed out in Matsuyama (2007), in addition to changes in the sectoral composition of output and employment, the term structural change has also been used refer to changes in the distribution of income and wealth, changes in demographics (population age distribution, etc.), and changes in institutions, such as the financial system, the organization of industry, and political institutions, that occur in the process of economic development.

relative importance of the manufacturing and service sectors. In the United States and other advanced economies, however, NBG consistent with the Kuznets facts has occurred against the backdrop of otherwise balanced economic growth. In particular, the Kuznets facts have been observed concurrently with the "Kaldor facts": the relative constancy of (i) the growth rate output, (ii) the real interest rate, (iii) the capital-to-output ratio, and (iv) the share of capital income in total output (see Barro and Sala-i-Martin 2004).

Several studies have attempted to reconcile NBG with the Kaldor facts (Kongsamut, Rebelo, and Xie 2001; Ngai and Pissarides 2007; Acemoglu and Guerrieri 2008; Foellmi and Zweimuller 2008). These studies have proposed theoretical models of economic growth that generate equilibrium growth paths that are consistent with the Kaldor facts and also feature NBG at the sector level. These studies, however, have focused on either supply-side NBG, where NBG results from differential rates of productivity growth across sectors (see Ngai and Pissarides 2007; Acemoglu and Guerrieri 2008), or demand-side NBG, where NBG results from changes in the composition of demand for sectoral output that occur as the economy grows (see Kongsamut, Rebelo, and Xie 2001; Foellmi and Zweimuller 2008).[2] The exclusive focus on only one of these two complementary explanations for NBG is problematic given that recent work has demonstrated the inability of either explanation, on its own, to generate patterns of NBG that match the dramatic changes in sectoral output and employment shares associated with the Kuznets facts for the United States and other advanced economies (Buera and Kaboski 2009; Iscan 2010; Guillo, Papageorgiou, and Perez-Sebastian 2011).

This essay develops and calibrates a dynamic general equilibrium model that in-

[2] Supply-side non-balanced growth is also referred to as *technology-based* non-balanced growth. Demand-side non-balanced growth is also referred to *preference-based* non-balanced growth.

tegrates supply-side and demand-side explanations for NBG, and uses it to evaluate the extent to which the two explanations for NBG – both alone and in combination – can account for patterns of NBG consistent with the Kuznets and Kaldor facts in the post-war United States (1949 to 2012). The model is calibrated using data from the National Income and Product Accounts (NIPA) from the U.S. Bureau of Economic Analysis (BEA). The analysis focuses on the model's ability to generate equilibrium dynamics that feature NBG consistent with the Kuznets facts and an approximately constant growth rate for output, real interest rate, capital-to-output ratio, and capital's share of total output, consistent with the Kaldor facts. The analysis examines the potential trade-offs between the model's ability to match the large changes in sectoral output and employment shares observed in the United State over the study period and its ability to generate a balanced growth path for the economy as described by the Kaldor facts.

In order to evaluate the Kuznets facts, the model divides the economy into three sectors: agriculture, manufacturing, and services. In doing so, the analysis builds on previous three-sector models of NBG. In the two studies closest to this essay, Buera and Kaboski (2009) and Iscan (2010) develop three-sector models that include the agricultural, manufacturing, and service sectors to evaluate the importance of supply- and demand-side explanations for NBG in explaining the Kuznets facts for the United States. These previous studies, however, only match the Kuznets facts in terms of sectoral employment shares, and either do not consider (Iscan 2010) or fail to match (Buera and Kaboski 2009) the Kuznets facts regarding sectoral output shares. Further, these previous studies do not fully explore their models' abilities to match the Kaldor facts. On the other hand, Kongsamut, Rebelo, and Xie (2001) and Foellmi and Zweimuller (2008) attempt to reconcile NBG with the Kaldor facts within a three-sector model of NBG that includes the agricultural, manufacturing, and service sectors, but only consider one causal mechanism for NBG in their models (demand-side NBG), and, hence, are subject to the critique mentioned above that models that focus on only one mechanism for NBG cannot match the Kuznets facts for the United States and other advanced economies.[3]

The analysis in this essay focuses on the model's ability to generate short-run transition dynamics that match the Kaldor and Kuznets facts in the post-war United States. The decision to focus on the short-run transition dynamics – and to ignore the characterization of the asymptotic equilibrium for the model – is a reflection of Buera and Kaboski (2009)'s result that any model that integrates supply-side and demand-side explanations for NBG will be inconsistent with balanced long-run growth for output. Buera and Kaboski (2009)'s result is confirmed by Iscan (2010), who is unable to characterize a long-run steady state that features constant growth rates for output or a constant capital-to-labor ratio for his model that integrates supply-side and demand-side explanations for NBG. Further, Acemoglu and Guerrieri (2008) note that the asymptotic equilibrium for models of NBG typically features one sector becoming "asymptotically dominant" with all other sectors becoming "vanishingly small" in terms of their shares of employment and other factors. This essay's focus on the short-run behavior of the economy allows the model's ability to match the Kaldor and Kuznets facts to be evaluated when there are comparable (i.e., non-trivial) levels of employment and capital in the agricultural, manufacturing, and service sectors.

---

[3] Other recent studies have developed three-sector models to analyze: changes in sectoral employment shares between the manufacturing, service, and agricultural sectors in a small open economy (South Korea) where exogenous price changes as a result of international trade influence the process of NBG (Mao and Yao 2012); changes in employment shares between manufacturing and the "progressive" and "asymptotically stagnant" components of the service sector (Kapur 2009); and recent changes in sectoral labor allocations in China between the agricultural, non-agricultural, and public sectors, where changes in the scale of the public sector and institutional barriers to labor mobility between sectors are important drivers of changes in sectoral labor allocations (Lei, Zhang, and Wu 2013).

As mentioned above, supply-side explanations for NBG posit that NBG occurs as the result of differential rates of productivity growth between sectors. Following Acemoglu and Guerrieri (2008), the model in this essay considers two drivers sectoral productivity growth: (i) sector-specific rates of Hicks-neutral technical change and (ii) capital-deepening. Capital-deepening refers to the increase in the capital-to-labor ratio over time in a growing economy.[4] Capital-deepening can contribute to NBG because more capital-intensive sectors derive greater benefit from the increase in the capital-to-labor ratio in the economy, which causes their output shares to increase relative to less capital-intensive sectors and influences the distribution of labor and other factors of production between sectors.

Following previous models of supply-side NBG (e.g., Ngai and Pissarides 2007; Acemoglu and Guerrieri 2008), the direction of supply-side NBG in the model is determined by the elasticity of substitution between the outputs of the more progressive and less progressive sectors in the economy. In particular, a less progressive sector's factor shares will increase as the economy grows when the elasticity of substitution between its output and the output of more progressive sectors is low (less than 1). This is because, when the elasticity of substitution is low, the increase in the relative price of output in the less progressive sector is more than proportional to the relative declines in the marginal products of capital, labor, and other factors of production in the sector, which causes the sector's factor shares to increase as the economy grows. The converse holds when the elasticity of substitution between the outputs in the more and less progressive sectors is high (greater than 1).

[4] Hicks-neutral technical change is a form of total factor productivity growth where technical change does not influence the marginal rate of substitution between inputs. Formally, total factor productivity is the ratio of output to an index of inputs (Chambers 1988). Growth in total factor productivity is a useful measure of technical change because the ratio of total factor productivity from two different time periods provides a measure of the change in the effectiveness of the index of inputs in producing output.

As mentioned above, demand-side explanations for NBG posit that NBG occurs as the result of changes in the composition of demand for sectoral output that occur as the economy grows. In most previous models of demand-side NBG, the changes in the composition of demand occur as a result of difference in the income elasticities of demand for sectoral output that arise because of non-homothetic consumer preferences (see Kongsamut, Rebelo, and Xie 2001; Gollin, Parente, and Rogerson 2002; Iscan 2010). In contrast to these previous models, changes in the composition of demand for sectoral output in the model presented in this essay occur as a result of differences in the elasticities of demand for sectoral output with respect to total output in the economy. This approach has three important advantages.

First, as is described in detail in Chapter 2.2, considering the elasticities of demand for sectoral output with respect to total output allows physical capital to be produced in the model using the output from all three sectors. That physical capital is produced using output from all three sectors helps the model to produce realistic behavior of both sectoral output and employment shares. It is likely that the inability of Buera and Kaboski (2009) or Iscan (2010) to match the Kuznets facts in terms of changes in sectoral output shares is due, in part, to their unrealistic restriction that physical capital is produced using output from only one sector (manufacturing).

Second, as detailed in Chapter 2.3.2, considering the elasticities of demand for sectoral output with respect to total output allows the model to be calibrated using data on nominal value added in each sector rather than data on consumption expenditures in each sector. Iscan (2010) describes the empirical issues associated with mapping the parameters for a utility function capturing non-homothetic consumer preferences to data on consumption expenditure from NIPA. These empirical issues are related to properly accounting for intermediate inputs and for the relationship between consumption categories and categories of sectoral output.[5]

[5] For this reason, Iscan (2010) uses total expenditures on food, non-food goods, and services to ap-

Third, it is unlikely that the model's predictions concerning changes over time in the composition of demand for sectoral output would change significantly if elasticities of demand for sectoral output were modeled with respect to consumption expenditure rather than total output. This is because over the period considered in this study, the ratio of consumption expenditure to total output – the consumption-to-output ratio – has been relatively constant in the United States. King et al. (1991) find evidence in support of the constancy of the consumption-to-output ratio for the United States for the period 1954 to 1988; more recent studies by Clemente et al. (1999) and Attfield and Temple (2006) confirm this result in econometric models that allow for structural breaks in the data.[6]

The remainder of this essay is structured as follows. In Chapter 2.2, the three-sector model of NBG used in the model calibration is developed and the model equilibrium conditions are presented. Chapter 2.3 describes the calibration procedure and explains how parameters and initial values are chosen to correspond to the observed data on the U.S. economy. In Chapter 2.4, the calibration results are presented and discussed. Finally, in Chapter 2.5, conclusions are given and directions for future research are discussed.

proximate consumptions shares for agricultural, manufacturing, and service sectors when calibrating his model.

[6]The constancy, or stationarity, of the consumption-to-output ratio – or, equivalenty, the constancy of the savings rate – is a standard feature of neoclassical economic growth models starting with Solow (1956). Indeed, neoclassical growth models predict a balanced growth path in terms of per capita output, consumption, and investment, which implies a constant consumption-to-output ratio. Empirical evidence suggests that the consumption-to-output has been approximately constant for some advanced countries, but not others (Clemente et al. 1999). For developing countries, Li and Daly (2009) find that the Chinese economy followed a growth path that featured a constant consumption-to-output prior to the 1970s, but not afterwards.

2.2 The Model

This section develops the three-sector model of NBG that is used in the model calibration. This section has two parts. First, Chapter 2.2.1 describes the preferences and the technology in the model. Second, Chapter 2.2.2 describes the equilibrium conditions for the model and demonstrates that the dynamic competitive equilibrium of the model can be represented by a boundary value system of ordinary differential equations. In the next section, it is shown that the boundary value system of ordinary differential equations that defines the dynamic equilibrium can be approximated by a system of difference equations whose parameters and initial values correspond to observed data on the U.S. economy. It is this parameterized system of difference equations that is used in the model calibration.

The competitive equilibrium for the model presented in this section is characterized by solving the social planner's problem of maximizing the utility of the representative household. The equivalence between the social planner's problem and the competitive equilibrium for the economy in this chapter follows the first and second fundamental theorems of welfare economics (Arrow and Debreu 1954).[7] Together,

[7] That the first and second welfare theorems hold for the economy in this chapter follows from the facts that (i) all households and firms in the economy are assumed to be price takers (i.e., all markets function competitively), (ii) there are perfectly defined property rights for all assets in the economy (i.e., markets are complete), (iii) household preferences are locally non-satiated (each household's utility function is strictly increasing) and convex (as is demonstrated in Appendix E for the economy in Chapter 1, that household preferences are convex follows directly from the concavity of household utility functions in this chapter), (iv) that all production sets in the economy are convex, and (v) the "cheaper consumption" condition is always satisfied (Mas Colell, Whinston, and Green 1995). That all production sets are convex follows from the fact that, as in Chapter 1, they represent the intersection of upper-contour sets of quasi-concave functions which are convex. The proof that the cheaper consumption condition always holds for the economy in this chapter is identical to proof offered in Appendix E for the economy in Chapter 1.

the first and second fundamental theorems of welfare economics establish that any competitive equilibrium is Pareto optimal and that any distributional objective (such as the solution to the social planner's problem) can be achieved through competitive markets using an appropriate lump-sum redistribution of wealth. The economy in this chapter admits a representative household because each household has the same linear Engel curve for the single consumption good. Linear Engel curves imply that the distribution of wealth between households does not influence aggregate demand, which further implies that the competitive equilibrium will not be influenced by the distribution of wealth across households and that, by the second welfare theorem, the competitive equilibrium will coincide with the solution to the social planner's problem.[8]

2.2.1 Preferences and Technology

Preferences

In the model, the population $L(t)$ at time t grows at the exponential rate $n \in [0, \delta)$, where δ is the rate of time preference for households in the economy, so that

$$L(t) = e^{nt} L(0) . \tag{2.1}$$

Each household member supplies inelastically one unit of labor per unit of time, so $L(t)$ also denotes total labor supply at time t.

The representative household derives utility at time t from per capita consumption

[8] A well-known necessary and sufficient condition for the economy to admit a representative household is for all households in the economy to have Gorman-form indirect utility functions (see Gorman 1953; Muellbauer 1976). The property of Gorman-form indirect utility functions that drives this result is that Gorman-form indirect utility functions imply identical linear Engel curves for all households.

of a unique final good

$$u[c(t)] = \frac{c(t)^{1-\theta}}{1-\theta}, \tag{2.2}$$

where $c(t)$ is per capita consumption at time t, and $\theta > 0$ is the inverse of the intertemporal elasticity of substitution. The representative household's dynamic optimization problem is presented below in Chapter 2.2.2.

Technology

The unique final good is produced competitively by combining manufacturing, service, and agricultural (M-, S-, and A-sector) output according to

$$Y(t) = \left[\mu_M (Y_M(t) - \gamma_M)^{\frac{\varepsilon-1}{\varepsilon}} + \mu_S (Y_S(t) - \gamma_S)^{\frac{\varepsilon-1}{\varepsilon}} + (1 - \mu_M - \mu_S)(Y_A(t) - \gamma_A)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{2.3}$$

where $Y_M(t) \geq 0$, $Y_S(t) \geq 0$, and $Y_A(t) \geq 0$ are M-, S-, and A-sector output at time t. $\varepsilon \in [0, \infty)$ is the elasticity of substitution between M-, S-, and A-sector output in the production of $Y(t)$; μ_M, $\mu_S \in (0, 1)$ determine the relative importance of $Y_M(t)$, $Y_S(t)$, and $Y_A(t)$ in the production of $Y(t)$; γ_M, γ_S, γ_A are parameters such that the elasticity of demand for $Y_M(t)$ ($Y_S(t)$, $Y_A(t)$) with respect to $Y(t)$ is less than 1 when $\gamma_M > 0$ (γ_S, $\gamma_A > 0$) and greater than 1 otherwise. Further, it is assumed that $Y_M(t) - \gamma_M \geq 0$, $Y_S(t) - \gamma_S \geq 0$, and $Y_A(t) - \gamma_A \geq 0$ for all t. As is discussed in Chapter 2.4, the parameters ε, γ_M, γ_S, and γ_A are key ingredients in the explanations for NBG considered in this essay.

It is assumed that physical capital is produced according to a linear production technology where $Y(t)$ is the only input. That $Y(t)$ is both the sole consumption good and the sole input in the production of physical capital implies a resource constraint for the economy. In particular, the resource constraint is that aggregate consumption, $C(t) \equiv c(t)L(t)$, plus aggregate investment in physical capital, $\dot{K}(t)$, cannot be greater than total output of the final good, $Y(t)$, net of capital depreciation, $dK(t)$, where $K(t)$ is capital stock at time t and $d \geq 0$ is the depreciation rate of capital. That is,

$$Y(t) \geq C(t) + \dot{K}(t) + dK(t). \tag{2.4}$$

(2.4) must hold with equality in equilibrium because if it did not the social planner could costlessly raise the utility of the representative household by increasing consumption.

M-, S-, and A-sector output is produced competitively using constant returns to scale Cobb-Douglas production technology that combines labor and capital to produce final output

$$\begin{aligned} Y_M(t) &= A_M(t) L_M(t)^{\alpha_M} K_M(t)^{1-\alpha_M}, \\ Y_S(t) &= A_S(t) L_S(t)^{\alpha_S} K_S(t)^{1-\alpha_S}, \end{aligned} \tag{2.5}$$

and

$$Y_A(t) = A_A(t) L_A(t)^{\alpha_A} K_A(t)^{1-\alpha_A},$$

where $L_M(t) \geq 0$, $K_M(t) \geq 0$, $L_S(t) \geq 0$, $K_S(t) \geq 0$, $L_A(t) \geq 0$, and $K_A(t) \geq 0$ are the levels of labor and capital used in the M-, S-, and A-sectors; and the parameters $\alpha_M, \alpha_S, \alpha_A \in [0, 1]$ determine labor and capital's shares of total cost in the M-, S-, and A-sectors (i.e., α_M, α_S, and α_A are the elasticity of output with respect to labor in the M-, S-, and A-sectors; and $(1 - \alpha_M)$, $(1 - \alpha_S)$, and $(1 - \alpha_A)$ are the elasticity of output with respect to capital in the M-, S-, and A-sectors).

Technical progress in the M-, S-, and A-sectors takes the form of exogenous Hicks-neutral technical change

$$\hat{A}_M(t) = a_M > 0,\ \hat{A}_S(t) = a_S > 0, \text{ and } \hat{A}_A(t) = a_A > 0. \tag{2.6}$$

The labor and capital market-clearing conditions at time t are given by

$$L_M(t) + L_S(t) + L_A(t) = L(t) \tag{2.7}$$

and

$$K_M(t) + K_S(t) + K_A(t) = K(t). \tag{2.8}$$

2.2.2 Equilibrium

The characterization of the equilibrium is divided into two parts: the static equilibrium and the dynamic equilibrium. The static equilibrium defines the allocations of factors across the M-, S-, and A-sectors – $L_M(t)$, $K_M(t)$, $L_S(t)$, $K_S(t)$, $L_A(t)$, and $K_A(t)$ – at a given time t that maximize output of the final good, $Y(t)$, taking the values of the state variables, $K(t)$, $L(t)$, $A_M(t)$, $A_S(t)$, and $A_A(t)$, as given. In the dynamic equilibrium, the state variables are determined endogenously and evolve according to the representative household's intertemporal utility maximization. The dynamic equilibrium is characterized by solving for the optimal dynamic allocation for the social planner's problem. The results from this section are used in Chapter 2.3 to describe the equilibrium dynamics of the economy used in the calibration.

Static equilibrium

In this section, the static optimal allocations for the social planner are presented. Note that four of the state variables – $L(t)$, $A_M(t)$, $A_S(t)$, and $A_A(t)$ – evolve exogenously ((2.1) and (2.6)), so that, for given the initial conditions, $L(0)$, $A_M(0)$, $A_S(0)$, and $A_A(0)$, their values at time t are exogenous. For this reason, the optimal static allocations derived in this section will be defined in terms of time, t, and capital, $K(t)$, which is the only endogenous state variable in the model.

The maximized value of current output at time t for a given level of capital, $K(t)$, is

$$Y(K(t), t) = \max_{L_M, K_M, L_S, K_S, L_A, K_A} Y(t), \tag{2.9}$$

where $Y(t)$ is defined in (2.3), subject to the labor and capital market-clearing conditions, (2.7) and (2.8), and given $L(0) > 0$, $A_M(0) > 0$, $A_S(0) > 0$, and $A_A(0) > 0$.

As the objective function in (2.9) is continuous and strictly concave, and as the constraint set defined by (2.7) and (2.8) is a convex set, the social planner's static optimization problem will have a unique solution that corresponds to the unique competitive equilibrium at time t.

The first-order conditions for the social planner's static maximization problem are

$$\begin{aligned}
\mu_M \alpha_M \left[\frac{Y(t)}{Y_M(t) - \gamma_M}\right]^{\frac{1}{\varepsilon}} \frac{Y_M(t)}{L_M(t)} - \lambda_L(t) &= 0 \qquad (2.10)\\
\mu_S \alpha_S \left[\frac{Y(t)}{Y_S(t) - \gamma_S}\right]^{\frac{1}{\varepsilon}} \frac{Y_S(t)}{L_S(t)} - \lambda_L(t) &= 0\\
(1 - \mu_M - \mu_S) \alpha_A \left[\frac{Y(t)}{Y_A(t) - \gamma_A}\right]^{\frac{1}{\varepsilon}} \frac{Y_A(t)}{L_A(t)} - \lambda_L(t) &= 0\\
\mu_M (1 - \alpha_M) \left[\frac{Y(t)}{Y_M(t) - \gamma_M}\right]^{\frac{1}{\varepsilon}} \frac{Y_M(t)}{K_M(t)} - \lambda_K(t) &= 0\\
\mu_S (1 - \alpha_S) \left[\frac{Y(t)}{Y_S(t) - \gamma_S}\right]^{\frac{1}{\varepsilon}} \frac{Y_S(t)}{K_S(t)} - \lambda_K(t) &= 0\\
(1 - \mu_M - \mu_S)(1 - \alpha_A) \left[\frac{Y(t)}{Y_A(t) - \gamma_A}\right]^{\frac{1}{\varepsilon}} \frac{Y_A(t)}{K_A(t)} - \lambda_K(t) &= 0\\
L(t) - L_M(t) - L_S(t) - L_A(t) &= 0\\
K(t) - K_M(t) - K_S(t) - K_A(t) &= 0
\end{aligned}$$

where $\lambda_L(t)$ and $\lambda_K(t)$ are the Lagrangean multipliers associated with (2.7) and (2.8). The first-order conditions can be used to determine the optimal allocations of $L_M(t)$, $K_M(t)$, $L_S(t)$, $K_S(t)$, $L_A(t)$, and $K_A(t)$ in terms of $K(t)$, and, hence, can be used to determine the maximized value of output in terms of $K(t)$, $Y(K(t), t)$.

The competitive prices for M-, S-, and A-sector output, $p_M(t)$, $p_S(t)$, and $p_A(t)$, as well as the prices for labor and capital, $w(t)$ and $R(t)$, are implied by the social planner's static optimization problem. To begin, the price of final output is normalized to one. Given this normalization, and setting the prices of M-, S-, and A-sector output equal to their marginal revenue products, yields

$$p_M(t) = \mu_M \left[\frac{Y(t)}{Y_M(t) - \gamma_M}\right]^{\frac{1}{\varepsilon}}, \; p_S(t) = \mu_S \left[\frac{Y(t)}{Y_S(t) - \gamma_S}\right]^{\frac{1}{\varepsilon}}, \qquad (2.11)$$

and

$$p_A(t) = (1 - \mu_M - \mu_S)\left[\frac{Y(t)}{Y_A(t) - \gamma_A}\right]^{\frac{1}{\varepsilon}}.$$

Next, the factor prices, $w(t)$ and $r(t)$, must equal the marginal revenue products of labor and capital in the M-, S-, and A-sectors (i.e., $w(t)$ and $R(t)$ must equal the Lagrangean multipliers for the constraints given by (2.7) and (2.8)). That is,

$$w(t) = p_M(t)\,\alpha_M \frac{Y_M(t)}{L_M(t)} = p_S(t)\,\alpha_S \frac{Y_S(t)}{L_S(t)} = p_A(t)\,\alpha_A \frac{Y_A(t)}{L_A(t)} \tag{2.12}$$

and

$$R(t) = p_M(t)(1-\alpha_M)\frac{Y_M(t)}{K_M(t)} = p_S(t)(1-\alpha_S)\frac{Y_S(t)}{K_S(t)} = p_A(t)(1-\alpha_A)\frac{Y_A(t)}{K_A(t)}, \tag{2.13}$$

where $R(t) \equiv r(t) + d$ is the rental price of capital that is equal to the interest rate, $r(t)$, plus the depreciation rate of capital, d.

Further, (2.10) implies the following expressions for the M-, S-, and A-sectors' shares of total employment at time t:

$$\psi_M(t) \equiv \frac{L_M(t)}{L(t)} = \left[\begin{array}{c} 1 + \frac{\mu_S\alpha_S}{\mu_M\alpha_M}\left(\frac{Y_M(t)-\gamma_M}{Y_S(t)-\gamma_S}\right)^{\frac{1}{\varepsilon}}\frac{Y_S(t)}{Y_M(t)} \\ +\frac{(1-\mu_M-\mu_S)\alpha_A}{\mu_M\alpha_M}\left(\frac{Y_M(t)-\gamma_M}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_M(t)}\end{array}\right]^{-1} \tag{2.14}$$

$$\psi_S(t) \equiv \frac{L_S(t)}{L(t)} = \left[\begin{array}{c} 1 + \frac{\mu_M\alpha_M}{\mu_S\alpha_S}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)} \\ +\frac{(1-\mu_M-\mu_S)\alpha_A}{\mu_S\alpha_S}\left(\frac{Y_S(t)-\gamma_S}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_S(t)}\end{array}\right]^{-1}, \tag{2.15}$$

and

$$\psi_A(t) \equiv \frac{L_A(t)}{L(t)} = 1 - \psi_M(t) - \psi_S(t).$$

Similarly, (2.10) implies the following expressions for the M-, S-, and A-sectors' shares of total capital at time t:

$$\phi_M(t) \equiv \frac{K_M(t)}{K(t)} = \left[\begin{array}{c} 1 + \frac{\mu_S(1-\alpha_S)}{\mu_M(1-\alpha_M)}\left(\frac{Y_M(t)-\gamma_M}{Y_S(t)-\gamma_S}\right)^{\frac{1}{\varepsilon}}\frac{Y_S(t)}{Y_M(t)} \\ +\frac{(1-\mu_M-\mu_S)(1-\alpha_A)}{\mu_M(1-\alpha_M)}\left(\frac{Y_M(t)-\gamma_M}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_M(t)}\end{array}\right]^{-1} \tag{2.16}$$

$$\phi_S(t) \equiv \frac{K_S(t)}{K(t)} = \left[\begin{array}{c} 1 + \frac{\mu_M(1-\alpha_M)}{\mu_S(1-\alpha_S)}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)} \\ +\frac{(1-\mu_M-\mu_S)(1-\alpha_A)}{\mu_S(1-\alpha_S)}\left(\frac{Y_S(t)-\gamma_S}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_S(t)}\end{array}\right]^{-1}, \tag{2.17}$$

and

$$\phi_A(t) \equiv \frac{K_A(t)}{K(t)} = 1 - \phi_M(t) - \phi_S(t).$$

Dynamic equilibrium

In this section, the dynamic equilibrium is characterized as the solution to the social planner's dynamic optimization problem. The social planner's objective is the maximize the discounted sum of the representative household's future instantaneous utilities,

$$\int_0^\infty \frac{c(t)^{1-\theta}}{1-\theta} e^{-(\delta-n)t} dt, \tag{2.18}$$

subject to the resource constraint for the economy, (2.4), which can be rewritten as

$$\dot{K}(t) = Y(K(t), t) - dK(t) - c(t) e^{nt} L(0), \tag{2.19}$$

and given the initial conditions $L(0) > 0$, $K(0) > 0$, $A_M(0) > 0, A_S(0) > 0$, and $A_A(0) > 0$.[9] Recall that $Y(K(t), t)$ is the maximum value of $Y(t)$ at time t for given $K(t)$ implied by the first-order conditions for the social planner's static maximization problem.

The solution to the representative household's program is found by maximizing

[9]A sufficient condition for the improper intergral in (2.18) to converge is for $\delta - n > 0$ (which is assumed) and for the representative household's instantaneous utility to be bounded for all $t \in [0, \infty)$, i.e., for $\left|\frac{c(t)^{1-\theta}}{1-\theta} e^{-(\delta-n)t}\right| \leq B \in \mathbb{R}_{++}$ for all $t \in [0, \infty)$. When these two conditions hold, which is assumed in this chapter, the improper integral in (2.18) will converge and not exceed $\frac{B}{\delta-n}$ (Caputo 2005, p.384). The assumption that the representative household's attainable instantaneous utility is bounded for all $t \in [0, \infty)$ is equivalent to assuming that the representative household's resources are constrained so that it cannot obtain an arbitrarily large level utility (i.e., an arbitrarily large level of consumption) at any time t.

the current value Hamiltonian[10]

$$H^c\left(c\left(t\right),K\left(t\right),\nu\left(t\right)\right)=\frac{c\left(t\right)^{1-\theta}}{1-\theta}+\nu\left(t\right)\left[Y^*\left(K\left(t\right),t\right)-dK\left(t\right)-c\left(t\right)e^{nt}L\left(0\right)\right] \tag{2.20}$$

where $\nu(t)$ is the co-state variable. Applying Theorem 14.3 from Caputo (2005), if $\{c^*\left(t\right),K^*\left(t\right)\}$ is an optimal solution, then it is necessary that there exists piecewise smooth function $v(t)$ such that for all $t\in[0,\infty)$,[11]

$$\begin{aligned}\frac{\partial H^c\left(c\left(t\right),K\left(t\right),\nu\left(t\right)\right)}{\partial c\left(t\right)} &= c\left(t\right)^{-\theta}-\nu\left(t\right)e^{nt}L\left(0\right)=0 &(2.21)\\ -\frac{\partial H^c\left(c\left(t\right),K\left(t\right),\nu\left(t\right)\right)}{\partial K\left(t\right)} &= -\nu\left(t\right)\left[\frac{\partial Y^*\left(K\left(t\right),t\right)}{\partial K\left(t\right)}-d\right] &(2.22)\\ &= \dot{\nu}\left(t\right)-\left(\delta-n\right)\nu\left(t\right),\end{aligned}$$

and

$$\frac{\partial H^c\left(c\left(t\right),K\left(t\right),\nu\left(t\right)\right)}{\partial\nu\left(t\right)}=\dot{K}\left(t\right)=Y\left(K\left(t\right),t\right)-dK\left(t\right)-c\left(t\right)e^{nt}L\left(0\right). \tag{2.23}$$

Applying Theorem 14.4 and Lemma 14.1 from Caputo (2005), given that H^c is a concave function of $c\left(t\right)$ and $K\left(t\right)$ for all $t\in[0,\infty)$ over the open convex set containing all the admissible values of $c\left(t\right)$ and $K\left(t\right)$, i.e., over the open convex set $\{\left(c\left(t\right),K\left(t\right)\right)|\,c\left(t\right)>0,K\left(t\right)>0\}$, the necessary conditions in (2.21), (2.22), and (2.23) are sufficient to identify the global maximum of the social planner's dynamic

[10] As (2.18) and (2.19) and the initial condition $K\left(0\right)>0$ form a maximization problem with a continuously differentiable objective function and a non-empty set of controls, Pontryagin's Maximum Principle is a necessary condition for the optimal solution to the social planner's dynamic optimization problem.

[11] The function $v(t)$ is piecewise smooth on the interval $t\in[0,\infty)$ if its deriative function $\dot{v}(t)$ is piecewise continous on the interval $t\in[0,\infty)$ (Caputo 2005, Definition 1.2). The function $\dot{v}(t)$ is piecewise continous on the interval $t\in[0,\infty)$ if the inteval $t\in[0,\infty)$ can be partitioned into a finite number of points $0=t_0<t_1<...<t_k<\infty$ such that (i) $\dot{v}(t)$ is continous on each open subinterval $t_{k-1}<t<t_k$ and (ii) $\dot{v}(t)$ approaches a finite limit as the end points of each subinteval are approached from within the subinterval (Caputo 2005, Definition 1.1).

optimization problem if the transversality condition

$$\lim_{t\to\infty} K(t)\,\nu(t)\,e^{-(\delta-n)t} = 0 \tag{2.24}$$

is satisfied.[12]

Combining (2.21) and (2.22), the optimal consumption path for the representative household is characterized by Euler equation

$$\dot{c}(t) = \frac{c(t)}{\theta}\left[\frac{\partial Y^{*}(K(t),t)}{\partial K(t)} - d - \delta\right] = \frac{c(t)}{\theta}\left[R(t) - d - \delta\right], \tag{2.25}$$

together with the initial conditions, $L(0) > 0$, $K(0) > 0$, $A_M(0) > 0, A_S(0) > 0$, and $A_A(0) > 0$, the state equation (2.19), and the transversality conditions (2.24).

The following proposition combines this characterization of the optimal consumption path for the representative household with the first-order conditions for the social planner's static maximization problem from (2.10) to provide a full representation of the dynamic equilibrium for the economy.

Proposition 14 *The dynamic equilibrium for the economy can be represented by a system of six non-autonomous differential equations representing the rates of change in per capita consumption and capital stock:*

$$\dot{c}(t) = \frac{c(t)}{\theta}\left[R(t) - d - \delta\right] \tag{2.26}$$

$$\dot{K}(t) = Y(t) - dK(t) - c(t)L(t) \tag{2.27}$$

where $R(t)$ is defined by (2.13) and $Y(t)$ is defined by (2.3); the M-sector's share of total employment

$$\begin{aligned}\dot{\psi}_M(t) &= -\psi_M(t)^2\left(1 + b_{\psi_M M}(t)\,\alpha_M\psi_M(t) - b_{\psi_M A}(t)\,\alpha_A\frac{1}{1-\psi_S(t)-\psi_M(t)}\psi_M(t)^2\right)^{-1} \\ &\quad\times\left(\begin{array}{c} b_{\psi_M M}(t)\left[f_M(t) - \alpha_M\frac{1}{\psi_M(t)}\dot{\psi}_M(t)\right] - b_{\psi_M S}(t)\,f_S(t) \\ -b_{\psi_M A}(t)\left[f_A(t) + \alpha_A\frac{1}{1-\psi_S(t)-\psi_M(t)}\dot{\psi}_M(t)\right]\end{array}\right),\end{aligned} \tag{2.28}$$

---

[12] Theorems 14.3 and 14.4 and Lemma 14.1 in Caputo (2005) are presented for the present-value Hamiltonian. Adjustments have been made for the use of the current-value Hamiltonian in this chapter.

where

$$
\begin{aligned}
b_{\psi_M M}(t) &= \left[\frac{\mu_S\alpha_S}{\mu_M\alpha_M}\left(\frac{Y_M(t)-\gamma_M}{Y_S(t)-\gamma_S}\right)^{\frac{1}{\varepsilon}}\frac{Y_S(t)}{Y_M(t)}+\frac{(1-\mu_S-\mu_M)\alpha_A}{\mu_M\alpha_M}\left(\frac{Y_M(t)-\gamma_M}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_M(t)}\right] \\
&\quad \times Y_M(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_M(t)-\gamma_M}-\frac{1}{Y_M(t)}\right] \\
b_{\psi_M S}(t) &= \left[\frac{\mu_S\alpha_S}{\mu_M\alpha_M}\left(\frac{Y_M(t)-\gamma_M}{Y_S(t)-\gamma_S}\right)^{\frac{1}{\varepsilon}}\frac{Y_S(t)}{Y_M(t)}\right]\times Y_S(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_S(t)-\gamma_S}-\frac{1}{Y_S(t)}\right] \\
b_{\psi_M A}(t) &= \left[\frac{(1-\mu_S-\mu_M)\alpha_A}{\mu_M\alpha_M}\left(\frac{Y_M(t)-\gamma_M}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_M(t)}\right]\times Y_A(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_A(t)-\gamma_A}-\frac{1}{Y_A(t)}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
f_M(t) &= a_M+\alpha_M n+(1-\alpha_M)\frac{1}{K(t)}\dot{K}(t)+\alpha_M\frac{1}{\psi_M(t)}\dot{\psi}_M(t)+(1-\alpha_M)\frac{1}{\phi_M(t)}\dot{\phi}_M(t) \\
f_S(t) &= a_S+\alpha_S n+(1-\alpha_S)\frac{1}{K(t)}\dot{K}(t)+\alpha_S\frac{1}{\psi_S(t)}\dot{\psi}_S(t)+(1-\alpha_S)\frac{1}{\phi_S(t)}\dot{\phi}_S(t) \\
f_A(t) &= a_A+\alpha_A n+(1-\alpha_A)\frac{1}{K(t)}\dot{K}(t)-\alpha_A\frac{1}{1-\psi_S(t)-\psi_M(t)}\left(\dot{\psi}_M(t)+\dot{\psi}_S(t)\right) \\
&\quad -(1-\alpha_A)\frac{1}{1-\phi_S(t)-\phi_M(t)}\left(\dot{\phi}_M(t)+\dot{\phi}_S(t)\right);
\end{aligned}
$$

the S-sector's share of total employment:

$$
\begin{aligned}
\dot{\psi}_S(t) &= -\psi_S(t)^2\left(1+b_{\psi_S S}(t)\alpha_S\psi_S(t)-b_{\psi_S A}(t)\alpha_A\frac{1}{1-\psi_S(t)-\psi_M(t)}\psi_S(t)^2\right)^{-1} \\
&\quad \times\left(\begin{array}{c} b_{\psi_S S}(t)\left[f_S(t)-\alpha_S\frac{1}{\psi_S(t)}\dot{\psi}_S(t)\right]-b_{\psi_S M}(t)f_M(t) \\ -b_{\psi_S A}(t)\left[f_A(t)+\alpha_A\frac{1}{1-\psi_S(t)-\psi_M(t)}\dot{\psi}_S(t)\right]\end{array}\right),
\end{aligned}
\tag{2.29}
$$

where

$$
\begin{aligned}
b_{\psi_S M}(t) &= \left[\frac{\mu_M\alpha_M}{\mu_S\alpha_S}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)}\right]\times Y_M(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_M(t)-\gamma_M}-\frac{1}{Y_M(t)}\right] \\
b_{\psi_S S}(t) &= \left[\frac{\mu_M\alpha_M}{\mu_S\alpha_S}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)}+\frac{(1-\mu_M-\mu_S)\alpha_A}{\mu_S\alpha_S}\left(\frac{Y_S(t)-\gamma_S}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_S(t)}\right] \\
&\quad \times Y_S(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_S(t)-\gamma_S}-\frac{1}{Y_S(t)}\right] \\
b_{\psi_S A}(t) &= \left[\frac{(1-\mu_M-\mu_S)\alpha_A}{\mu_S\alpha_S}\left(\frac{Y_S(t)-\gamma_S}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_S(t)}\right]\times Y_A(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_A(t)-\gamma_A}-\frac{1}{Y_A(t)}\right];
\end{aligned}
$$

the M-sector's share of total capital stock:

$$\dot{\phi}_M(t) = -\phi_M(t)^2 \left(1 + b_{\phi_M M}(t)(1-\alpha_M)\phi_M - b_{\phi_M A}(t)(1-\alpha_A)\frac{1}{1-\phi_S(t)-\phi_M(t)}\phi_M(t)^2\right)^{-1} \times \left(\begin{array}{c} b_{\phi_M M}\left[f_M(t) - (1-\alpha_M)\frac{1}{\phi_M(t)}\dot{\phi}_M(t)\right] - b_{\phi_M S} f_S(t) \\ -b_{\phi_M A}(t)\left[f_A(t) + (1-\alpha_A)\frac{1}{1-\phi_S(t)-\phi_M(t)}\dot{\phi}_M(t)\right] \end{array}\right), \quad (2.30)$$

where

$$b_{\phi_M M}(t) = \left[\frac{\mu_S(1-\alpha_S)}{\mu_M(1-\alpha_M)}\left(\frac{Y_M(t)-\gamma_M}{Y_S(t)-\gamma_S}\right)^{\frac{1}{\varepsilon}}\frac{Y_S(t)}{Y_M(t)} + \frac{(1-\mu_S-\mu_M)(1-\alpha_A)}{\mu_M(1-\alpha_M)}\left(\frac{Y_M(t)-\gamma_M}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_M(t)}\right] \times Y_M(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_M(t)-\gamma_M} - \frac{1}{Y_M(t)}\right]$$

$$b_{\phi_M S}(t) = \left[\frac{\mu_S(1-\alpha_S)}{\mu_M(1-\alpha_M)}\left(\frac{Y_M(t)-\gamma_M}{Y_S(t)-\gamma_S}\right)^{\frac{1}{\varepsilon}}\frac{Y_S(t)}{Y_M(t)}\right] \times Y_S(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_S(t)-\gamma_S} - \frac{1}{Y_S(t)}\right]$$

$$b_{\phi_M A}(t) = \left[\frac{(1-\mu_S-\mu_M)(1-\alpha_A)}{\mu_M(1-\alpha_M)}\left(\frac{Y_M(t)-\gamma_M}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_M(t)}\right] \times Y_A(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_A(t)-\gamma_A} - \frac{1}{Y_A(t)}\right];$$

and the S-sector's share of total capital stock:

$$\dot{\phi}_S(t) = -\phi_S(t)^2 \left(1 + b_{\phi_S S}(t)(1-\alpha_S)\phi_S - b_{\phi_S A}(t)(1-\alpha_A)\frac{1}{1-\phi_S(t)-\phi_M(t)}\phi_S(t)^2\right)^{-1} \times \left(\begin{array}{c} b_{\phi_S S}(t)\left[f_S(t) - (1-\alpha_S)\frac{1}{\phi_S(t)}\dot{\phi}_S(t)\right] - b_{\phi_S M}(t) f_M(t) \\ -b_{\phi_S A}(t)\left[f_A(t) + (1-\alpha_A)\frac{1}{1-\phi_S(t)-\phi_M(t)}\dot{\phi}_S(t)\right] \end{array}\right), \quad (2.31)$$

where

$$b_{\phi_S M}(t) = \left[\frac{\mu_M(1-\alpha_M)}{\mu_S(1-\alpha_S)}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)}\right] \times Y_M(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_M(t)-\gamma_M} - \frac{1}{Y_M(t)}\right]$$

$$b_{\phi_S S}(t) = \left[\frac{\mu_M(1-\alpha_M)}{\mu_S(1-\alpha_S)}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)} + \frac{(1-\mu_M-\mu_S)(1-\alpha_A)}{\mu_S(1-\alpha_S)}\left(\frac{Y_S(t)-\gamma_S}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_S(t)}\right] \times Y_S(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_S(t)-\gamma_S} - \frac{1}{Y_S(t)}\right]$$

$$b_{\phi_S A}(t) = \left[\frac{(1-\mu_M-\mu_S)(1-\alpha_A)}{\mu_S(1-\alpha_S)}\left(\frac{Y_S(t)-\gamma_S}{Y_A(t)-\gamma_A}\right)^{\frac{1}{\varepsilon}}\frac{Y_A(t)}{Y_S(t)}\right] \times Y_A(t)\left[\frac{1}{\varepsilon}\frac{1}{Y_A(t)-\gamma_A} - \frac{1}{Y_A(t)}\right];$$

four autonomous differential equations:

$$\dot{L}(t) = nL(t) > 0, \quad (2.32)$$

$$\dot{A}_M(t) = a_M A_M(t) > 0,\ \dot{A}_S(t) = a_S A_S(t) > 0,\ \text{and } \dot{A}_A(t) = a_A A_A(t) > 0;$$

along with the initial conditions

$$L(0) > 0, K(0) > 0, A_M(0) > 0, A_S(0) > 0, \text{ and } A_A(0) > 0,$$

and the transversality condition

$$\lim_{t \to \infty} K(t)\, \nu(t)\, e^{-(\delta - n)t} = 0. \tag{2.33}$$

Proof. To begin, several of the equations in Proposition 14 are unchanged from above. In particular, (2.26) is (2.25), (2.27) is (2.19), (2.32) follows from (2.1) and (2.6), and (2.33) is (2.24).

Next, (2.28), (2.29), (2.30), and (2.31) are obtained by differentiating the expressions for the M-sector's employment share, (2.14), the S-sector's employment share, (2.15), M-sector's capital share, (2.16), and the S-sector's capital share, (2.17), with respect to time and substituting the expressions for $\dot{Y}_M(t)$, $\dot{Y}_S(t)$, and $\dot{Y}_A(t)$ obtained by differentiating $Y_M(t)$, $Y_S(t)$, and $Y_A(t)$ from (2.5) with respect to time. ■

In the next section, it is shown that the dynamic equilibrium for the economy described in Proposition 14 can be approximated by a system of difference equations. It is this system of difference equations that is used in the calibration.

2.3 Calibration Procedure

This section describes the procedure used in the calibration. This section has two parts. First, in Chapter 2.3.1, it is shown that the dynamic equilibrium for the economy described in Proposition 14 can be approximated by a system of difference equations. Second, in Chapter 2.3.2, it is explained how the parameters and initial values for this system of difference equations are chosen to correspond to the observed data on the U.S. economy. It is this parameterized system of difference equations that is used to predict the time paths of sectoral employment, sectoral output, and other variables in the model calibration.

2.3.1 Model for Calibration

In the following proposition, Euler's method is used to generate an approximate solution to the system of differential equations that represent the dynamic equilibrium for the economy presented in Proposition 14. Euler's method converts the system of ordinary differential equations defined in Proposition 14 into a system of difference equations that give an approximate solution to the dynamic equilibrium of the economy. As is mentioned above, this system of difference equations is used to generate the calibration results presented in Chapter 2.4.

The advantage of Euler's method is that its implementation, described in Judd (1998), is straightforward, but its drawback is that it does not provide an estimate of the error in the approximation. This drawback of the Euler method, however, is not relevant to this application. The periodicity of the NIPA data – most of the relevant series used in the calibration are only available on an annual basis – means that in the model calibration, which involves second-order difference equations, the step-size in the Euler method must be set equal to one year. In the Euler method, the error in the approximation is proportional to the step-size. Even if an alternative means of approximating the dynamic equilibrium of the economy were used which did provide an estimate of the error, it would not be possible to decrease the step-size to improve the approximation given the periodicity of the available data.

Before continuing, it is necessary to introduce new notation to distinguish variables and parameters in the calibration from those in the theoretical model developed in Chapter 2.2. For starters, the variables in the calibrated model are discrete and are denoted with a t subscript. For example, $K(t)$ denotes total capital stock at time t in the theoretical model, while K_t denotes total capital stock in year t in the calibration. Similarly, the parameters in the calibration are denoted with a tilda to distinguish them from the parameters in the theoretical model. For example, α_M is labor's share of total cost in the M-sector in the theoretical model, while $\tilde{\alpha}_M$ is labor's share of total cost in Manufacturing Sector in the calibration, which is calculated directly from NIPA data (see Chapter 2.3.2). Finally, the terms Manufacturing sector, Service sector, and Agricultural sector are capitalized in the context of the calibration, where they denote specific groupings of industries. The industries included in the Manufacturing, Service Sector, and Agricultural Sectors in the calibration are described in Chapter 2.3.2. The subscripts M, S, and A are used in the calibration to denote the Manufacturing, Service, and Agricultural Sectors.

Proposition 15 *Using the Euler method, the dynamic equilibrium of the economy can be approximated by a system of five non-autonomous, second-order difference equations that capture the evolution of capital stock,*

$$K_{t+1} = K_t + h\left(Y_t - \tilde{d}K_t\right) - \frac{L_t}{L_{t-1}}\left[h\left(Y_{t-1} - \tilde{d}K_{t-1}\right) - (K_t - K_{t-1})\right]\left[1 + \frac{h}{\tilde{\theta}}\left(R_{t-1} - \tilde{d} - \tilde{\delta}\right)\right] \tag{2.34}$$

where

$$Y_t = \left[\tilde{\mu}_M\left(Y_{M,t} - \tilde{\gamma}_M\right)^{\frac{\tilde{\varepsilon}-1}{\tilde{\varepsilon}}} + \tilde{\mu}_S\left(Y_{S,t} - \tilde{\gamma}_S\right)^{\frac{\tilde{\varepsilon}-1}{\tilde{\varepsilon}}} + \left(1 - \tilde{\mu}_M - \tilde{\mu}_S\right)\left(Y_{A,t} - \tilde{\gamma}_A\right)^{\frac{\tilde{\varepsilon}-1}{\tilde{\varepsilon}}}\right]^{\frac{\tilde{\varepsilon}}{\tilde{\varepsilon}-1}}, \tag{2.35}$$

$$\begin{aligned} Y_{M,t} &= A_{M,t} L_t^{\tilde{\alpha}_M} K_t^{1-\tilde{\alpha}_M} \psi_{M,t}^{\tilde{\alpha}_M} \phi_{M,t}^{1-\tilde{\alpha}_M}, \\ Y_{S,t} &= A_{S,t} L_t^{\tilde{\alpha}_S} K_t^{1-\tilde{\alpha}_S} \psi_{S,t}^{\tilde{\alpha}_S} \phi_{S,t}^{1-\tilde{\alpha}_S} \\ Y_{A,t} &= A_{A,t} L_t^{\tilde{\alpha}_A} K_t^{1-\tilde{\alpha}_A} \left(1 - \psi_{M,t} - \psi_{S,t}\right)^{\tilde{\alpha}_A} \left(1 - \phi_{M,t} - \phi_{S,t}\right)^{1-\tilde{\alpha}_A}, \end{aligned} \tag{2.36}$$

$$R_t = \tilde{\mu}_P\left(1 - \tilde{\alpha}_P\right)\left[\frac{Y_t}{Y_{P,t} - \tilde{\gamma}_P}\right]^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{P,t}}{K_t \psi_{P,t}},\ P = A, M, S, \tag{2.37}$$

$$\psi_{M,t} = \left[\begin{array}{c} 1 + \frac{\tilde{\mu}_S \tilde{\alpha}_S}{\tilde{\mu}_M \tilde{\alpha}_M}\left(\frac{Y_{M,t} - \tilde{\gamma}_M}{Y_{S,t} - \tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{S,t}}{Y_{M,t}} \\ + \frac{(1 - \tilde{\mu}_M - \tilde{\mu}_S)\tilde{\alpha}_A}{\tilde{\mu}_M \tilde{\alpha}_M}\left(\frac{Y_{M,t} - \tilde{\gamma}_M}{Y_{A,t} - \tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{M,t}} \end{array}\right]^{-1} \tag{2.38}$$

$$\psi_{S,t} = \left[\begin{array}{c} 1 + \frac{\tilde{\mu}_M \tilde{\alpha}_M}{\tilde{\mu}_S \tilde{\alpha}_S}\left(\frac{Y_{S,t} - \tilde{\gamma}_S}{Y_{M,t} - \tilde{\gamma}_M}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{M,t}}{Y_{S,t}} \\ + \frac{(1 - \tilde{\mu}_M - \tilde{\mu}_S)\tilde{\alpha}_A}{\tilde{\mu}_S \tilde{\alpha}_S}\left(\frac{Y_{S,t} - \tilde{\gamma}_S}{Y_{A,t} - \tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{S,t}} \end{array}\right]^{-1}, \tag{2.39}$$

$$\phi_{M,t} = \left[\begin{array}{c} 1+\frac{\tilde{\mu}_S(1-\tilde{\alpha}_S)}{\tilde{\mu}_M(1-\tilde{\alpha}_M)} \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{S,t}}{Y_{M,t}} \\ +\frac{(1-\tilde{\mu}_M-\tilde{\mu}_S)(1-\tilde{\alpha}_A)}{\tilde{\mu}_M(1-\tilde{\alpha}_M)} \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{M,t}} \end{array} \right]^{-1}, \tag{2.40}$$

$$\phi_{S,t} = \left[\begin{array}{c} 1+\frac{\tilde{\mu}_M(1-\tilde{\alpha}_M)}{\tilde{\mu}_S(1-\tilde{\alpha}_S)} \left(\frac{Y_{S,t}-\tilde{\gamma}_S}{Y_{M,t}-\tilde{\gamma}_M}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{M,t}}{Y_{S,t}} \\ +\frac{(1-\tilde{\mu}_M-\tilde{\mu}_S)(1-\tilde{\alpha}_A)}{\tilde{\mu}_S(1-\tilde{\alpha}_S)} \left(\frac{Y_{S,t}-\tilde{\gamma}_S}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{S,t}} \end{array} \right]^{-1}; \tag{2.41}$$

the M-sector's share of total employment,

$$\psi_{M,t+1} = \psi_{M,t} + h \left[\begin{array}{c} -\psi_{M,t}^2 \left(1 + b_{\psi_M M,t}\tilde{\alpha}_M \psi_{M,t} - b_{\psi_M A,t}\tilde{\alpha}_A \frac{1}{1-\psi_{S,t}-\psi_{M,t}} {\psi_{M,t}}^2\right)^{-1} \\ \times \left(\begin{array}{c} b_{\psi_M M,t}\left[f_{M,t} - \tilde{\alpha}_M \frac{1}{\psi_{M,t}} \left(\frac{\psi_{M,t}-\psi_{M,t-1}}{h}\right)\right] - b_{\psi_M S,t} f_{S,t} \\ -b_{\psi_M A,t}\left[f_{A,t} + \tilde{\alpha}_A \frac{1}{1-\psi_{S,t}-\psi_{M,t}} \left(\frac{\psi_{M,t}-\psi_{M,t-1}}{h}\right)\right] \end{array} \right) \end{array} \right] \tag{2.42}$$

where

$$b_{\psi_M M,t} = \begin{array}{c} \left[\frac{\tilde{\mu}_S\tilde{\alpha}_S}{\tilde{\mu}_M\tilde{\alpha}_M} \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{S,t}}{Y_{M,t}} + \frac{(1-\tilde{\mu}_S-\tilde{\mu}_M)\tilde{\alpha}_A}{\tilde{\mu}_M\tilde{\alpha}_M} \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{M,t}}\right] \\ \times Y_{M,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{M,t}-\tilde{\gamma}_M} - \frac{1}{Y_{M,t}}\right] \end{array}$$

$$b_{\psi_M S,t} = \left[\frac{\tilde{\mu}_S\tilde{\alpha}_S}{\tilde{\mu}_M\tilde{\alpha}_M} \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{S,t}}{Y_{M,t}}\right] \times Y_{S,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{S,t}-\tilde{\gamma}_S} - \frac{1}{Y_{S,t}}\right]$$

$$b_{\psi_M A,t} = \left[\frac{(1-\tilde{\mu}_S-\tilde{\mu}_M)\,\tilde{\alpha}_A}{\tilde{\mu}_M\tilde{\alpha}_M} \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{M,t}}\right] \times Y_{A,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{A,t}-\tilde{\gamma}_A} - \frac{1}{Y_{A,t}}\right],$$

and

$$\begin{aligned} f_{M,t} &= a_M + \tilde{\alpha}_M n + (1-\tilde{\alpha}_M)\frac{1}{K_t}\left(\frac{K_t-K_{t-1}}{h}\right) + \tilde{\alpha}_M\frac{1}{\psi_{M,t}}\left(\frac{\psi_{M,t}-\psi_{M,t-1}}{h}\right) \\ &\quad + (1-\tilde{\alpha}_M)\frac{1}{\phi_{M,t}}\left(\frac{\phi_{M,t}-\phi_{M,t-1}}{h}\right) \\ f_{S,t} &= a_S + \tilde{\alpha}_S n + (1-\tilde{\alpha}_S)\frac{1}{K_t}\left(\frac{K_t-K_{t-1}}{h}\right) + \tilde{\alpha}_S\frac{1}{\psi_{S,t}}\left(\frac{\psi_{S,t}-\psi_{S,t-1}}{h}\right) \\ &\quad + (1-\tilde{\alpha}_S)\frac{1}{\phi_{S,t}}\left(\frac{\phi_{S,t}-\phi_{S,t-1}}{h}\right) \\ f_{A,t} &= a_A + \tilde{\alpha}_A n + (1-\tilde{\alpha}_A)\frac{1}{K_t}\left(\frac{K_t-K_{t-1}}{h}\right) \\ &\quad -\tilde{\alpha}_A\frac{1}{1-\psi_{S,t}-\psi_{M,t}}\left[\left(\frac{\psi_{M,t}-\psi_{M,t-1}}{h}\right) + \left(\frac{\psi_{S,t}-\psi_{S,t-1}}{h}\right)\right] \\ &\quad - (1-\tilde{\alpha}_A)\frac{1}{1-\phi_{S,t}-\phi_{M,t}}\left(\left(\frac{\phi_{M,t}-\phi_{M,t-1}}{h}\right) + \left(\frac{\phi_{S,t}-\phi_{S,t-1}}{h}\right)\right); \end{aligned}$$

the S-sector's share of total employment,

$$\psi_{S,t+1}=\psi_{S,t}+h\left[\begin{array}{c} -\psi_{S,t}^{2}\left(1+b_{\psi_S S,t}\tilde{\alpha}_S\psi_{S,t}-b_{\psi_S A,t}\tilde{\alpha}_A\frac{1}{1-\psi_{S,t}-\psi_{M,t}}\psi_{S,t}^{2}\right)^{-1} \\ \times\left(\begin{array}{c} b_{\psi_S S,t}\left[f_{S,t}-\tilde{\alpha}_S\frac{1}{\psi_{S,t}}\left(\frac{\psi_{S,t}-\psi_{S,t-1}}{h}\right)\right]-b_{\psi_S M,t}f_{M,t} \\ -b_{\psi_S A,t}\left[f_{A,t}+\tilde{\alpha}_A\frac{1}{1-\psi_{S,t}-\psi_{M,t}}\left(\frac{\psi_{S,t}-\psi_{S,t-1}}{h}\right)\right] \end{array}\right) \end{array}\right], \tag{2.43}$$

where

$$\begin{aligned}
b_{\psi_S M,t} &= \left[\frac{\tilde{\mu}_M\tilde{\alpha}_M}{\tilde{\mu}_S\tilde{\alpha}_S}\left(\frac{Y_{S,t}-\tilde{\gamma}_S}{Y_{M,t}-\tilde{\gamma}_M}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{M,t}}{Y_{S,t}}\right]\times Y_{M,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{M,t}-\tilde{\gamma}_M}-\frac{1}{Y_{M,t}}\right] \\
b_{\psi_S S,t} &= \begin{array}{c}\left[\frac{\tilde{\mu}_M\tilde{\alpha}_M}{\tilde{\mu}_S\tilde{\alpha}_S}\left(\frac{Y_{S,t}-\tilde{\gamma}_S}{Y_{M,t}-\tilde{\gamma}_M}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{M,t}}{Y_{S,t}}+\frac{(1-\tilde{\mu}_M-\tilde{\mu}_S)\tilde{\alpha}_A}{\tilde{\mu}_S\tilde{\alpha}_S}\left(\frac{Y_{S,t}-\tilde{\gamma}_S}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{A,t}}{Y_{S,t}}\right] \\ \times Y_{S,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{S,t}-\tilde{\gamma}_S}-\frac{1}{Y_{S,t}}\right]\end{array} \\
b_{\psi_S A,t} &= \left[\frac{(1-\tilde{\mu}_M-\tilde{\mu}_S)\,\tilde{\alpha}_A}{\tilde{\mu}_S\tilde{\alpha}_S}\left(\frac{Y_{S,t}-\tilde{\gamma}_S}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{A,t}}{Y_{S,t}}\right]\times Y_{A,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{A,t}-\tilde{\gamma}_A}-\frac{1}{Y_{A,t}}\right];
\end{aligned}$$

the M-sector's share of total capital stock,

$$\phi_{M,t+1}=\phi_{M,t}+h\left[\begin{array}{c} -\phi_{M,t}^{2}\left(1+b_{\phi_M M,t}\left(1-\tilde{\alpha}_M\right)\phi_M-b_{\phi_M A,t}\left(1-\tilde{\alpha}_A\right)\frac{1}{1-\phi_{S,t}-\phi_{M,t}}\phi_{M,t}^{2}\right)^{-1} \\ \times\left(\begin{array}{c} b_{\phi_M M}\left[f_{M,t}-\left(1-\tilde{\alpha}_M\right)\frac{1}{\phi_{M,t}}\left(\frac{\phi_{M,t}-\phi_{M,t-1}}{h}\right)\right]-b_{\phi_M S}f_{S,t} \\ -b_{\phi_M A,t}\left[f_{A,t}+\left(1-\tilde{\alpha}_A\right)\frac{1}{1-\phi_{S,t}-\phi_{M,t}}\left(\frac{\phi_{M,t}-\phi_{M,t-1}}{h}\right)\right] \end{array}\right) \end{array}\right], \tag{2.44}$$

where

$$\begin{aligned}
b_{\phi_M M,t} &= \begin{array}{c}\left[\frac{\tilde{\mu}_S(1-\tilde{\alpha}_S)}{\tilde{\mu}_M(1-\tilde{\alpha}_M)}\left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{S,t}}{Y_{M,t}}+\frac{(1-\tilde{\mu}_S-\tilde{\mu}_M)(1-\tilde{\alpha}_A)}{\tilde{\mu}_M(1-\tilde{\alpha}_M)}\left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{A,t}}{Y_{M,t}}\right] \\ \times Y_{M,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{M,t}-\tilde{\gamma}_M}-\frac{1}{Y_{M,t}}\right]\end{array} \\
b_{\phi_M S,t} &= \left[\frac{\tilde{\mu}_S\left(1-\tilde{\alpha}_S\right)}{\tilde{\mu}_M\left(1-\tilde{\alpha}_M\right)}\left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{S,t}}{Y_{M,t}}\right]\times Y_{S,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{S,t}-\tilde{\gamma}_S}-\frac{1}{Y_{S,t}}\right] \\
b_{\phi_M A,t} &= \left[\frac{\left(1-\tilde{\mu}_S-\tilde{\mu}_M\right)\left(1-\tilde{\alpha}_A\right)}{\tilde{\mu}_M\left(1-\tilde{\alpha}_M\right)}\left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{A,t}-\tilde{\gamma}_A}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{A,t}}{Y_{M,t}}\right]\times Y_{A,t}\left[\frac{1}{\tilde{\varepsilon}}\frac{1}{Y_{A,t}-\tilde{\gamma}_A}-\frac{1}{Y_{A,t}}\right];
\end{aligned}$$

and the S-sector's share of total capital stock,

$$\phi_{S,t+1} = \phi_{S,t} + h \left[\begin{array}{c} -\phi_{S,t}^2 \left(1 + b_{\phi_S S,t} \left(1 - \tilde{\alpha}_S\right) \phi_S - b_{\phi_S A,t} \left(1 - \tilde{\alpha}_A\right) \frac{1}{1 - \phi_{S,t} - \phi_{M,t}} \phi_{S,t}^2 \right)^{-1} \\ \times \left(\begin{array}{c} b_{\phi_S S,t} \left[f_{S,t} - \left(1 - \tilde{\alpha}_S\right) \frac{1}{\phi_{S,t}} \left(\frac{\phi_{S,t} - \phi_{S,t-1}}{h} \right) \right] - b_{\phi_S M,t} f_{M,t} \\ -b_{\phi_S A,t} \left[f_{A,t} + \left(1 - \tilde{\alpha}_A\right) \frac{1}{1 - \phi_{S,t} - \phi_{M,t}} \left(\frac{\phi_{S,t} - \phi_{S,t-1}}{h} \right) \right] \end{array} \right) \end{array} \right], \tag{2.45}$$

where

$$\begin{aligned} b_{\phi_S M,t} &= \left[\frac{\tilde{\mu}_M \left(1 - \tilde{\alpha}_M\right)}{\tilde{\mu}_S \left(1 - \tilde{\alpha}_S\right)} \left(\frac{Y_{S,t} - \tilde{\gamma}_S}{Y_{M,t} - \tilde{\gamma}_M} \right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{M,t}}{Y_{S,t}} \right] \times Y_{M,t} \left[\frac{1}{\tilde{\varepsilon}} \frac{1}{Y_{M,t} - \tilde{\gamma}_M} - \frac{1}{Y_{M,t}} \right] \\ b_{\phi_S S,t} &= \begin{array}{c} \left[\frac{\tilde{\mu}_M (1 - \tilde{\alpha}_M)}{\tilde{\mu}_S (1 - \tilde{\alpha}_S)} \left(\frac{Y_{S,t} - \tilde{\gamma}_S}{Y_{M,t} - \tilde{\gamma}_M} \right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{M,t}}{Y_{S,t}} + \frac{(1 - \tilde{\mu}_M - \tilde{\mu}_S)(1 - \tilde{\alpha}_A)}{\tilde{\mu}_S (1 - \tilde{\alpha}_S)} \left(\frac{Y_{S,t} - \tilde{\gamma}_S}{Y_{A,t} - \tilde{\gamma}_A} \right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{S,t}} \right] \\ \times Y_{S,t} \left[\frac{1}{\tilde{\varepsilon}} \frac{1}{Y_{S,t} - \tilde{\gamma}_S} - \frac{1}{Y_{S,t}} \right] \end{array} \\ b_{\phi_S A,t} &= \left[\frac{\left(1 - \tilde{\mu}_M - \tilde{\mu}_S\right) \left(1 - \tilde{\alpha}_A\right)}{\tilde{\mu}_S \left(1 - \tilde{\alpha}_S\right)} \left(\frac{Y_{S,t} - \tilde{\gamma}_S}{Y_{A,t} - \tilde{\gamma}_A} \right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{A,t}}{Y_{S,t}} \right] \times Y_{A,t} \left[\frac{1}{\tilde{\varepsilon}} \frac{1}{Y_{A,t} - \tilde{\gamma}_A} - \frac{1}{Y_{A,t}} \right]; \end{aligned}$$

four autonomous, first-order difference equations

$$L_{t+1} = \left(1 + h\tilde{n}\right) L_t, \tag{2.46}$$

$$A_{M,t+1} = \left(1 + h\tilde{a}_M\right) A_{M,t},\ A_{S,t+1} = \left(1 + h\tilde{a}_S\right) A_{S,t},\ \text{and}\ A_{A,t+1} = \left(1 + h\tilde{a}_A\right) A_{A,t};$$

and the initial conditions

$$\begin{gathered} L_0, L_1, K_0, K_1, A_{M,0}, A_{M,1}, A_{S,0}, A_{S,1}\ A_{A,0}, A_{A,1}, \\ \psi_{M,0}, \psi_{M,0}, \psi_{S,0}, \psi_{S,1}, \phi_{M,0}, \phi_{M,1}, \phi_{S,0},\ \text{and}\ \phi_{S,1}. \end{gathered}$$

The step size, h, is set equal to one. This implies that $t_{t+1} = t_t + 1$.

Proof. The proof of Proposition 15 is a straightforward application of the Euler method as described in Judd (1998). To begin, many of the expressions in Proposition 15 are simply discretized version of expressions presented in Chapter 2.2. In particular, (2.35) is (2.3), (2.36) is (2.5), (2.37) is (2.13), (2.38) and (2.39) are (2.14) and (2.15), and (2.40) and (2.41) are (2.16) and (2.17).

(2.34) is derived in three steps. First, discretizing (2.19) and (2.25) using the Euler method yields

$$K_{t+1} = K_t + h\left[Y_t - \tilde{d}K_t - c_tL_t\right] \tag{2.47}$$

and

$$c_{t+1} = c_t + h\left[\frac{c_t}{\tilde{\theta}}\left(R_t - \tilde{d} - \tilde{\delta}\right)\right]. \tag{2.48}$$

Second, (2.47) can be re-expressed as

$$c_t = \frac{Y_t - \tilde{d}K_t}{L_t} - \frac{K_{t+1} - K_t}{hL_t}.$$

Substituting this expression for c_t into (2.48) gives

$$c_{t+1} = \left(\frac{Y_t - \tilde{d}K_t}{L_t} - \frac{K_{t+1} - K_t}{hL_t}\right)\left[1 + \frac{h}{\tilde{\theta}}\left(R_t - \tilde{d} - \tilde{\delta}\right)\right]. \tag{2.49}$$

Third, moving the time subscript in (2.49) back one period and substituting the resulting expression for c_t into (2.47) yields (2.34). It is necessary that (2.47) and (2.48) be combined to form (2.34) because household consumption, c_t, while performing a useful role in the model, does not correspond to any observed data on the U.S. economy. Note that (2.34) is a second-order difference equation and thus requires two years of initial values. For this reason, Table 2.3 reports initial values for both 1949 and 1950.

The remaining difference equations follow from the straightforward application of the Euler method. In particular, (2.42) follows from (2.28), (2.43) from (2.29), (2.44) from (2.30), and (2.45) from (2.31). Similarly, (2.46) follows from (2.32). ■

The system of difference equations described by Proposition 15 is used to derive the calibration results presented in Chapter 2.4. Before continuing, there are two features of Proposition 15 that require further explanation. First, the expressions for (2.42), (2.43), (2.44), and (2.45) approximate the expressions for the rates of change in $\psi_M(t)$, $\psi_S(t)$, $\phi_M(t)$, and $\phi_S(t)$ on the right hand sides of (2.28), (2.29), (2.30), and (2.31) by $\dot{\psi}_M(t) \simeq h^{-1}\left(\psi_{M,t} - \psi_{M,t-1}\right)$, $\dot{\psi}_S(t) \simeq h^{-1}\left(\psi_{S,t} - \psi_{S,t-1}\right)$, etc. This approximation has the advantage that the value of $h^{-1}\left(\psi_{M,t}-\psi_{M,t-1}\right)$, etc., is known at the time step that $\psi_{M,t+1}$, $\psi_{M,t+1}$, $\psi_{M,t+1}$, and $\psi_{M,t+1}$ are calculated. An alternative approximation would be $\dot{\psi}_M(t) \simeq h^{-1}\left(\psi_{M,t+1}-\psi_{M,t}\right)$, etc. This approximation would change the difference equations (2.42), (2.43), (2.44), and (2.45) into a system of four equations and four unknowns (i.e., $\psi_{M,t+1}$, $\psi_{S,t+1}$, $\phi_{M,t+1}$, and $\phi_{S,t+1}$) to be solved at each step of the calibration. This system of four equations could be solved simultaneously at each stage of the calibration, which would introduce the issue of how to deal with multiple real roots, or solved recursively to obtain an expression for $\psi_{M,t+1}$, $\psi_{S,t+1}$, $\phi_{M,t+1}$, or $\phi_{S,t+1}$ in terms of variables determined before $t+1$ and fixed parameters, which would further complicated an already complex calibration procedure.

Second, the transversality condition in Proposition 15, (2.33), is not represented in Proposition 15. The transversality condition is not represented in Proposition 15 because, as is discussed in the Introduction, the analysis in this essay focuses on the calibrated model's ability to generate short-run patterns of NBG consistent with those observed in the U.S. economy in the post-war era, and, as such, does not consider whether the dynamic equilibrium for the economy that is approximated in Proposition 15 converges to a asymptotic growth path along which the transversality condition is satisfied.

2.3.2 Variables, Initial Values, and Parameters

In this section, it is explained how the variables and parameters in the calibration are chosen and how they correspond to the observed data on the U.S. economy. As is mentioned in the Introduction, the data used in the calibration is taken from the NIPA from the U.S. BEA.

In the calibration, industries are classified according to the North American Industry Classification System (NAICS). Using the NAICS, the definitions of the Manufacturing, Service, and Agricultural sectors in the calibration are as follows. To begin, the Agricultural sector corresponds to the BEA's definition of the *Agriculture, Forestry, Fishing, and Hunting* industry. The Manufacturing sector corresponds to the BEA's definition of the set of "Private goods-producing industries" minus the Agricultural sector. The Manufacturing sector includes (i) *Mining*, (ii) *Construction*, and (iii) *Manufacturing* (durable and non-durable goods). The Service sector corresponds to BEA's definition of the set of "Private services-producing industries". The Service sector includes (i) *Utilities*, (ii) *Wholesale Trade*, (iii) *Retail Trade*, (iv) *Transportation and Warehousing*, (v) *Information*, (vi) *Finance, Insurance, Real Estate, Rental, and Leasing* (without *Real Estate, Rental, and Leasing*), (vii) *Professional and Business Services*, (viii) *Educational Services, Health Care, and Social Assistance*, (ix) *Arts, Entertainment, Recreation, Accommodation, and Food Services*, and (x) *Other Services.*[13] These definitions of the Manufacturing, Service, and Agricultural sectors include all private industries in the economy, so that only the *Government* sector from the BEA is not included in the analysis.

The initial values of the variables in the calibration are calculate for the base years 1949 and 1950. The calibration procedure requires two years of initial data because, from Proposition 15, the system of difference equations that approximates the dynamic equilibrium for the economy includes several second-order difference equations that require two years of data to be initialized. The base years of 1949 and 1950 where used despite the fact that employment data was available from the BEA at the industry-level starting in 1948. Using 1948 and 1949 as the base year, however, undermined the performance of the calibrated model because total employment and

[13]The BEA industry *Real Estate, Rental, and Leasing* is excluded because it has a large capital share that does not reflect the share of capital in the industry production function. Rather, the high capital share in the *Real Estate, Rental, and Leasing* – an average of 0.94 between 1987 and 2011 – reflects the large value of assets in the industry.

Sector	**Industry***
Manufacturing	Mining
	Construction
	Manufacturing
Services	Utilities
	Wholesale Trade
	Retail Trade
	Transportation and Warehousing
	Information
	Finance, Insurance, Real Estate, Rental, and Leasing**
	Professional and Business Services
	Educational Services, Health Care, and Social Assistance
	Arts, Entertainment, Recreation, Accommodation, and Food Services
	Other Services
Agriculture	Agriculture, Forestry, Fishing, and Hunting
*According to the 2002 North American Industry Classification System (NAICS)	
***Real Estate, Rental, and Leasing* not included. See text for explanation.	

Table 2.1: Industries in the Manufacturing, Service, and Agricultural Sectors

real output in all three sectors was lower in 1949 than in 1948. The calibrated model is constructed to analyze upward secular trends in output, employment, and capital stock, and not fluctuations in these variables.

Variables and Initial Values

Real and nominal value-added by industry is available according to the NAICS from 1947 to 2012. Nominal value-added in (industry i) sector P in year t is expressed as $(Y_{i,t}^{N})$ $Y_{P,t}^{N}$. As is demonstrated below, nominal value-added is used to calculate the initial sectoral employment and capital shares used in the calibration. Real value-added expressed in 2005 dollars (billions) is calculated as the product value-added quantity index for industry i in year t and nominal value-added for industry i in 2005 (divided by 100).[14] Nominal and real value-added in the Manufacturing, Service, and Agricultural sectors is the sum of the nominal and real value-added in their component industries.

Nominal value-added in the base years of 1949 and 1950 in the Manufacturing, Service, and Agricultural sectors are $Y_{M,1949}^{N} = 86.1$, $Y_{M,1950}^{N} = 99.9$, $Y_{S,1949}^{N} = 109.7$, $Y_{S,1950}^{N} = 118.5$, $Y_{A,1949}^{N} = 18.6$, and $Y_{A,1950}^{N} = 19.9$. Real value-added in the base years of 1949 and 1950 in the Manufacturing, Service, and Agricultural sectors are $Y_{M,1949} = 577.6$, $Y_{S,1950} = 650.5$, $Y_{S,1949} = 839.7$, $Y_{S,1950} = 887.1$, $Y_{A,1949} = 28.5$, and $Y_{A,1950} = 29.7$ (in billions of 2005 dollars). The variables and initial values of real and nominal value added and other quantities used in the calibration are summarized in Tables 2.2 and 2.3.

[14] This procedure for calculating real value-added implies that the accuracy of the real value-added figures used in the calibration depends on the accuracy of the BEA's value-added quantity and price indices. The dependence of the results on the BEA's quantity and price indices is an important caveat to the analysis in this essay given that several previous authors have questioned the appropriateness of the methodology employed by the BEA to construct these indices (e.g., Nordhaus 2001).

Variables and Data	
$Y_{i,t}^N$	**Nominal value-added of industry i at time t.** Nominal value-added by industry (in billions) is available according to NAICS from 1947 to 2012.
$Y_{P,t}^N = \sum_{i \in \Omega_P} Y_{i,t}^N$	**Nominal value-added of the P-sector, $P = M, S, A$ at time t.** (Ω_P denotes the subset of industries in the P-sector.)
$Y_{i,t}$	**Real value-added of industry i at time t.** $Y_{i,t}$ is the value-added of industry i at time t expressed in 2005 dollars (billions). $Y_{i,t}$ is the product of the value-added quantity index (chain-weighted) for industry i at time t, $VAQI_{i,t}$, divided by 100, and the industry nominal value added in 2005, $Y_{i,2005}^N$. $VAQI_{i,t}$ is available according to NAICS from 1947 to 2012.
$Y_{P,t} = \sum_{i \in \Omega_P} Y_{i,t}$	**Real value added of the P-sector, $P = M, S, A$ at time t.**
L_t	**Employment at time t.** Employment is total full-time and part-time private sector employment, in thousands of employees. Total full-time and part-time employment data is available according to NAICS from 1948-2011.
$W_{i,t}$	**Total employee compensation in industry i at time t** (in millions). Total employee compensation data, in billions, are available according to NAICS from 1987 to 2011.
K_t^N	**Nominal net stock of private fixed assets at time t** (in billions). Private fixed asset data are available from 1924 to 2012.
K_t	**Real net stock of private fixed assets at time t** (in billions). K_t is expressed in 2005 dollars (billions). K_t is the product of the "chain-type quantity index" of private fixed assets, divided by 100, and the nominal total value of private fixed assets in 2005.
Note: The data in Table 2.2 are from the National Income and Product Accounts (NIPA) from the U.S. Bureau of Economic Analysis (BEA). Nominal value-added and value-added quantity indexes use the 2007 NAICS and are available in the 25 April 2013 release. Total full-time and part-time private sector employment from 1948 to 1997 use the 2002 NAICS and from 1997 to 2011 use the 2017 NAICS and are available in the 13 November 2012 release. Compensation of employees by industry use the 2007 NAICS and are avaiable 13 November 2012 release. "Current-cost" net stock of private fixed assets and "chain-type quantity indexes" for net stock of private fixed assests are available in the 30 September 2013 release.	

Table 2.2: Variables and Data Sources used in the Calibration

The measure of employment used in the calibration is total full-time and part-time employment by industry from NIPA (i.e., it is assumed in the calibration that the labor input is proportional to total full-time and part-time employment in a given year). Total full-time and part-time employment data are available from 1948 to 2011. Total full-time and part-time employment is used because it is the only employment measure that extends back to 1948 using the NAICS classifications. All other employment measures are only available according to the Standard Industrial Classifications (SIC) prior to 1997. Total private-sector employment in the base years of 1949 and 1950 are $L_{1949} = 41,306$ and $L_{1950} = 43,054$.

The total real value of capital stock in the economy is calculated as the real value of the net stock of private fixed assets in the economy in 2005 dollars. Real values of the net stock of private fixed assets are calculated for years 1924 to 2012 as the product of the nominal value of the net stock of private fixed assets in 2005 and the quantity index for the net stock of private fixed assets for the years 1924 to 2012. The total real value of capital stock in the base years of 1949 and 1950 are $K_{1949} = 5,286.5$ and $K_{1950} = 5,522.1$ (in billions of 2005 dollars).

Using the expressions for sectoral output prices from (2.11), the expression the Manufacturing and Service sectors' shares of total employment in year t from (2.14) and (2.15) can be expressed (in discrete time form as) as

$$\psi_{M,t} = \left[1 + \frac{\tilde{\alpha}_S}{\tilde{\alpha}_M} \frac{Y_{S,t}^N}{Y_{M,t}^N} + \frac{\tilde{\alpha}_A}{\tilde{\alpha}_M} \frac{Y_{A,t}^N}{Y_{M,t}^N} \right]^{-1} \tag{2.50}$$

and

$$\psi_{S,t} = \left[1 + \frac{\tilde{\alpha}_M}{\tilde{\alpha}_S} \frac{Y_{M,t}^N}{Y_{S,t}^N} + \frac{\tilde{\alpha}_A}{\tilde{\alpha}_S} \frac{Y_{A,t}^N}{Y_{S,t}^N} \right]^{-1}, \tag{2.51}$$

where $Y_{M,t}^N$, $Y_{S,t}^N$, and $Y_{A,t}^N$ are nominal value-added (output) in the Manufacturing, Service, and Agricultural sectors in year t. Similarly, again using the expressions for sectoral output prices from (2.11), the expressions for the Manufacturing and Service

Initial Values	
$Y^N_{M,1949} = 86.1$, $Y^N_{M,1950} = 99.9$	Nominal value-added in 1949 and 1950 in the Manufacturing sector (in billions).
$Y^N_{S,1949} = 109.7$, $Y^N_{S,1950} = 118.5$	Nominal value-added in 1949 and 1950 in the Service sector (in billions).
$Y^N_{A,1949} = 18.6$, $Y^N_{A,1950} = 19.9$	Nominal value-added in 1949 and 1950 in the Agricultural sector (in billions).
$Y_{M,1949} = 577.6$, $Y_{S,1950} = 650.5$	Real value-added in 1949 and 1950 in the Manufacturing sector (in billions of 2005 dollars).
$Y_{S,1949} = 839.7$, $Y_{S,1950} = 887.1$	Real value-added in 1949 and 1950 in the Service sector (in billions of 2005 dollars).
$Y_{A,1949} = 28.5$, $Y_{A,1950} = 29.7$	Real value-added in 1949 and 1950 in the Agricultural sector (in billions of 2005 dollars).
$L_{1949} = 41,306$, $L_{1950} = 43,054$	Total private-sector employment in 1949 and 1950 (in thousands).
$K_{1949} = 5,286.5$, $K_{1950} = 5,522.1$	Real capital stock in 1949 and 1950 (in billions of 2005 dollars).
$\psi_{M,1949} = 0.4213$, $\psi_{M,1950} = 0.4390$	Manufacturing sector's shares of employment in 1949 and 1950.
$\psi_{S,1949} = 0.5390$, $\psi_{S,1950} = 0.5229$	Service sector's shares of employment in 1949 and 1950.
$\phi_{M,1949} = 0.3746$, $\phi_{M,1950} = 0.3920$	Manufacturing sector's shares of capital in 1949 and 1950.
$\phi_{S,1949} = 0.4742$, $\phi_{S,1950} = 0.4621$	Service sector's shares of capital in 1949 and 1950.
$A_{M,1949} = 0.0781$, $A_{M,1950} = 0.0807$	Total factor productivity in the Manufacturing Sector in 1949 and 1950.
$A_{S,1949} = 0.0886$, $A_{S,1950} = 0.0923$	Total factor productivity in the Service Sector in 1949 and 1950.
$A_{A,1949} = 0.0292$, $A_{A,1950} = 0.0305$	Total factor productivity in the Agricultural Sector in 1949 and 1950.

Table 2.3: Initial Values used in the Calibration

sectors' shares of total capital at time t from (2.16) and (2.17) can be expressed as

$$\phi_{M,t} = \left[1 + \frac{(1-\tilde{\alpha}_S)}{(1-\tilde{\alpha}_M)}\frac{Y_{S,t}^N}{Y_{M,t}^N} + \frac{(1-\tilde{\alpha}_A)}{(1-\tilde{\alpha}_M)}\frac{Y_{A,t}^N}{Y_{M,t}^N}\right]^{-1} \tag{2.52}$$

and

$$\phi_{S,t} = \left[1 + \frac{(1-\tilde{\alpha}_M)}{(1-\tilde{\alpha}_S)}\frac{Y_{M,t}^N}{Y_{S,t}^N} + \frac{(1-\tilde{\alpha}_A)}{(1-\tilde{\alpha}_S)}\frac{Y_{A,t}^N}{Y_{S,t}^N}\right]^{-1}. \tag{2.53}$$

Given the values of $\tilde{\alpha}_M$ and $\tilde{\alpha}_S$ (which are calculated below), initial values of the Manufacturing and Service sectors' shares of total employment and capital in the base years of 1949 and 1950 can be calculated using (2.50), (2.51), (2.52), and (2.53) and the values for nominal value added in the base years of 1949 and 1950. The Manufacturing and Service sectors' shares of total employment and capital in 1949 and 1950 are $\psi_{M,1949} = 0.4213$, $\psi_{M,1950} = 0.4390$, $\psi_{S,1949} = 0.5390$, $\psi_{S,1950} = 0.5229$, $\phi_{M,1949} = 0.3746$, $\phi_{M,1950} = 0.3920$, $\phi_{S,1949} = 0.4742$, and $\phi_{S,1950} = 0.4621$.

Given that data on real value-added, real capital stock, and employment, the calculated values for the Manufacturing and Service sectors' shares of employment and capital from (2.50), (2.51), (2.52), and (2.53), and the values of $\tilde{\alpha}_M$ and $\tilde{\alpha}_S$ (which are calculated below), (2.5) can be used to calculate initial values for total factor productivity. The calculated values of total factor productivity in 1949 and 1950 are $A_{M,1949} = 0.0781$, $A_{M,1950} = 0.0807$, $A_{S,1949} = 0.0886$, $A_{S,1950} = 0.0923$, $A_{A,1949} = 0.0292$, $A_{A,1950} = 0.0305$.

Parameters

There are 16 fixed parameters in the model, of which ten, $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, $\tilde{\alpha}_S$, $\tilde{a}_M$, $\tilde{a}_S$, $\tilde{a}_A$, $\tilde{n}$, $\tilde{d}$, $\tilde{\delta}$, and $\tilde{\theta}$, are determined by the available data or from the literature and are fixed across simulation runs. The remaining parameters, $\tilde{\varepsilon}$, $\tilde{\gamma}_M$, $\tilde{\gamma}_S$, $\tilde{\gamma}_A$, $\tilde{\mu}_M$, and $\tilde{\mu}_S$, are varied between simulation runs in order to evaluate the relative importance of the competing explanations for NBG at explaining the observed patterns of sectoral output and employment growth in the U.S. economy. In this section, it is explained how the values for the parameters used in the calibration were chosen.

To begin, labor's share total cost in each sector in year t is computed as $\tilde{\alpha}_{P,t} = W_{P,t}/Y^N_{P,t}$, $P = M, S, A$, where $Y^N_{S,t}$ is nominal value-added in sector P in year t and $W_{P,t}$ is total employee compensation in sector P in year t. The values of $\tilde{\alpha}_P$, $P = M, S, A$, used in the calibration are calculated as the average value of $\tilde{\alpha}_{P,t}$ between 1987 and 2011 (data on "compensation of employees" by industry is only available from 1987 to 2011 using the NAICS industry definitions).[15] The values for labor's share of total cost in the Manufacturing, Service, and Agricultural sectors used in the calibration are $\tilde{\alpha}_M = 0.6065$, $\tilde{\alpha}_S = 0.6090$, and $\tilde{\alpha}_A = 0.2647$. Capital's share of total costs in the Manufacturing, Service, and Agricultural sectors are then calculated as $(1 - \tilde{\alpha}_P)$, $P = M, S, A$.

The growth rate of total employment, $\tilde{n}$, is chosen so that the calibrated model matches employment growth over the sample period. That is, $\tilde{n}$ is chosen so that

$$L_{2011} = (1 + \tilde{n})^{62} L_{1949}. \tag{2.54}$$

Equation 2.54 gives $\tilde{n} = 0.0153$.

Estimates of the rates of Hick's neutral technical change (i.e., the rates of total factor productivity growth) in the Agricultural, Manufacturing, and Service sectors are obtained from several different sources. Estimates of the productivity growth rate in Agriculture are obtained from U.S. Department of Agriculture, Economics Research Service (USDA-ERS). The USDA-ERS reports an average total factor productivity growth of 1.5% in the agricultural sector for the period 1948-2006. For the manufacturing sector, Bureau of Labor Statistics reports annual productivity growth rates of

[15] The use of aveage values of $\tilde{\alpha}_{P,t}$, $P = M, S, A$, to calculate $\tilde{\alpha}_P$ is justified because there is not a significant trent in $\tilde{\alpha}_{M,t}$, $\tilde{\alpha}_{S,t}$, or $\tilde{\alpha}_{A,t}$ between 1987 and 2011. Indeed, when $\tilde{\alpha}_{M,t}$, $\tilde{\alpha}_{S,t}$, and $\tilde{\alpha}_{A,t}$ are regressed on a constant and a time trend, the coefficient on the time trend is very close to zero and not significant at the 10% level in all three regressions.

Parameter Values	
$\tilde{\alpha}_M = 0.6065$, $\tilde{\alpha}_S = 0.6090$, $\tilde{\alpha}_A = 0.2647$	**Labor's share of value-added in the Manufacturing, Service, and Agricultural Sectors.** $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, and $\tilde{\alpha}_A$ are computed as the average values between 1987 and 2011 using data on (i) nominal value added by industry and (ii) total employee compensation by industry from the BEA. Data on "compensation of employees" by industry is only available from 1987 to 2011 using the NAICS industry definitions.*
$\tilde{a}_M = 0.0135$, $\tilde{a}_S = 0.005$, $\tilde{a}_A = 0.015$	**Rates of Hick's neutral technical change in the Agricultural, Manufacturing, and Service sectors.** Calculated from a variety of sources. See text for details.
$\tilde{n} = 0.0153$	**Employment growth rate.** Calculated from growth in total full-time and part-time employment in the private sector from 1948-2011 using (2.54).
$\tilde{d}$=0.05	**Depreciation rate.** Taken from Barro and Sala-i-Martin (2004).
$\tilde{\delta}$=0.02	**Discount rate.** Taken from Barro and Sala-i-Martin (2004).
$\tilde{\theta}^{-1}$=0.594	**Intertemporal elasticity of substitution.** Taken from Havranek et al. (2013).
Note: All the data used to compute $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, and $\tilde{\alpha}_A$ are described in Table 2.2.	

Table 2.4: Parameter Values used in the Calibration

1.35% for the period 1987-2006 (reported in Iscan 2010). For the service sector, there is evidence of significant sector-specific acceleration and deceleration of productivity growth over the study period. In particular, Triplett and Bosworth (2003) report service sector productivity growth in the United States of 0.1% for the period 1977 to 1995, while Bosworth and Triplett (2007) report an average service sector productivity growth of 0.99% for the period 1987 to 2006. For this reason, three scenarios for service sector productivity growth are considered that encompasses these extremes (i.e., $\tilde{a}_S = 0.001, 0.005, \text{and } 0.01$) . In addition, in order to appraise the importance of capital-deepening alone in driving supply-side NBG in model, an additional scenario is considered where the rate of total factor productivity growth is set equal to 0.76% in all three sectors. This corresponds to the average annual total factor productivity growth rate in the United States as a whole from 1960 to 1995 reported in Jorgenson and Yip (2001).[16]

The values of the discount rate and the depreciation rate of capital used in the calibration are $\tilde{\delta} = 0.02$ and $\tilde{d} = 0.05$. The values of $\tilde{\delta}$ and $\tilde{d}$ are standard in the literature and are the same as those used in Barro and Sala-i-Martin (2004) and Acemoglu and Guerrieri (2008). The value of the intertemporal elasticity of substitution used in the calibration is $\tilde{\theta}^{-1} = 0.594$ ($\tilde{\theta} = 1.684$). This corresponds to the average value of the intertemporal elasticity of substitution for the United States from the meta-analysis in Havranek et al. (2013).[17]

As is mentioned above, the parameters $\tilde{\varepsilon}$, $\tilde{\gamma}_M$, $\tilde{\gamma}_S$, and $\tilde{\gamma}_A$ are varied between calibration runs in order to evaluate their relative importance in explaining observed patterns of NBG in the United States over the study period. The expressions for sectoral output prices from (2.11), and the fact that nominal value added in the P-sector $Y_P^N(t) \equiv p_P(t) Y_P(t)$, $P = M, S, A$, imply that the following two equations must hold when the economy is in static equilibrium:

$$\frac{Y_M^N(t)}{Y_S^N(t)} = \frac{\mu_M}{\mu_S}\left(\frac{Y_S(t)-\gamma_S}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_S(t)} \text{ and } \frac{Y_M^N(t)}{Y_A^N(t)} = \frac{\mu_M}{1-\mu_M-\mu_S}\left(\frac{Y_A(t)-\gamma_A}{Y_M(t)-\gamma_M}\right)^{\frac{1}{\varepsilon}}\frac{Y_M(t)}{Y_A(t)}.$$

These two equations (in discrete time form) can be solved for $\tilde{\mu}_M$ and $\tilde{\mu}_S$:

$$\tilde{\mu}_M = \left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{S,t}}{Y_{M,t}}\frac{Y_{M,t}^N}{Y_{S,t}^N}\left[1+\left(\frac{Y_{M,t}-\tilde{\gamma}_M}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{S,t}}{Y_{M,t}}\frac{Y_{M,t}^N}{Y_{S,t}^N}+\left(\frac{Y_{A,t}-\tilde{\gamma}_A}{Y_{S,t}-\tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}}\frac{Y_{S,t}}{Y_{A,t}}\frac{Y_{A,t}^N}{Y_{S,t}^N}\right]^{-1} \tag{2.55}$$

[16] Christenson, Cummings, and Jorgenson (1980) report an average annual total factor productivity growth rate in the United States from 1947 to 1973 of 1.35%. This suggests that the 0.76% figure from Jorgenson and Yip (2001) is likely to cause the calibrated model to underperform relative to the U.S. economy in the earlier years of the calibration.

[17] This value of the intertemporal elasticity of substitution of $\tilde{\theta}^{-1} = 0.594$ is similar to estimates reported by Beaudry and van Wincoop (1996) use a panel of state-level data on the U.S. economy for the period 1953 to 1991, and report an intertemporal elasticity of substitution for nondurables consumption close to 1.

and

$$\tilde{\mu}_S = \left[1 + \left(\frac{Y_{M,t} - \tilde{\gamma}_M}{Y_{S,t} - \tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{S,t}}{Y_{M,t}} \frac{Y_{M,t}^N}{Y_{S,t}^N} + \left(\frac{Y_{A,t} - \tilde{\gamma}_A}{Y_{S,t} - \tilde{\gamma}_S}\right)^{\frac{1}{\tilde{\varepsilon}}} \frac{Y_{S,t}}{Y_{A,t}} \frac{Y_{A,t}^N}{Y_{S,t}^N}\right]^{-1}. \tag{2.56}$$

For given values of $\tilde{\varepsilon}$, $\tilde{\gamma}_M$, $\tilde{\gamma}_S$, and $\tilde{\gamma}_A$, the values of $\tilde{\mu}_M$, and $\tilde{\mu}_S$ are chosen so that the following expressions hold in the initial year of the calibration, 1949.

2.4 Calibration Results

This section presents the results of the model calibration in three parts. First, Chapter 2.4.1 presents results for the case when only supply-side NBG is considered. Next, Chapter 2.4.2 presents results for the case when only demand-side NBG is considered. Finally, Chapter 2.4.3 presents results for the case when supply- and demand-side NBG are considered simultaneously. In all three cases, the calibrated model is evaluated based on its ability to match both the Kuznets and Kaldor Facts for the United States over the study period.

Several previous studies have used a goodness-of-fit criteria to select the parameterization of their model calibration that produces the best fit to the observed data (see Buera and Kaboski 2009; Iscan 2010). For example, Iscan (2010) uses a "root mean squared error" criteria to select the "preferred" parameterization of his model of the U.S. economy. In this essay, however, evaluating the ability of the calibrated model to match both the Kuznets and Kaldor facts does not lend itself to a simple goodness-of-fit criteria. This is because evaluating the model's ability to match the Kaldor facts includes analyzing whether the output growth rate, the real interest rate, the capital-to-output ratio, and the share of capital income in total output are relatively constant over time, i.e., have low variances and do no exhibit significant time trends. If the only objective of this essay were to match the Kuznets facts in terms of changes in sectoral output and employment shares over time, then a goodness-of-fit criteria similar to those used in previous studies could be appropriate.

Supply-Side Non-Balanced Growth: Kuznets Facts						
	1949	2011				
Variable	U.S. Data	U.S. Data	$\tilde{\varepsilon} = 1$	$\tilde{\varepsilon} = 0.85$	$\tilde{\varepsilon} = 0.65$	$\tilde{\varepsilon} = 0.50$
$\psi_M = L_M/L$	0.4213*	0.1638	0.4377	0.4182	0.3843	0.3527
$\psi_S = L_S/L$	0.5390*	0.8245	0.5244	0.5478	0.5871	0.6224
$\psi_A = L_A/L$	0.0397*	0.0116	0.0379	0.0340	0.0286	0.0249
$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.2304	0.5411	0.5211	0.4887	0.4624
$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.7572	0.4174	0.4404	0.4775	0.5074
$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0124	0.0414	0.0384	0.0338	0.0302
*$\psi_{M,1949}$, $\psi_{S,1949}$, and $\psi_{A,1949} = 1 - \psi_{M,1949} - \psi_{S,1949}$ are calculate using (2.50) and (2.51).						

Table 2.5: Supply-Side Non-Balanced Growth: Kuznets Facts Results

2.4.1 Supply-Side Non-Balanced Growth: Results

Supply-Side Non-Balanced Growth: Kuznets Facts Results

Table 2.5 presents results on the calibrated model's ability to match the Kuznets facts for the case when only the supply-side NBG is considered (i.e., when $\tilde{\gamma}_M = \tilde{\gamma}_S = \tilde{\gamma}_A = 0$). Specifically, Table 2.5 reports on the employment and output shares for the Manufacturing, Services, and Agricultural sectors in the initial year and final year of the calibration, as well as the value of these variables in the U.S. data. The results reported in Table 2.5 consider values of $\tilde{\varepsilon} < 1$ because this is the empirically relevant case for the U.S. economy over the study period, where the less progressive Service sector has gained in employment relative to the more progressive Agricultural and Manufacturing sectors. Previous studies that have analyzed supply-side NBG in the United States that have spanned the study period considered in this essay have also used $\tilde{\varepsilon} < 1$ (e.g., Acemoglu and Guerrieri 2008; Buera and Kaboski 2009).[18]

The intuition for the mechanism for the supply-side NBG in the model is that differences between sectors in rates of Hick's neutral technical change and in the im-

[18] Acemoglu and Guerrieri (2008) use $\varepsilon = 0.76$ in their baseline calibration of the U.S. economy for the period $1948 - 2005$, and also run thier model with values of ε as low as 0.56 and as high as 0.86. Buera and Kaboski (2009) use $\varepsilon = 0.5$ in their preferred model calibration.

pact of capital-deepening in the economy have two countervailing effects. All else equal, sectors with higher rates of Hick's neutral technical change and that are more capital intensive (i.e., benefit more from capital-deepening) will experience stronger growth in labor and capital relative to other sectors. However, faster productivity growth and capital accumulation also causes output in more progressive sectors to grow faster, which will increase the relative price of output of less-progressive sectors. This increase in relative price will induce labor and capital to move into the less progressive sectors. When $\tilde{\varepsilon} < 1$, and the output from the less-progressive sector(s) (e.g., Services) and more-progressive sector(s) (e.g., Agriculture and Manufacturing) sectors are gross complements, the increase in the relative price of output in the less-progressive sector is more proportional to the increase in the marginal products of labor and capital in the more-progressive sectors, which causes the less-progressive sector's capital and labor shares to increase as the economy grows.[19] Table 2.5 confirms that when $\tilde{\varepsilon} = 1$ the model calibration predicts that the sectoral employment are approximately constant over the study period.

Table 2.5 shows that the model calibration's predictions concerning changes in sectoral employment shares is improved as $\tilde{\varepsilon}$ decreases. Table 2.5 also shows, however, that the model predicts a smaller movement of labor out of Agriculture and Manufacturing and into Service than is observed in the U.S. data even for low values of $\tilde{\varepsilon}$. For example, when $\tilde{\varepsilon} = 0.50$, the model predicts that the Service Sector will comprise almost 62% of total employment in 2011, compared to over 82% in the data. Further calibrations (not reported) revealed that considering values of $\tilde{\varepsilon}$ less than 0.50 does not improve the fit of the model substantially regarding sectoral employment shares. For this reason, 0.50 is used in subsequent calibrations of the models as a

[19] The role of $\tilde{\varepsilon}$ in determining the direction of the change factor shares is consistent with previous studies that have emphasized supply-side explanations for non-balanced growth (e.g., Ngai and Pissarides 2007; and Acemoglu and Guerrieri 2008).

low but reasonable value for $\tilde{\varepsilon}$. In using $\tilde{\varepsilon} = 0.50$, this essay follows Buera and Kaboski (2009), who use an elasticity of substitution of $\varepsilon = 0.5$ in their preferred model calibration.[20]

Table 2.5 also shows that the model calibration predicts that the Manufacturing and Agricultural sectors' shares of total output increase relative to the Service sector's over the study period, which is counter to what occurred in reality. The failure of the model to match the increase in the Service sector's share of output relative to the Manufacturing is the result of an insufficiency of supply-side explanations for NBG to explain the Kuznets facts as they apply to both sectoral output and employment shares. To understand why supply-side explanations are insufficient at explaining the Kuznets facts, recall from the discussion above that higher rates Hick's neutral technical change in the Agriculture/Manufacturing leads to an increase in Service sector's employment share because the price of Service sector output increases relative to the price of Agricultural/Manufacturing sector output, which attracts factors of production, including labor, into the Service sector. This change in relative prices in favor of the Service sector will only occur, however, if real output growth is more rapid in Agriculture and/or Manufacturing than in Services. As such, it is not possible for a model that considers only supply-side NBG to match the Kuznets facts in terms of both sectoral output and employment shares in the U.S. over the study period.

As is mentioned in Chapter 2.3.2, there is evidence of significant sector-specific acceleration and deceleration of productivity growth in the Service sector over the study period. For this reason, three scenarios for the rate of Hick's neutral technical change in the Service sector are considered on Table 2.6: the baseline scenario of 0.5% growth (which is also reported in Table 2.5), as well as 0.1% and 1.0% growth, which correspond to the estimates of service sector productivity growth in the United States

[20] Buera and Kaboski (2009) note that the best fit of their model would be under the assumption of Leontief preferences (i.e., $\varepsilon \rightarrow 0$), which the authors' deem to be unrealistic.

Supply-Side Non-Balanced Growth: Kuznets Facts Varying the Rate of Productivity Growth in the Service Sector ($\tilde{\varepsilon} = 0.50$)						
	1949	2011				
			$\tilde{a}_S$			$\tilde{a}_M = \tilde{a}_A = \tilde{a}_S$
Variable	U.S. Data	U.S. Data	0.001	0.005	0.01	0.0076
$\psi_M = L_M/L$	0.4213*	0.1638	0.3197	0.3527	0.4003	0.4388
$\psi_S = L_S/L$	0.5390*	0.8245	0.6566	0.6224	0.5733	0.5304
$\psi_A = L_A/L$	0.0397*	0.0116	0.0237	0.0249	0.0264	0.0308
$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.2304	0.4860	0.4624	0.4353	0.4190
$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.7572	0.4828	0.5074	0.5358	0.5564
$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0124	0.0313	0.0302	0.0289	0.0246

*$\psi_{M,1949}$, $\psi_{S,1949}$, and $\psi_{A,1949} = 1 - \psi_{M,1949} - \psi_{S,1949}$ are calculate using (2.50) and (2.51).

Table 2.6: Supply-Side Non-Balanced Growth: Varying the Rate of Productivity Growth in the Service Sector

reported in Triplett and Bosworth (2003) for the period 1977 to 1995 and in Bosworth and Triplett (2007) for the period 1987 to 2006. Table 2.6 demonstrates that the 0.1% scenarios provides the best fit for sectoral employment shares, while the 1.0% scenario provides the best fit for sectoral output shares. Given this apparent trade-off between the model's ability to match sectoral employment and output shares, the remaining calibrations reported in this essay will continue using the 0.5% for the rate of Hick's neutral technical change in the Service sector.

Table 2.6 also reports the results for the case where the rate of Hick's neutral technical change is set equal to 0.76% in all three sectors. In this case, the NBG observed will be due to capital-deepening given the differences in capital intensities across the three sectors ($\tilde{\alpha}_M = 0.6065$, $\tilde{\alpha}_S = 0.6090$, and $\tilde{\alpha}_A = 0.2647$). The results reported in Table 2.6 suggest that supply-side NBG driven by capital-deepening accounts for very little of the observed changes in sectoral employment or output shares over the study period.

As is mentioned above in Chapter 2.3.2, labor and capital's shares of total cost

Supply-Side Non-Balanced Growth: Kuznets Facts **Alternative Capital/Labor Cost Shares** ($\tilde{\alpha}_M = 0.67$; $\tilde{\alpha}_S = 0.66$; $\tilde{\alpha}_A = 0.46$)						
	1949	2011				
Variable	U.S. Data	U.S. Data	$\tilde{\varepsilon} = 1$	$\tilde{\varepsilon} = 0.85$	$\tilde{\varepsilon} = 0.65$	$\tilde{\varepsilon} = 0.50$
$\psi_M = L_M/L$	0.4016*	0.1638	0.4338	0.4148	0.3834	0.3539
$\psi_S = L_S/L$	0.5117*	0.8245	0.5069	0.5312	0.5699	0.6048
$\psi_A = L_A/L$	0.0868*	0.0116	0.0593	0.0540	0.0342	0.0413
$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.2304	0.5375	0.5169	0.4850	0.4582
$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.7572	0.4213	0.4447	0.4808	0.5108
$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0124	0.0412	0.0383	0.0342	0.0309
*$\psi_{M,1949}$, $\psi_{S,1949}$, and $\psi_{A,1949} = 1 - \psi_{M,1949} - \psi_{S,1949}$ are calculate using (2.50) and (2.51).						

Table 2.7: Supply-Side Non-Balanced Growth: Alternative Capital/Labor Cost Shares

in each sector (i.e., $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, and $\tilde{\alpha}_A$) are calculated using data from 1987 to 2011.[21] The value of capital's share of total costs calculated using this data (i.e., $(1 - \tilde{\alpha}_M) = 0.3935$, $(1 - \tilde{\alpha}_S) = 0.3910$, and $(1 - \tilde{\alpha}_A) = 0.7353$), however, are substantially higher capital cost shares for the U.S. economy as a whole used in previous macroeconomic calibrations. In particular, Gomme and Rupert (2007) report that there is "reasonable agreement" in macroeconomic calibration literature for a value of capital's share of total costs (i.e., the elasticity of output with respect to capital) for the U.S. economy of approximately 0.283.[22] In addition, Valentinyi and Herrendorf (2008) report sector-level cost shares (factor income shares) for capital for the U.S. economy that are lower than those calculated in this essay. In particular, Valentinyi and Herrendorf report $(1 - \tilde{\alpha}_M) = 0.33$, $(1 - \tilde{\alpha}_S) = 0.34$, and $(1 - \tilde{\alpha}_A) = 0.54$. In order to investigate the impact of our calculated values $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, and $\tilde{\alpha}_A$ on the calibrated models performance, Table 2.7 reports results using Valentinyi and Herrendorf's estimates of capital and

[21] Recall that data on "compentation of employees" by industry is only available from 1987 to 2011 using NAICS industry definitions.

[22] Citing Gomme and Rupert (2007) , Iscan (2010) used 0.283 for capital's share of total of cost for the argicultural, manufacturing, and service sectors in his calibration of the U.S. economy.

labor's cost shares in the Agricultural, Manufacturing, and Service sectors.

Comparing the baseline results from Table 2.5 with the results reported in Table 2.7 suggests that the model calibration's ability to match changes in sectoral employment and output shares is not substantially different when Valentinyi and Herrendorf (2008)'s estimates of $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, and $\tilde{\alpha}_A$ are used. As is to be expected, the fact that labor's cost share in Agriculture is larger in Valentinyi and Herrendorf (2008)'s estimates means that the calibrated model predicts a smaller movement of labor out of Agriculture in Table 2.7 compared to the baseline results. This increase in employment in Agriculture comes at the expense of employment in the Service Sector. Predicted Service Sector employment is reduced when Valentinyi and Herrendorf's estimates are used because the Service Sector is less labor intensive relative to the Manufacturing and Agricultural sectors compared to the baseline estimates of $\tilde{\alpha}_M$, $\tilde{\alpha}_S$, and $\tilde{\alpha}_A$. Given the similarity between the results reported in Tables 2.5 and Table 2.7, the remaining calibrations reported in this essay will continue using the baseline values of $\tilde{\alpha}_M = 0.6065$, $\tilde{\alpha}_S = 0.6090$, and $\tilde{\alpha}_A = 0.2647$.

Supply-Side Non-Balanced Growth: Kaldor Facts							
		1949 – 2011					
$\tilde{\varepsilon}$	Variable	Mean	Min.	Max.	Variance	Time Trend	p-Value
0.85	Growth Rate of Output ($\hat{Y}_t$)	0.0342	−0.0122	0.0835	7.680×10^{-5}	-7.306×10^{-5}	0.2432
	Real Interest Rate (R_t)	0.0406	0.0368	0.0495	9.549×10^{-6}	-1.621×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	10.76	8.869	11.76	0.5718	4.045×10^{-2}	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.4344	0.4219	0.4394	4.687×10^{-6}	-5.644×10^{-5}	0.0000
0.50	Growth Rate of Output ($\hat{Y}_t$)	0.0349	−0.0100	0.0831	7.362×10^{-5}	-8.848×10^{-5}	0.1477
	Real Interest Rate (R_t)	0.0445	0.0378	0.0586	2.650×10^{-5}	-2.743×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	9.883	7.678	11.45	1.115	0.0573	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.4347	0.4219	0.4496	8.883×10^{-5}	-8.579×10^{-5}	0.0000
Calculated Using U.S. Data							
0.50	Growth Rate of Output ($\hat{Y}_t$)	0.0284	−0.0400	0.1102	8.103×10^{-4}	-4.382×10^{-4}	0.0288
	Real Interest Rate (R_t)	0.0694	0.0452	0.1123	4.855×10^{-4}	-6.855×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	8.233	7.492	9.501	0.2366	0.0176	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.2877	0.1700	0.4645	0.0084	-4.881×10^{-3}	0.0000
Calculated Using U.S. Data – Hodrick-Prescott Filter							
0.50	Growth Rate of Output ($\hat{Y}_t$)	0.0280	0.0031	0.0436	8.064×10^{-4}	-3.581×10^{-4}	0.0000
	Real Interest Rate (R_t)	0.0694	0.0466	0.1096	4.851×10^{-4}	-6.904×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	8.221	7.645	8.919	0.1840	0.0176	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.2884	0.1710	0.4598	0.0084	-4.921×10^{-3}	0.0000

Table 2.8: Supply-Side Non-Balanced Growth: Kaldor Facts Results

Supply-Side Non-Balanced Growth: Kaldor Facts Results

Table 2.8 describes the calibrated model's ability to match the Kaldor facts for the case of supply-side NBG. Recall that the Kaldor facts describe the relative constancy of (i) the growth rate of output, (ii) the real interest rate, (iii) the capital-to-output ratio, and (iv) the share of capital income in total output. Table 2.8 evaluates the calibrated model's ability to match the Kaldor facts on two dimensions: the variance of the four variables relative to their means and whether the variables exhibit significant time trends. Note that the last four rows of Table 2.8 report results when the four variables relevant to the Kaldor facts are calculated using NIPA data on real output by sector and total capital stock for the case when $\tilde{\varepsilon} = 0.5$ (and, because only supply-side NBG is considered, $\tilde{\gamma}_M = \tilde{\gamma}_S = \tilde{\gamma}_A = 0$).[23]

Table 2.8 shows that the predicted values of all four variables relevant to the Kaldor facts have low variances relative to their means. In particular, the coefficients of variation (the ratio of standard deviation to the mean) are low for all four variables: 0.2562 for the growth rate of output, 0.0761 for the real interest rate, 0.0703 for the capital-to-output ratio, and 0.0050 for the share of capital income in total output.[24] Moreover, the coefficients of variation for all four of these variables are higher – i.e., have higher standard deviations relative to their means – when they are

[23] The values of Y_t from (2.35) and R_t from (2.37) depend on the values of real output by sector and total capital stock, as well as $\tilde{\varepsilon}$, $\tilde{\gamma}_M$, $\tilde{\gamma}_S$, $\tilde{\gamma}_A$, $\tilde{\mu}_M$, and $\tilde{\mu}_S$. Recall that $\tilde{\mu}_M$ and $\tilde{\mu}_S$ are determined by $\tilde{\varepsilon}$, $\tilde{\gamma}_M$, $\tilde{\gamma}_S$, and $\tilde{\gamma}_A$ and NIPA data on nominal output by sector in the base year through (2.55) and (2.56).

[24] In general, the coefficient of variation should be computed for data measured on a ratio scale, i.e., for data that only take on non-negative values and that include a zero value. While none of the four variables relevant for the Kaldor facts satisfy these conditions – the growth rate of output can take on non-zero values and the other three variables cannot reasonably take on a value of zero – they come close enough to holding that the coefficient of variation is still useful in describing the magnitude of the variance in the data relative to the mean.

calculated using NIPA data on the U.S. economy (1.002, 0.3179, 0.0599, and 0.3186, respectively).

It is not surprising that the coefficients of variation are higher for all four variables relevant for the Kaldor facts when they are calculated using U.S. data. This is because the U.S. data contains short-term fluctuations related to the business cycle and idiosyncratic factors that are not considered in the calibrated model. In an attempt to control for short-term fluctuations, Table 2.8 reports values for the four variables calculated after applying a Hodrick-Prescott (HP) filter to the U.S. data. HP filters are used to remove short-term fluctuations and cyclical components from time series (Hodrick and Prescott 1998). The coefficients of variation for the four variables when the HP filter is used are: 1.014 for the growth rate of output, 0.3174 for the real interest rate, 0.0522 for the capital-to-output ratio, and 0.3178 for the share of capital income in total output.[25] These coefficients of variation are higher than those for the calibrated model's predicted values of the four variables, which suggests that the lower variances of the calibrated model's predicted values cannot be attributed entirely to short-term fluctuations in the U.S. data.

Table 2.8 also shows that while the model's predictions for the four variables relevant to the Kaldor facts have relatively low variance, three of the four variables exhibit a significant time trend. In particular, Table 2.8 reports results for four regressions where the predicted time series for each of the four variables is regressed on a constant and a time trend. The coefficient and p-value for the time trend are reported in Table 2.8 for each of the four regressions. In each parameterizations of the model considered, the real interest rate and the share of capital income in total output exhibit a slight but statistically significant decline over time, the capital-to-output ratio exhibits a slight but statistically significant secular increase, and the

[25]The value of the HP filters "smoothing parameter" is set equal to 100, which is the value suggested by Hodrick and Prescott (1998) for annual data.

growth rate of output does not exhibit a statistically significant time trend. These same time trends, however, are also observed when the four variables are calculated using NIPA data on the U.S. economy. This fact, combined with the low variances of the variables relative to their means, suggests that overall the model of supply-side NBG is relatively consistent with the Kaldor facts as they apply to the U.S. economy over the study period.

Table 2.8 also show that calibrated model's predictions for the means of the four variables relevant to the Kaldor facts are significantly different than the means calculated using NIPA data. In particular, the model calibration's predictions for the mean growth rate of output (3.5% v. 2.8%), the capital-to-output ratio (9.883 v. 8.233), and the share of capital income in total output (0.4347 v. 0.2877) are higher than the mean values of these variables calculated using U.S. data, while the predicted value of the rental rate of capital is lower (4.5% v. 8.2%).[26] The differences between the model calibrations predictions for these variables and their values computed using NIPA data can be attributed to the fact that the model calibration predicts higher rate of economic growth than is observed in the U.S. data. The model calibration overstates economic growth because it model predicts larger factor shares for the progressive Manufacturing and Agricultural sectors than occurred in reality. Larger input shares in the Agricultural and Manufacturing sectors increase the impact of higher rates of Hick's neutral technical change in the two sectors on the overall growth rate of output. The higher predicted rate of economic growth leads to greater capital accumulation in the calibrated model than occurred in the U.S. data, which causes the calibration to overstate the capital-to-output ratio and the share of capital income in

[26] The 8.2% figure for the average rate of return on capital calculated using the NIPA data is inline with previous studies that have examined rates of return on capital in the United States over the study period. For example, Poterba (1998) finds an average return on capital of 8.6% for the period 1959–1996 and Siegel (1992) finds an average return on capital of 7.77% for the period 1800–1990.

total output, and understate the rate of return on capital.

The results in this section suggest that the calibrated model of supply-side NBG can produce equilibrium dynamics that are relatively consistent with the Kaldor facts while at the same time producing large changes in sectoral output and employment shares. The model's predicted changes in employment shares, however, are smaller in magnitude than those observed in the U.S. data over the study period. Furthermore, any model that only considers supply-side mechanisms for NBG is not capable, by construction, of matching the Kuznets facts in terms of the expansion of the Service sector's shares of both output and employment relative to the Agricultural and Manufacturing sectors. The next section investigates whether demand-side explanations for NBG can overcome these shortcomings of supply-side explanations for NBG, while also still matching the Kaldor facts.

2.4.2 Demand-Side Non-Balanced Growth: Results

Demand-side explanations for NBG posit that NBG is a result of changes in the composition of demand for sectoral output that occur as the economy grows. In this essay, changes in the composition of demand for sectoral output occur as a result of differences in the elasticities of demand for sectoral output with respect to the total production of the unique final good (i.e., total output in the economy), where, as is described in Chapter 2.2, the unique final good can be consumed by households or used to produce capital stock. NBG occurs in the model because growth in sectoral employment and capital shares will be biased towards the sector(s) whose output has a higher elasticity of demand with respect to the unique final good.

In this essay, it is assumed that $\gamma_A > 0$, $\gamma_M = 0$, and $\gamma_S < 0$. These assumptions imply an elasticity of demand for Agricultural sector output with respect to the unique final good less than 1, an elasticity of demand for Manufacturing sector output with respect to total output equal to 1, and an elasticity of demand for Service sector output with respect to total output greater to 1. That is, these assumptions imply that demand for sectoral output will be biased away from Agriculture and towards Services as the economy grows. The values of $\tilde{\gamma}_A$ and $\tilde{\gamma}_S$ used in the calibration are selected so that $\tilde{\gamma}_A$ is set equal to 25% of total output in the Agricultural sector in the base year (i.e., $\tilde{\gamma}_A/Y_{A,1949} \times 100\% = 25\%$) and $|\gamma_S|$ is set equal to 25%, 50%, or 75% of total output in the Service sector output in the base year ($|\tilde{\gamma}_S/Y_{S,1949}| = 25\%$, 50%, or 75%).[27]

Demand-Side Non-Balanced Growth: Kuznets Facts Results

Table 2.9 presents results on the calibrated model's ability to match the Kuznets facts for the case when only the demand-side NBG is considered (i.e., when $\tilde{\varepsilon} = 0$). Table 2.9 show that the ability of the model to match the Kuznets facts improves as both $|\tilde{\gamma}_S|$ and $\tilde{\gamma}_A$ are increased. As expected, higher values of $|\tilde{\gamma}_S|$ increase the

[27] In models that have emphasized non-homothetic consumer preferences in their explanations for demand-side NBG (Kongsamut, Rebelo, and Xie 2001; Gollin, Parente, and Rogerson 2002; Iscan 2010), non-homothetic preferences are explained by a positive "consumption requirement" for agricultural sector output and a negative consumption requirement for service sector output (these models typically assume no consumption requirement for manufacturing sector output). A positive consumption requirement for agricultural sector output is meant to reflect the fact that a minimum level of agricultural output is required to meet a populations subsistence needs. This subsisitence requirement will cause Agricultural output to comprise a large share of total output in the early stages of development, but also implies that the demand for Agricultural output will increase less than proportionaty to the increases in societal expenditure as the economy grows. Conversely, the negative consumption requirement for services is meant to reflect the basic endowment of service in the economy due to the home production of services. This basic endowment of services will cause service sector output to comprise a small share of total output (as reported in the national accounts) in the early stages of development, but also implies that as the economy develops households will purchase an increasing number of services outside of the home, causing the demand for service sector output to increase more than proportionately to increases in societal expenditure.

Demand-Side Non-Balanced Growth: Kuznets Facts							
		1949	2011				
$\tilde{\gamma}_A/Y_{A,1949}$				$\lvert\tilde{\gamma}_S/Y_{S,1949}\rvert \times 100\%$			
$\times 100\%$	Variable	U.S. Data	U.S. Data	0%	25%	50%	75%
0%	$\psi_M = L_M/L$	0.4213*	0.1638	0.4377	0.3947	0.3631	0.3382
	$\psi_S = L_S/L$	0.5390*	0.8245	0.5244	0.5710	0.6054	0.6327
	$\psi_A = L_A/L$	0.0397*	0.0116	0.0379	0.0343	0.0315	0.0291
	$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.2304	0.5411	0.4964	0.4619	0.4337
	$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.7572	0.4174	0.4646	0.5012	0.5311
	$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0124	0.0414	0.0390	0.0369	0.0351
50%	$\psi_M = L_M/L$	0.4213*	0.1638	0.4458	0.4003	0.3672	0.3410
	$\psi_S = L_S/L$	0.5390*	0.8245	0.5333	0.5808	0.6155	0.6428
	$\psi_A = L_A/L$	0.0397*	0.0116	0.0209	0.0189	0.0173	0.0160
	$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.2304	0.5474	0.5003	0.4643	0.4352
	$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.7572	0.4283	0.4771	0.5144	0.5447
	$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0124	0.0242	0.0227	0.0213	0.0202
*$\psi_{M,1949}$, $\psi_{S,1949}$, and $\psi_{A,1949} = 1 - \psi_{M,1949} - \psi_{S,1949}$ are calculate using (2.50) and (2.51).							

Table 2.9: Demand-Side Non-Balanced Growth: Kuznets Facts Results

growth in the Service sector's employment share over the study period, while higher values of $\tilde{\gamma}_A$ lead to larger declines in the Agricultural sector's employment share. The model, however, predicts a substantially smaller shift in employment towards Services than occurred in reality. In particular, under the best performing parametrization ($\tilde{\gamma}_A/Y_{A,1949} \times 100\% = 50\%$ and $\lvert\tilde{\gamma}_S/Y_{S,1949}\rvert \times 100\% = 75\%$), the model calibration explains 36% of the movement of labor into Services over the study period (the equivalent figure for the model of supply-side NBG considered in the previous section is 29%). The model's predictions concerning changes in the Agricultural sector's employment share, however, are improved considerably under the assumption of demand-side NBG. Under the best performing parameterization, the model of demand-side NBG explain 84% of the decline in Agricultural sector's employment share over the study period (compared to the only 33% for the model of supply-side NBG).

Table 2.9 shows that the model calibration predicts that the Agricultural and

Manufacturing sector's output shares rise over the study period, while the Service sector's output share declines, which is the exact opposite of what occurred in reality. The model's failure the match the direction of change in sectoral output shares is due to the fact that the model does not produce a sufficient movement of labor and capital in Services to offset the lower rate of Hick's neutral technical change in the sector and allow the sector's output share to increase. Recall that for supply-side NBG, the mechanism for NBG itself precluded increases in both the Service sector's output and employment shares. In contrast, the failure of the model of demand-side NBG considered in this section to match the changes in sectoral output shares is simply the consequences of the economic forces driving NBG in the model not producing a sufficiently large movement of factors of production into the Service sector.

Demand-Side Non-Balanced Growth: Kaldor Facts ($\tilde{\gamma}_A/Y_{A,1949} \times 100\% = 50\%$)							
$\lvert\tilde{\gamma}_S/Y_{S,1949}\rvert$		1949 − 2011					
$\times 100\%$	Variable	Mean	Min.	Max.	Variance	Time Trend	p-Value
25%	Growth Rate of Output ($\hat{Y}_t$)	0.0331	−0.0107	0.0769	6.413×10^{-5}	-5.808×10^{-5}	0.31075
	Real Interest Rate (R_t)	0.0422	0.0358	0.0559	2.719×10^{-5}	-2.678×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	9.703	7.215	11.43	1.48	0.0660	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.3957	0.3836	0.4024	1.901×10^{-4}	2.340×10^{-4}	0.0000
75%	Growth Rate of Output ($\hat{Y}_t$)	0.0298	−0.0087	0.0619	4.171×10^{-5}	-2.623×10^{-5}	0.57147
	Real Interest Rate (R_t)	0.0443	0.0359	0.0610	4.360×10^{-5}	-3.395×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	8.416	5.420	10.88	2.696	0.0893	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.3523	0.3103	0.3812	4.354×10^{-4}	1.128×10^{-3}	0.0000
Calculated Using U.S. Data							
75%	Growth Rate of Output ($\hat{Y}_t$)	0.0266	−0.0382	0.0762	4.686×10^{-4}	-2.288×10^{-4}	0.1377
	Real Interest Rate (R_t)	0.0541	0.0452	0.0647	2.449×10^{-5}	-2.415×10^{-4}	0.0000
	Capital-to-Output Ratio (K_t/Y_t)	6.399	5.369	7.310	0.2589	0.0213	0.0000
	Share of Capital Income in Output (R_tK_t/Y_t)	0.3439	0.3147	0.3730	2.115×10^{-4}	-3.321×10^{-4}	0.0006

Table 2.10: Demand-Side Non-Balanced Growth: Kaldor Facts Results

Demand-Side Non-Balanced Growth: Kaldor Facts Results

Table 2.10 describes the calibrated model's ability to match the Kaldor facts for the case of demand-side NBG. Table 2.10 shows that the calibrated model's ability to match the Kaldor facts for the case of demand-side NBG is similar to the case of supply-side NBG presented in Table 2.8. In particular, the predicted values of all four variables relevant to the Kaldor facts have low variances relative to their means (both in absolute terms and relative to the means and variances of the four variables when they are calculated using NIPA data), and the three of the four variables (the interest rate, the share of capital income, and the capital-to-output ratio) display slight but statistically significant time trends that are also observed when the four variables are calculated using NIPA data. Together, Tables 2.8 and 2.10 suggest that the model is capable of approximating the Kaldor facts as they apply to the U.S. economy over the study period under the assumptions of supply- and demand-side NBG and for a wide range of different parameterizations.

Table 2.10 indicates that the model predicts a higher mean growth rate output compared to when the growth rate is calculated using U.S. data (3.0% v. 2.7%). As with the case of supply-side NBG, the model predicts a higher growth rate of output because the it predicts larger factor shares for the progressive Manufacturing and Agricultural sectors than occurred in reality, and these larger factors share lead to a higher effective rate of technical process in the economy. As in the case of supply-side NBG, the higher predicted rate of economic growth leads to greater capital accumulation in the calibrated model than occurred in the U.S. data, which causes the calibration to overstate the capital-to-output ratio (8.416 v. 6.399) and the share of capital income in total output (0.3523 v. 0.3439), and understate the rate of return on capital (4.4% v. 5.4%).

The results in this section suggest that, like the calibrated model of supply-side NBG presented in the previous section, the calibrated model of demand-side NBG can produce equilibrium dynamics that are relatively consistent with the Kaldor facts while at the same time producing significant changes in sectoral output and employment shares. In addition, the model of demand-side NBG presented in this sections improves on the model of supply-side NBG in its ability to reproduce the decline in the Agricultural sector's employment share over the study period. The model, however, predicts a much smaller increase in the Service sector's employment share than is observed in the U.S. data over the study period, and this shortcoming causes the model to predict a decline in the Service sector's output share and a rise in the Manufacturing and Agricultural sector's output shares, which is counter to the data. The next section investigates whether simultaneously considering supply- and demand-side explanations for NBG can overcome this shortcoming and match the Kuznets facts concerning both changes in output and employment shares, while still matching the Kaldor facts.

2.4.3 Joint Model: Results

Tables 2.11 and 2.13 report results on the model calibration's ability to match the Kuznets and Kaldor facts, respectively, when both supply- and demand-side mechanisms for NBG are considered simultaneously. The results on Tables 2.11 and 2.13 for the joint model of supply- and demand-side NBG assume the parameter values that produced the best fits to the observed data on the U.S. economy when supply- and demand-side NBG were considered individually (i.e., $\tilde{\varepsilon} = 0.50$, $|\tilde{\gamma}_S/Y_{S,1949}| \times 100\% = 75\%$, $\tilde{\gamma}_A/Y_{A,1949} \times 100\% = 50\%$, and $\tilde{\gamma}_M = 0$). The values of the other parameters in this section are set at their baseline levels described in Chapter 2.3.2.

Joint Model: Kuznets Facts Results

Table 2.11 shows that of all the model parameterizations considered in this essay, the joint model of supply- and demand-side NBG provides the best fit to the U.S. data

Joint Model: Kuznets Facts $\tilde{\varepsilon}=0.50$, $\tilde{\gamma}_A/Y_{A,1949}\times 100\% = 50\%$, $\tilde{\gamma}_M=0$, and $\lvert\tilde{\gamma}_S/Y_{S,1949}\rvert\times 100\% = 75\%$							
	1949	1969		1990		2011	
Variable	U.S. Data	U.S. Data	Model	U.S. Data	Model	U.S. Data	Model
$\psi_M = L_M/L$	0.4213*	0.3795	0.3484	0.2505	0.2919	0.1638	0.2570
$\psi_S = L_S/L$	0.5390*	0.5970	0.6300	0.7344	0.6946	0.8245	0.7332
$\psi_A = L_A/L$	0.0397*	0.0235	0.0216	0.0151	0.0135	0.0116	0.0098
$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.3858	0.3683	0.2838	0.3480	0.2304	0.3476
$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.6024	0.6171	0.7034	0.6390	0.7572	0.6392
$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0119	0.0146	0.0128	0.0130	0.0124	0.0133
p_S/p_M **	0.8764	1.322	1.064	0.6975	1.173	0.4543	1.622
p_S/p_A **	0.1984	0.1880	0.2927	0.3590	0.4271	0.5601	0.7412

*$\psi_{M,1949}$, $\psi_{S,1949}$, and $\psi_{A,1949} = 1-\psi_{M,1949}-\psi_{S,1949}$ are calculate using (2.50) and (2.51).
**Relative output prices are calculated using (2.11).

Table 2.11: Joint Model of Supply- and Demand-Side NBG: Kuznets Facts Results

regarding the Kuznets facts. Regarding sectoral employment shares, the joint model explains 68% of the increase in the Service sector's employment share, 64% of the decline in the Manufacturing sector's decline in employment share, and predicts a 6% larger decline in the Agricultural sector's employment share than is observed in the data in the period 1949 to 2011. These results comport with Iscan (2010), who finds that, taken together, supply- and demand-side explanations for NBG explain roughly two-thirds of the reallocation of labor into the Service sector from the Manufacturing and Agricultural sectors in the United States from 1800 to 2000.

Concerning output shares, Table 2.11 shows that the joint model predicts sufficient movement of labor and capital into Services for the sector's output share to increase, and for the output shares of the Agricultural and Manufacturing sectors to decline. As such, unlike when both explanations for NBG are considered individually, the joint model is able to qualitatively match the Kuznets facts regarding changes in both employment and output shares. The joint model, however, only explains 31% of the decrease in the Manufacturing sector's output share, 87% of the decrease in the Agricultural sector's output share, and 33% of the increase in the Service sector's output share.

Table 2.11 reports the model calibration's predictions for three years: 1969, 1990, and 2011 (these years correspond to dividing the 63 year study period into three 21-year segments). Table 2.11 shows that the joint model overpredicts the growth in the Service sectors output and employment shares in the first 21 years of the calibration, and then underpredicts growth in these two variables in the latter two 21-year segments. An important reason why the joint model is not able to reproduced the magnitude of changes in output and employment shares in the later periods of the calibration is that the influence of $\tilde{\gamma}_S$ and $\tilde{\gamma}_A$ on changes in sectoral output and employment shares is most significant in the early periods of the calibration, when the values of these two parameters are large relative to real output in the Service and Agricultural sectors. In particular, as $\tilde{\gamma}_S$ and $\tilde{\gamma}_A$ become small in magnitude relative to output in the Service and Agricultural sectors in the latter part of the calibration, the elasticities of demand for Service and Agricultural sector output with respect to the unique final good to become closer to one, which reduces the importance of demand-side NBG (i.e., exogeneous changes in the compostion of demand for sectoral output) as a driver of changes in sectoral output and employment shares in the model. One way for the joint model to match the large movement of employment and output shares towards Services that is observed in the U.S. data in the latter part of the study period would be to include a second large, delayed exogenous shift in the composition of demand for sectoral output towards Services and away from Agriculture in the model.

To check the robustness of the joint model results reported in Table 2.11 to the assumption that $\tilde{\mu}_M$ and $\tilde{\mu}_S$ are chosen so that (2.55) and (2.56) hold in the initial year of the calibration (i.e., 1949), separate values of $\tilde{\mu}_M$ and $\tilde{\mu}_S$ were calculated using U.S. data on sectoral nominal and real value added for each year of the study period. Using the values of $\tilde{\varepsilon}$, $\tilde{\gamma}_M$, $\tilde{\gamma}_S$, and $\tilde{\gamma}_A$ from the joint model, the calculated values

Joint Model: Kuznets Facts **Sensitivity Analysis on μ_M and μ_S** $\tilde{\varepsilon} = 0.50$, $\tilde{\gamma}_A/Y_{A,1949} \times 100\% = 50\%$, $\tilde{\gamma}_M = 0$, and $\lvert\tilde{\gamma}_S/Y_{S,1949}\rvert \times 100\% = 75\%$						
	1949	2011				
Variable	U.S. Data	U.S. Data	$\tilde{\mu}_S = \tilde{\mu}_{S,1949}$**	$\tilde{\mu}_S = \widehat{\tilde{\mu}_S}$	$\tilde{\mu}_S = \tilde{\mu}_S^{\min}$	$\tilde{\mu}_S = \tilde{\mu}_S^{\max}$
$\psi_M = L_M/L$	0.4213*	0.1638	0.2570	0.2580	0.2896	0.2150
$\psi_S = L_S/L$	0.5390*	0.8245	0.7332	0.7331	0.7026	0.7755
$\psi_A = L_A/L$	0.0397*	0.0116	0.0098	0.0090	0.0078	0.0095
$\frac{Y_M}{Y_M+Y_S+Y_A}$	0.3996	0.2304	0.3476	0.3461	0.3754	0.3006
$\frac{Y_S}{Y_M+Y_S+Y_A}$	0.5809	0.7572	0.6392	0.6419	0.6141	0.6886
$\frac{Y_A}{Y_M+Y_S+Y_A}$	0.0195	0.0124	0.0133	0.0121	0.0105	0.0104

*$\psi_{M,1949}$, $\psi_{S,1949}$, and $\psi_{A,1949} = 1 - \psi_{M,1949} - \psi_{S,1949}$ are calculate using (2.50) and (2.51).
**The baseline assumption used throughout this essay is that $\tilde{\mu}_M$ and $\tilde{\mu}_S$ are chosen so that (2.55) and (2.56) hold in the initial year of the calibration (1949).

Table 2.12: Joint Model: Kuznets Facts - Sensitivity Analysis

of $\tilde{\mu}_M$ and $\tilde{\mu}_S$ in the initial year (0.1498 and 0.8498) are close to their means over the study period (0.1506 and 0.8491). The calculated values of $\tilde{\mu}_M$ and $\tilde{\mu}_S$, however, vary over the study period. In particular, $\tilde{\mu}_M$ ranges from 0.0709 and 0.2278 with a standard deviation of 0.0479 , while $\tilde{\mu}_S$ ranges from 0.7720 and 0.9236 with a standard deviation also of 0.0479. Given this variation in the calculated values of $\tilde{\mu}_M$ and $\tilde{\mu}_S$, Table 2.12 reports results for the joint model with $\tilde{\mu}_S$ set equal to its mean, maximum value, and minimum value (conveniently, $\tilde{\mu}_M$ attains its maximum value in the year that $\tilde{\mu}_S$ attains its minimum value, and vice versa).

Table 2.12 demonstrates that the joint model is able to qualitatively match the Kuznets facts regarding changes in both employment and output shares for the full range of values of $\tilde{\mu}_M$ and $\tilde{\mu}_S$ implied by the U.S. data over the study period. This implies that the main results in this section do not depend on the assumption for how $\tilde{\mu}_M$ and $\tilde{\mu}_S$ are calculated. In addition, Table 2.12 shows that using the mean values for $\tilde{\mu}_M$ and $\tilde{\mu}_S$ has almost no effect on the joint model's predicted changes in sectoral labor and output shares. This results is not surprising given that the calculated values of $\tilde{\mu}_M$ and $\tilde{\mu}_S$ in the initial year are so close to their means over the study period.

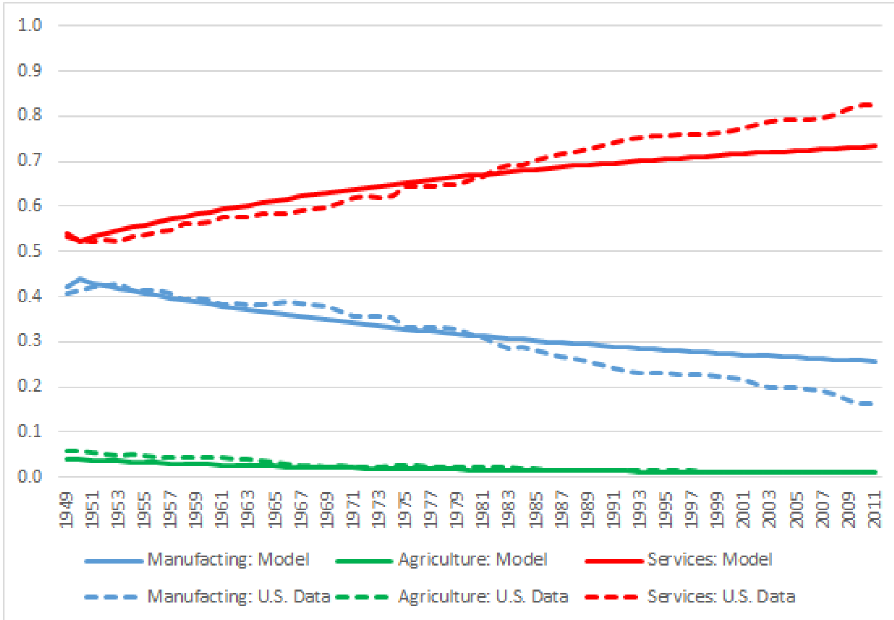

Figure 2.1: Joint Model Results: Employment Shares by Sector

Table 2.12 also shows that when $\tilde{\mu}_S^{\min} = 0.7720$ is used, the joint model predicts weaker growth in the Service Sector's employment and output shares compared to the baseline model, and that the opposite result holds when $\tilde{\mu}_S^{\max} = 0.9236$ is used. Indeed, the joint model does a better job of matching the Kuznets facts for the U.S. economy relative to the baseline model when $\tilde{\mu}_S^{\max}$ is used.

Figures 2.1 and 2.2 confirm that the joint model overpredicts the growth in the Service sector's output and employment shares in the early periods of the calibration relative to the U.S. data, but underpredicts the growth in these two variables in later periods. Indeed, the joint model predicts almost no change in either the Manufacturing or Service sector's output shares in the second half of the study period, while the Service sectors' output share continued to increase at the expense of the Manufacturing sector in the U.S. data. Figures 2.1 and 2.2 show that in the later stages of the calibration, when the importance of demand-side NBG as a driver of changes in sectoral output and employment shares has declined, the predictions of the joint model begin to resemble those of the model of supply-side NBG considered

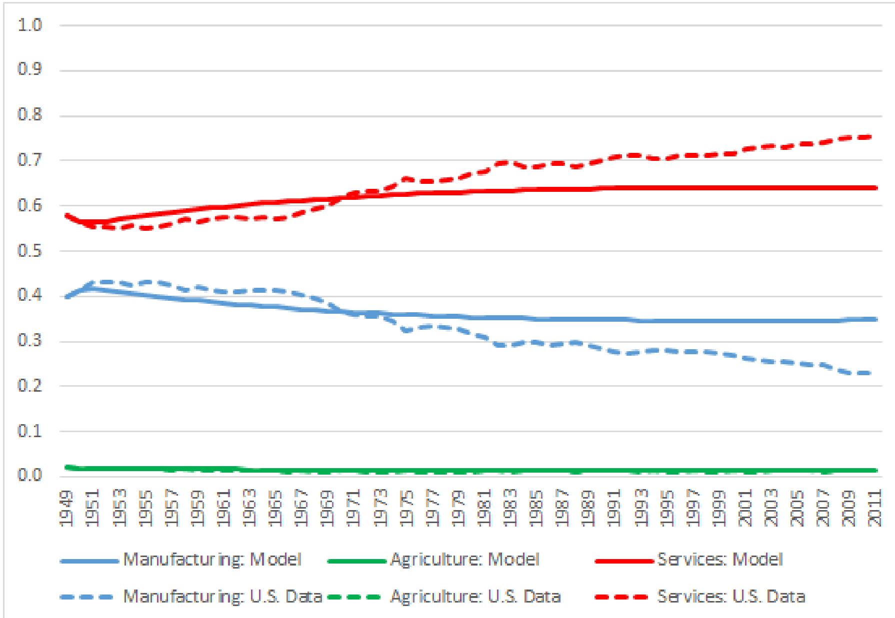

Figure 2.2: Joint Model Results: Output Shares by Sector

in Chapter 2.4.1, which underpredicts the growth in the Service sector's employment share relative to the U.S. data, and predicts a slight decline in the Service sector's output share over time.

Table 2.11 also shows that the joint model predicts that the price of Service sector output is increasing relative to the prices of Manufacturing and Agricultural sector output over the study period. This increase in the relative price of Service sector output is necessary to attract factors of production into the sector despite its lower rate of Hick's neutral technical change. Buera and Kaboski (2009) identify the inability of their model of NBG to generate the simultaneous increases in the relative price and output share of the Service sector relative to the Manufacturing and Agricultural sectors that occurred in the post-war United States as a shortcoming of supply- and demand-side explanations of NBG. The results reported here suggest that the inability of Buera and Kaboski's model to match changes the relative prices and output shares across sectors is an artifact of their assumptions and parameterization, rather than a more general failing of supply- and demand-side explanations of NBG to qualitatively match the Kuznets facts for the U.S. economy.[28]

[28] Following Caselli and Coleman (2001), Buera and Kaboski (2009)'s constuct thier model in large part to explain the deviations between the changes in sectoral employment and nominal values added (nominal output) shares in the agricultural, manufactuing, and service sectors in the late 19^{th} and early 20^{th} centuries. In particular, Buera and Kaboski (2009) evalute whether the supply- and demand-side explanations of NBG can explain the persistently high employment share in agriculture relative to its share of nominal value added over the period.

Joint Model: Kaldor Facts

$\tilde{\varepsilon} = 0.50$, $\tilde{\gamma}_A/Y_{A,1949} \times 100\% = 50\%$, $\tilde{\gamma}_M = 0$, and $|\tilde{\gamma}_S/Y_{S,1949}| \times 100\% = 75\%$

	1949 – 2011					
Variable	Mean	Min.	Max.	Variance	Time Trend	p-Value
Model Calibration Results						
Growth Rate of Output ($\hat{Y}_t$)	0.0304	−0.0071	0.0611	3.967×10^{-5}	-4.804×10^{-5}	0.2862
Real Interest Rate (R_t)	0.0507	0.0385	0.0688	7.326×10^{-5}	-4.637×10^{-4}	0.0000
Capital-to-Output Ratio (K_t/Y_t)	7.313	4.505	10.10	2.863	0.0923	0.0000
Share of Capital Income in Output (R_tK_t/Y_t)	0.3565	0.2978	0.3890	6.383×10^{-4}	1.349×10^{-3}	0.0000
Calculated Using U.S. Data						
Growth Rate of Output ($\hat{Y}_t$)	0.0264	−0.0433	0.0766	4.764×10^{-4}	-2.375×10^{-4}	0.1263
Real Interest Rate (R_t)	0.0652	0.0484	0.0812	8.309×10^{-5}	-4.590×10^{-4}	0.0000
Capital-to-Output Ratio (K_t/Y_t)	5.338	4.473	6.097	0.1814	0.0183	0.0000
Share of Capital Income in Output (R_tK_t/Y_t)	0.3450	0.2890	0.3935	7.757×10^{-4}	-1.183×10^{-3}	0.0000

Table 2.13: Joint Model of Supply- and Demand-Side NBG: Kaldor Facts Results

Joint Model: Kaldor Facts Results

Table 2.13 illustrates the joint model's ability to match the Kaldor facts. The results reported on Table 2.13 show that the joint model's behavior concerning the Kaldor facts is almost identical to the results when only supply-side NBG is considered (Table 2.8) and when only demand side NBG is considered (Table 2.10). In particular, the predicted values of all four variables relevant to the Kaldor facts have low variances relative to their means, the three of the four variables (the interest rate, the share of capital income, and the capital-to-output ratio) display slight but statistically significant time trends that are also observed when the four variables are calculated using NIPA data, and the model calibration predicts a slightly higher mean growth rate (and thus a higher rate of capital accumulation) than is observed in the U.S. data. Taken together with the results Tables 2.8 and 2.10, the results on Table 2.13 further confirm that the model is capable of approximating the Kaldor facts as they apply to the U.S. economy over the study period for a wide range of different parameterizations.

The results in this section show that the joint model that considers both supply- and demand-side mechanisms for NBG is capable of producing equilibrium dynamics that qualitatively match the Kuznets and the Kaldor facts as they apply to the U.S. economy over the study period. Unlike the case when only demand-side NBG is considered, the joint model predicts sufficient movement of labor and capital into Services for the sector's output share to increase, and for the output shares of the Agricultural and Manufacturing sectors to decline. The model calibration, however, predicts a smaller increase in the Service sector's employment share and a smaller decrease in the Manufacturing sector's employment share than is observed in the U.S. data over the study period in the latter stages of the calibration. It is argued that this shortcoming of the model is due to the fact that the parameters $\tilde{\gamma}_S$ and $\tilde{\gamma}_A$ exert a stronger influence on changes in sectoral output and employment shares in the early stages of the calibration, when their values are relatively close to the values for real output in the Service and Agricultural sectors.

2.5 Conclusions

This essay developed a dynamic general equilibrium model that integrates supply-side and demand-side explanations for NBG and calibrated it using data on the post-war U.S. economy (1949 to 2012) from the U.S. National Income and Product Accounts. This essay evaluated the extent that these two complementary explanations for NBG – both alone and in combination – can generate changes in sectoral output and employment shares between the agricultural, manufacturing, and service sectors that are consistent with the Kuznets facts as they occurred in the United States, while simultaneously predicting approximately constant values of the (i) the growth rate output, (ii) the real interest rate, (iii) the capital-to-output ratio, and (iv) the share of capital income in total output consistent with the Kaldor facts. Following previous studies, supply-side NBG occurs in the model as a result of differential productivity growth between sectors, while demand-side NBG occurs as the result of changes in the composition of demand for sectoral output that take place as the economy grows.

It is shown that for a set of realistic parameter values, the calibrated model is capable of generating changes in sectoral output and employment shares that are qualitatively consistent with the Kuznets facts as they occurred in the United States over the study period. The model, however, is only capable of qualitatively matching the Kuznets facts when the supply-side and demand-side mechanisms are considered simultaneously. Individually, both mechanisms are only able to match observed changes in sectoral employment shares. In the case of supply-side NBG, the inability of the model to match observed changes in sectoral output shares is a consequence of the mechanism for supply-side NBG itself, which precludes concurrent increases in a sector's output and employment shares, such as occurred in the service sector over the study period. In the case of demand-side NBG, the inability of the model to match observed changes in sectoral output shares is due to the fact that the model predicts smaller increases in the service sector's employment and capital shares than occurred in reality. The increases in the service sector's factor shares predicted by the model are not sufficient to offset the lower rate of productivity growth in the sector, and thus cause the model to predict that the sector's output share will fall over time.

On the other hand, the joint model that considers both supply- and demand-side mechanisms for NBG predicts sufficient movement of labor and capital into services for the sector's output share to increase, and for the output shares of the agricultural and manufacturing sectors to decline. The joint model, however, predicts smaller changes in sectoral output and employment shares in the latter stages of the calibration than occurred in reality. It is argued that this is shortcoming of the model is related to the fact that the parameters that drive demand-side NBG exert a stronger influence on changes in sectoral output and employment shares in the early stages of the calibration. This shortcoming of the model implies that while the calibrated model is able to do a reasonable job of matching changes in sectoral output and factor shares over the study period as a whole, it is not likely to perform well at predicting future patterns of non-balanced growth in the United States.

As is mentioned in the Introduction, Buera and Kaboski (2009) developed three-sector models that integrates supply-side and demand-side explanations for NBG that is capable of matching the Kuznets facts for the U.S. economy regarding sectoral employment shares, but not regarding sectoral output shares. In particular, Buera and Kaboski (2009) where unable to predict the simultaneous increases the relative price of service sector output (which is necessary to draw labor and capital into the service sector) and in the service sector's output share that occurred in the post-war U.S. economy. The results reported in this essay indicate that Buera and Kaboski

(2009)'s inability to qualitatively match the Kuznets facts for the U.S. in terms of both output and employment shares is not a general failing of supply- and demand-side explanations for NBG to explain the Kuznets facts in the United States, but rather an artifact of their particular modelling assumptions and parameterization.

In addition to producing significant changes in sectoral output and employment shares consistent with the Kuznets facts, the model generates equilibrium dynamics that are broadly consistent with the Kaldor facts as they apply to the U.S. economy over the study period. As is noted in the Introduction, Buera and Kaboski (2009) demonstrate that any model that integrates supply-side and demand-side explanations for NBG will be inconsistent with balanced long-run growth for output. The analysis in this essay demonstrates that despite this result, a relatively simple three-sector model of NBG can generate short-run transition dynamics that approximately match the Kaldor facts over a wide range of different parameterizations. Moreover, the analysis did not reveal any significant trade-off between the magnitude of the changes in sectoral output and employment predicted by the model and its ability to match the Kaldor facts. These results suggest that the desire to develop a model that integrates supply-side and demand-side explanations for NBG and produces a long-run balanced growth path for the economy as described by the Kaldor facts is an issue of theoretical rather than practical interest.

The mechanism for demand-side NBG in this essay assumes that changes in the composition of demand for sectoral output occur as a result of differences in the elasticities of demand for sectoral output with respect to the total production of a unique final good, where the unique final good can be consumed by households or used to produce new capital stock. As is explained in the Introduction, this assumption has the advantages that it allows physical capital to be produced in the model using the output from all three sectors, which is important for the model's ability to match changes in sectoral output shares, and that it allows the parameters of the model to be more easily calibrated to U.S. data. This assumption has the drawback, however, that it implies that the parameters that determine changes the composition of demand for sectoral output also influence the quantity of the final output produced, and, as a result, capital accumulation in the economy. While this assumption blurs the distinction between "preference" and "technology" parameters in the model, it is in keeping with previous models of supply- and demand-side NBG. In many previous models of supply-side NBG, the parameter that determines the elasticity of substitution between output from different sectors is modelled on the technology side of the economy (e.g., Acemoglu and Guerrieri 2008), and thus influences total output and capital accumulation, despite the fact that the elasticity of substitution is ultimately determined by household preferences. Similarly, previous studies of demand-side NBG such Kongsamut, Rebelo, and Xie (2001) and Foellmi and Zweimuller (2005) have placed restrictions on the relationship between preference and technology parameters in their models in order to reconciled demand-side NBG with the Kaldor facts. Future work may be directed towards developing models that reconcile the Kuznets facts and the Kaldor facts within a theoretical framework that maintains the distinction between preference and technology parameters that is traditionally emphasized in macroeconomic models.

Chapter 3

Regulatory Policy Design for Agroecosystem Management on Public Rangelands

This essay is co-authored with Tigran Melkonyan and has been published in the *American Journal of Agricultural Economics* (Volume 95, Issue 3; pages 606-627).

3.1 Introduction

Rangeland is the dominant land type in the United States, comprising 34.2% of total land area (731 million acres), compared to 32.4% forested, 17% agricultural, and 2% urban (Loomis 2002). Over 235 million acres of this rangeland is under the management of the federal government and is used for livestock grazing via contractual arrangements between ranchers and federal land management agencies (FLMAs). Two FLMAs – the Bureau of Land Management (BLM) and the U.S. Forest Service (USFS) – manage the livestock grazing leases on over 98% of these 235 million acres (GAO 2005). Given the amount of rangeland managed by FLMAs through federal grazing leases, the regulation of ranching on public rangelands plays a central role in rangeland management, as well as overall natural resource management, in the United States.

Ranching is, in many respects, a prototypical agroecosystem management problem. It generates both private economic gains for agricultural producers as well as externalities. The latter are caused by the influence of livestock grazing on rangeland vegetation. Grazing stresses native perennial grasses, reducing their ability to compete with native shrubs, non-native annual grasses, and noxious weeds. The consequences of a change in rangeland vegetation away from native perennials include a reduction in the quality of wildlife habitat for game animals and sensitive species, increased frequency and severity of wildfires, and increased soil erosion.[1]

While ranchers have private incentives to maintain ecosystem health (healthy rangeland provides more productive and sustainable forage base for livestock) their private objectives differ from social goals. Reflecting the importance of external costs on public lands, the "multiple use" and "sustainable yield" mandates of the BLM, USFS, and other FLMAs require these government agencies to address the externalities associated with ranching when setting regulation. These two mandates require that FLMAs take into account wildlife, watershed health, and recreation as well as commercial interests such as ranching (multiple use), and that they work to ensure that the resource values on public lands are available at current levels in perpetuity (sustainable yield).[2]

[1]Keith and Lyon (1985), Cory and Martin (1985), Roach, Loomis, and Motroni (1996), and Shonkwiler and Englin (2005) find that livestock grazing and recreation have competing values on public rangelands. Other externalities include the influence of rangeland degradation on soil erosion (Knapp 1996), carbon sequestration (Follett, Kimble, and Lal 2001; Verburg et al. 2004; Brown et al. 2006; Havstad et al. 2007), and wildfire activity (Billings 1990) and its effects on ranch profits (Maher 2007) and wildfire suppression costs (Taylor et al. 2011). One of the most robust findings of this literature is that these external costs increase dramatically with changes in rangeland vegetation away from native perennial grasses.

[2]The USFS adopted the principles of multiple use and sustainable yield with the "Multiple Use, Sustained Yield Act of 1960." The BLM followed suit in 1964 with the "Classification and Multiple Use Act of 1964".

This article analyzes regulatory design for agroecosystem management on public rangeland. We develop a model with two parties: an agent (rancher) that uses the agroecosystem for private economic gain and a principal (FLMA or regulator) that manages the agroecosystem for both economic gains of the rancher and public goods related to ecosystem health. We consider an informational environment where the rancher is better informed than the FLMA about the effectiveness of her/his actions in achieving both her/his private economic objectives and in influencing the public good aspects of ecosystem health, and where there is moral hazard in the implementation of any regulatory scheme because some of the rancher's actions cannot be observed by the FLMA. In addition, high costs of monitoring ranch-level ecological conditions make it infeasible for FLMAs to engage in regular and detailed monitoring of ranch-level ecological conditions on public rangelands.[3] As a result, FLMAs base regulation on imperfect signals of how the ranchers' activities influence ecosystem health.

In addition to these informational constraints, we model institutional constraints faced by FLMAs. It is assumed that the FLMA is constrained by its exogenous budget to fund policy but it can supplement this exogenous budget through taxation. This feature of the model reflects the current practice on public rangelands, where FLMAs have fixed budgets in the short-run but are able to use revenues collected through grazing fees to fund their activities.[4] We also consider ranchers' participation constraints, which require that a rancher's profit from ranching on public rangeland

[3] Monitoring costs on public rangelands are high relative to FLMA budgets. Indeed, the high cost in terms of personnel and other resources to monitor ecosystem health has been cited as a reason why FLMAs often do not perform the regular monitoring activities that are specified in federal grazing leases (GAO 1992).

[4] FLMAs, however, are constrained to spend their grazing-fee revenue on "range improvement" projects and have to give a large portion (roughly 50%) of their grazing-fee revenues to state governments to return to counties as "Payments in Lieu of Taxes" and to the U.S. Treasury (Watts, Shimshack, and LaFrance 2006).

exceed her/his outside option. As a result, the FLMA is constrained in the monetary and non-monetary penalties it can assess.

The modeled informational and institutional environment allows us to define and compare the most prominent regulatory instruments on public rangelands. These instruments are *input mandates*, where the regulator mandates the level of usage of certain inputs, *cost-sharing/taxation*, where the regulator subsidizes or taxes the use of certain inputs, and *performance regulation*, where the regulator compensates or penalizes the rancher based on the value of an observed performance measure.

We begin by analyzing the efficiency of the three regulatory instruments in light of the informational and institutional constraints faced by FLMAs. We characterize conditions under which each of the three instruments improves welfare and dominates the other instruments. When FLMAs are unconstrained in the level of bonus or penalty they can assess and when there is perfect monitoring, the first-best outcome can be achieved through performance regulation. In a more realistic setting, however, the FLMA is constrained in the level of bonus/penalty it can assess and/or monitoring is imperfect. Under these circumstances, both input mandates and cost-sharing/taxation can dominate performance regulation. After considering each regulatory instrument in isolation, we examine relative attractiveness of a joint use of the regulatory instruments.

To our knowledge, this article is the first to compare the merits of these three regulatory instruments in a setting that captures the salient informational and institutional constraints faced by FLMAs on public rangelands. By analyzing these three regulatory instruments in the same model, we provide a platform to compare the optimal mix of regulatory instruments with the existing FLMA regulations for ranching on public rangelands. This allows us to consider how FLMAs' informational and institutional constraints have shaped the existing regulation and evaluate possible explanations for the continued reliance of FLMAs on input mandates, in the form of grazing restrictions, despite their demonstrated inefficiency in reaching a target level of environmental performance (e.g., Zhao 2008).

The results in this article apply specifically to regulation on public rangeland, as opposed to regulation on private rangeland, because regulators have a greater ability to restrict herd size, mandate infrastructure investments, and penalize ranchers for non-compliance when the latter operate on public land. The informational and institutional constraints on efficient management considered in this article, however, apply equally to the regulation of ranching on both public and private rangelands, as well as to many other agroecosystem management problems. In particular, in many agricultural and natural resource extraction activities the (i) externalities associated with the activity cannot be monitored perfectly, (ii) some management efforts are not perfectly observed, (iii) there is uncertainty about the effect of management actions on external costs, and (iv) the public agency tasked with management faces a budget constraint.[5] As such, our findings have implications that extend beyond the regulation of ranching on public rangelands.

3.2 Regulation on Public Rangelands

FLMAs use several regulatory instruments to reduce the negative externalities associated with ranching and to ensure that public rangelands are managed in accordance with their multiple use and sustainable yield mandates. The most prominent of these

[5]For example, (i) the external cost of chemical (herbicide and pesticide) use in agriculture on watershed health cannot be monitored perfectly; (ii) farmer effort to mitigate the external cost of chemical use cannot be observed perfectly by a regulatory agency; (iii) the external cost of chemical use cannot be inferred with certainty from usage rates due to a complex, farm-specific relationship between chemical use and watershed health; and (iv) budget considerations likely constrain a regulatory agency's ability to promote efficient chemical use through performance bonuses or cost-sharing of alternative technologies with lower social costs.

regulatory instruments are restrictions on the number of livestock ranchers can graze on their public land allotments.[6] Livestock grazing restrictions are aimed at ensuring the long-term ecological health of public rangeland allotments by limiting the rancher's ability to inflict ecological harm through over-grazing. In addition, as we explain below, FLMAs can use the possibility of expanded or reduced grazing privileges to motivate the rancher to manage their rangeland in accordance with FLMAs' ecological health objectives.

In principle, grazing restrictions specify the maximum number of livestock a rancher can run on her/his public land allotment. In practice, however, ranchers are also required to make "substantial use" of range forage or risk possible loss of grazing privileges. Reduced grazing privileges lower ranchers' potential profits from ranching and can diminish the sale value of their grazing permit and base ranch. The combination of the maximum grazing restrictions and non-use provisions amounts to a de facto mandate that forces most ranchers to choose a number of livestock they graze on their public rangeland allotment from a narrow interval of possible herd sizes.

In addition to facing grazing restrictions, ranchers must pay a per animal, per month grazing fee. An efficient grazing fee would be set equal to the marginal social value of forage that incorporates the marginal forage value for ranchers and the marginal external environmental costs. Grazing fees on public rangelands, however, are set nationally, and are thus inefficient for most ranches because of the heterogeneity of range conditions. Johnson and Watts (1989) find that despite this inefficiency and the existence of non-use provisions, stocking rates on public land allotments are

[6] These are often referred to as Animal Unit Month (AUM) restrictions, where an AUM is the amount of forage needed to sustain one cow and her calf, or one horse, or five sheep or goats for one month.

somewhat responsive to changes in grazing fees.[7] It has also been suggested that under certain degraded rangeland conditions, it may be socially efficient to subsidize grazing above privately optimal levels as a means of noxious weed control and wildfire fuels reduction (Papanastasis 2009).

Besides grazing restrictions and fees, ranchers operating on public rangeland are often obligated to engage in construction or maintenance of "range improvements" as part of the conditions of their lease (USDI BLM 2008). These include enhancing livestock grazing management, improving watershed conditions, and enhancing wildlife habitat. "Range improvements" can be structural, such as water pipes, wells, and fences, or non-structural, such as re-seeding and prescribed burns. While an FLMA and a rancher will often work jointly to achieve desired "range improvements" (FLMAs have budgets for "range improvements" that are funded through grazing fees) these activities add to the rancher's cost of operating on public rangelands (Torell and Doll 1991; Xu, Mittelhammer, and Torell 1994).

FLMAs pursue a strategy of both long- and short-term monitoring of the ecological conditions on public rangeland allotments. Monitoring is performed in order to assess the ranchers' compliance with their contractual obligations on their allotments and with the "Standards of Rangeland Health," which are a series of ecological health goals set forth by the FLMA (USDI BLM 2007).[8] Long-term monitoring is focused on changes in the status of vegetation on an allotment and is generally performed at the time of permit renewal.[9] In contrast, short-term monitoring includes monitoring

---

[7] Watts, Shimshack, and LaFrance (2006) argue that because non-use provisions encourage ranchers to graze their maximum allowed number of livestock, grazing fees represent a fixed cost rather than a variable cost for most ranchers.

[8] The Standards for Rangeland Health that apply to a given allotment are set by local 15-member Resource Advisory Councils that have flexibility to adapt the Standards of Rangeland Health to local conditions and priorities (Swanson 2008, personal interview).

[9] Long-term monitoring is generally performed at the time of permit renewal unless there is

the time and intensity of grazing, the total number of animals on the allotment, pasture, or use area, and utilization, which is an estimate of the amount of forage removed from the land after the grazing season. The dates of use can be used to understand plant stresses and determine the amount of time that forage is allowed to rejuvenate after grazing (Swanson 2006; Swanson 2008: Personal Interview).

If monitoring reveals that current management is degrading rangeland health, or that the rancher is failing to comply with her/his contractual obligations, reductions in the rancher's grazing privileges or mandatory range improvements may be imposed.[10] Both of these consequences of monitoring serve as penalties on ranchers for violations of contractual obligations. Conversely, monitoring can result in the expansion of grazing privileges if current grazing is found to do limited ecological harm. In this way, monitoring and the associated penalties/bonuses provide the rancher with incentives to manage their allotments in accordance with the FLMAs' ecological health objectives.

3.3 Related Literature

Of the three regulatory instruments that we consider in this article, cost-sharing/taxation in the form of grazing fees has received the most attention in the previous literature. This focus on grazing fees can be explained in part by the considerable controversy that federal grazing fees have generated (Hess and Holecheck 1995). Some authors argue that federal grazing fees are set too low relative to the market value of forage

a serious resource concern on the allotment, such as soil erosion or degraded riparian areas, or the rancher is involved in an ongoing range improvement project that involves a comprehensive monitoring program.

[10] While obligatory range improvements may benefit the rancher, they will impose costs on the rancher if the latter would not have otherwise undertaken the range improvement or if the rancher would have undertaken it in a different manner than was mandated by the FLMA.

(Fowler, Torell and Gallacher 1994; LaFrance and Watts 1995), while others maintain that federal grazing fees are set appropriately given the cost of compliance with federal regulations on public rangelands (Torell and Doll 1991; Xu, Mittelhammer, and Torell 1994). Several studies construct theoretical models to characterize the optimal grazing fee in the presence of externalities (McCarl and Brokken 1985; Huffaker, Wilen, and Gardner 1989) and informational asymmetries between the FLMA and the rancher (Watts, Shimshack, and LaFrance 2006). Relative to this literature, we consider the optimal grazing fee taking into account both the externalities associated with ranching and a richer set of informational and institutional constraints faced by FLMAs. In addition, we consider how the optimal grazing fee is influenced by other regulatory instruments in concurrent use on public rangelands.

Regulatory mechanisms other than grazing fees have received substantially less attention in the literature. Torell, Lyon, and Godfrey (1991) consider the relative economic importance of current-period animal performance and future forage production for a yearling stocker operation in eastern Colorado and find that current period animal performance defined by weight gain drives economic stocking-rate decisions. In an analysis of incentive-based mechanisms, Huffaker, Wilen and Gardner (1989) propose the use of a grazing fee in conjunction with "transfer payments" based on observed range conditions as a potential mechanism to induce ranch compliance with the FLMA's ecological health objectives. Our work builds on these studies by considering both input mandates (stocking-rates) and performance regulation (incentive-based mechanisms) in a setting that captures the informational and institutional constraints on public rangelands and allows for a comparison of the efficiency properties of these regulatory instruments, along with cost-sharing/taxation.

There is a large and burgeoning economics literature on regulatory design under asymmetric information. A number of studies examine the relative merits of quantity instruments (input mandates), price instruments (cost-sharing/taxation), and performance regulation. Weitzman (1974) demonstrates how a quantity instrument can dominate price instruments when there is uncertainty and asymmetric information in policy design. The simultaneous use of different regulatory instruments has also been analyzed (Shavell 1984; Innes 1998; Kolstad, Ulen, and Johnson 1990; and Hueth and Melkonyan 2009).

The present article makes three innovations relative to the received literature. First, by constructing a model where closed-form expressions for the social welfare under the three regulatory instruments can be derived, we directly compare the efficiency of the three regulatory instruments and examine how their relative efficiency is influenced by the informational and institutional constraints faced by the regulator. Second, we examine the simultaneous use of multiple regulatory instruments, identify circumstances under which it is most advantageous for the regulator to use multiple instruments, and examine how the efficient use of each individual instrument changes when a mixture of instruments is optimal. Third, we consider optimal regulation in the presence of budget constraints. In doing this, we are contributing to the literature on optimal contracting with a budget-constrained principal operating under imperfect information. The previous literature on optimal contracting under a budget constraint has focused on procurement problems, where the principal designs contracts to overcome adverse selection (Levaggi 2004; Gautier 2004; Anthon et al. 2007). In contrast, the informational environment in our model entails both asymmetric information and moral hazard. We find that the budget constraint causes the regulator to rely on instruments that would otherwise be inefficient.

An important component of the economics literature on regulatory design is the large and growing body of work on optimal environmental regulation in agriculture. By and large, this literature (e.g., Bontems, Turpin, and Rotillon 2005; Feng 2007; and Sheriff 2008) has focused on the case where the principal has the dual goals of limiting environmental externalities and providing income support to agricultural producers, and where the efficiency of regulation is undermined by adverse selection. In a context similar to these studies, Bontems and Bourgeon (2005) investigate the properties of the optimal environmental taxation and enforcement policy assuming that emissions can be observed through costly audits and that private information remains even when an audit is performed. In a related analysis, Bontems and Thomas (2006) present a model of pollution regulation for a risk averse farmer facing production risk from nitrogen leaching. Our framework is similar to their model which incorporates moral hazard and private information on the producer's part about farm-level ecological conditions. In contrast to these studies, we do not consider adverse selection, nor does income support for producers enter the principal's objective (unlike in many other agricultural contexts, income support for ranchers is not an explicit goal of FLMA policy). We also do not consider monitoring of compliance with environmental regulation.[11] Indeed, our model departs from the literature by examining how the combination of the institutional and informational constraints faced by FLMAs influences the relative efficiency of three pervasive and relatively unsophisticated regulatory instruments. In another related work, Anthon, Garcia, and Stenger (2010) analyze environmental regulation in the presence of both asymmetric information and moral hazard, but do not consider a budget constrained regulator/principal.

Finally, on a purely formal level, our model extends the work of Baker (1992), Prendergast (2002), and Hueth and Melkonyan (2009) by examining multiple regulatory instruments under an alternative informational environment and by considering additional institutional constraints. In addition, while there are similarities between our model and Holmstrom and Milgrom (1991), we assume linear separability between actions and, as such, do not examine the multi-tasking issues they consider.

[11] Rather, we consider a setting where a bonus/penalty is assessed to the rancher based on imperfect ex post monitoring of ranch-level ecological conditions, but where compliance with regulation is not considered.

3.4 Model

We consider a strategic interaction between two parties, a regulator (she) and a rancher (he). The rancher utilizes a production process with two inputs, denoted by $e_1 \geq 0$ and $e_2 \geq 0$. In addition to influencing the rancher's private payoff, these inputs affect the health of the ecosystem where the rancher carries out his production activities and, by doing so, cause externalities.

In the absence of regulation, the rancher's private payoff function[12] is given by

$$\pi (e_1, e_2) \equiv F (e_1, e_2) - w_1 e_1 - w_2 e_2, \tag{3.1}$$

where $F (e_1, e_2)$ has the quadratic form

$$F (e_1, e_2) = \left[\theta_1 e_1 - \gamma_1 (e_1)^2\right] + \left[\theta_2 e_2 - \gamma_2 (e_2)^2\right], \tag{3.2}$$

with $\gamma_1, \gamma_2 > 0$, $\theta_1 > w_1 > 0$, and $\theta_2 > w_2 > 0$. The function $F (e_1, e_2)$ represents the revenue from ranching. We assume that e_1 is purchased from a market and it can be observed, and, hence, regulated, by the regulator. We let w_1 denote the market price of e_1. In contrast, e_2 represents the rancher's effort directed toward enhancing production and/or the ecosystem health and it is not observed by the regulator. We let w_2 denote the constant marginal cost of effort e_2. The most important inputs chosen by ranchers and observed and used by regulators on public rangelands are the scale and intensity of livestock grazing. Other observed inputs include certain

[12] Previous research has found that ranchers receive compensation from ranching in the form of "consumptive amenities" related to the "ranching lifestyle" (Torell et al. 2005). This amenity value of ranching, however, is unlikely to vary on the margin with the number of cattle grazed or with the use of any other input. As such, including the amenity value of ranching in our analysis would only influence the rancher's participation constraint. As described below, we normalize the rancher's reservation utility, without any loss of generality, to zero. If the amenity value of ranching is significant, then this normalized reservation utility can be thought to include the expected utility of not ranching minus any consumptive amenities related to the ranching lifestyle.

structural (e.g., water pipes, wells, and fences) and non-structural (e.g., re-seeding and prescribed burns) range improvements. The most important actions performed by ranchers that are often not observed by regulators are ranchers' grazing management techniques, such as rotational grazing and preventing cattle from damaging riparian areas, that reduce the ecological harm from livestock grazing.

In the absence of regulation, the rancher will choose e_1 and e_2 to maximize his private payoff:

$$\max_{e_1,e_2} \pi\left(e_1, e_2\right). \tag{3.3}$$

Under our assumption that $\theta_1 - w_1 > 0$ and $\theta_2 - w_2 > 0$, the rancher's optimal choice of inputs in the absence of regulation is given by

$$e_i^{nr} = \frac{\theta_i - w_i}{2\gamma_i} > 0 \text{ for } i = 1, 2. \tag{3.4}$$

The ex post social value of the public good related to ecosystem health is given by

$$V\left(e_1, e_2, \varepsilon_1, \varepsilon_2\right) = \left(\mu_{11} + \mu_{12}\varepsilon_1\right) e_1 + \left(\mu_{21} + \mu_{22}\varepsilon_2\right) e_2, \tag{3.5}$$

where μ_{11}, μ_{21}, μ_{12} and μ_{22} are parameters of the model and ε_i $(i = 1, 2)$ is a random variable with support $\left[-\varepsilon_i', \varepsilon_i'\right]$, $E\left(\varepsilon_i\right) = 0$, and $Var\left(\varepsilon_i\right) = \sigma_{\varepsilon_i}^2$. Thus, the social value of the public good is a random variable whose distribution is affected by the rancher's choice of inputs.[13] It is assumed that neither the rancher nor the regulator observes the realizations of ε_1 and ε_2. Given our specification, an increase in input e_i $(i = 1, 2)$ leads to an increase in the variance of V.

We have assumed that both $F\left(e_1, e_2\right)$ and $V\left(e_1, e_2, \varepsilon_1, \varepsilon_2\right)$ are separable in e_1 and e_2 to focus on the direct effects of the regulatory instruments on the rancher's

[13] In contrast to the social value of ecosystem health, the rancher's private payoff is assumed to be deterministic. In reality, many random factors (e.g., weather) affecting the former likely impact the latter. One could easily utilize the techniques in Baker (2002) to analyze this more general case. Although this would yield additional insights, the qualitative results reported in the paper would not be affected by such enrichment of the model.

input choices. When the above functional forms are not separable, a regulatory instrument targeted at a specific input (in our model, an input mandate, cost-sharing, or taxation) will also have an indirect effect on the use of the other input. As long as these indirect effects are small compared to the direct effects, our qualitative results remain unchanged. Our choice of a quadratic functional form for $F(e_1, e_2)$ and a linear functional form for $V(e_1, e_2, \varepsilon_1, \varepsilon_2)$ is driven mainly by space considerations.

The ex post social welfare is defined as the sum of the rancher's private payoff, $F(e_1, e_2) - w_1 e_1 - w_2 e_2$, and the value of the public good related to ecosystem health, $V(e_1, e_2, \varepsilon_1, \varepsilon_2)$. Since any payment related to a regulatory instrument is a transfer between the regulator and the rancher, it does not influence the social welfare.[14]

Using e_1^{nr} and e_2^{nr} from (3.4), we obtain the expected social welfare in the unregulated case:

$$\begin{aligned} U^{nr} &\equiv E\left[\pi\left(e_1^{nr}, e_2^{nr}\right) + V\left(e_1^{nr}, e_2^{nr}, \varepsilon_1, \varepsilon_2\right)\right] \\ &= \sum_{i=1}^{2} \frac{\left(\theta_i - w_i\right)^2}{4\gamma_i} + \sum_{i=1}^{2} \frac{\mu_{i1}\left(\theta_i - w_i\right)}{2\gamma_i}. \end{aligned} \tag{3.6}$$

The expected social welfare without regulation will be used as a benchmark against which the efficiency of each regulatory mechanism will be compared.

Another important benchmark for assessing the efficacy of the regulatory mechanisms is the full-information social optimum, also called first-best, for which the rancher's choice of e_i is conditioned on the realization of the random variable ε_i to maximize the ex post social welfare:

$$\max_{e_1, e_2} \left[\pi\left(e_1, e_2\right) + V\left(e_1, e_2, \varepsilon_1, \varepsilon_2\right)\right]. \tag{3.7}$$

The full-information socially optimal input choices are given by

$$e_i^*\left(\varepsilon_i\right) = \begin{cases} 0, & \text{if } \theta_i + \mu_{i1} + \mu_{i2}\varepsilon_i \leq w_i \\ \frac{\theta_i - w_i + \mu_{i1} + \mu_{i2}\varepsilon_i}{2\gamma_i}, & \text{otherwise} \end{cases} \quad \text{for } i = 1, 2. \tag{3.8}$$

[14] We do not model the potential welfare costs associated with raising revenue (e.g., due to distortionary taxation) to fund regulation on public rangeland.

We assume that $\theta_i + \mu_{i1} + \mu_{i2}\varepsilon_i > w_i$ $(i = 1, 2)$ for all $\varepsilon_i \in [-\varepsilon_i', \varepsilon_i']$, so that the full-information socially optimal choices of e_1 and e_2 are always strictly positive. When parameter μ_{i2} is strictly positive, the first-best calls for the rancher to increase the use of input e_i in response to increases in ε_i.

The expected social welfare evaluated at the full-information social optimum is given by

$$U^* = \sum_{i=1}^{2} \gamma_i E\left\{[e_i^*(\varepsilon_i)]^2\right\} = U^{nr} + \sum_{i=1}^{2} \frac{\mu_{i1}^2 + \mu_{i2}^2 \sigma_{\varepsilon_i}^2}{4\gamma_i} \quad (3.9)$$
$$= U^{nr} + \sum_{i=1}^{2} \frac{[E(V_{e_i})]^2 + Var(V_{e_i})}{4\gamma_i},$$

where $V_{e_i} \equiv \frac{\partial V(e_1, e_2, \varepsilon_1, \varepsilon_2)}{\partial e_i}$ is the marginal social value of e_i, $i = 1, 2$ (or, alternatively, V_{e_i} is the marginal product of e_i on ecosystem health).

It follows immediately from (3.6) and (3.9) that when the expected marginal social value of e_i is non-zero – i.e., $E(V_{e_i}) = \mu_{i1} \neq 0$, $i = 1, 2$ – the rancher's private optimal choices of e_1 and e_2 in the absence of regulation are socially inefficient, and that the extent of this inefficiency is positively affected by $|E(V_{e_1})|$ and $|E(V_{e_2})|$. Expression (3.9) also reveals that the inefficiency of the rancher's private optimum in the absence of regulation is independent of whether input e_i is detrimental ($E(V_{e_i}) < 0$) or beneficial ($E(V_{e_i}) > 0$) to ecosystem health. In addition, the degree of the inefficiency of the rancher's privately optimal choice e_i^{nr}, $i = 1, 2$, is increasing in the variance of the marginal social value of the rancher's effort, $Var(V_{e_i})$, $i = 1, 2$, which measures the heterogeneity of the effect of e_i on ecosystem health between ranchers.

An alternative benchmark for assessing the efficiency of the regulatory mechanisms corresponds to the socially optimal policy that, in contrast to the previous benchmark, takes into account only the information observed by at least one of the two parties. As discussed in the next section, the regulator can observe an imperfect signal, termed performance measure, of the impact of the rancher's choices of e_1 and e_2 on ecosystem health. The rancher receives private signals η_1 and η_2, correlated with ε_1 and ε_2, that determine how his choices of e_1 and e_2, respectively, influence the performance measure. We call this alternative benchmark – where the input choices can be conditioned on the observations of η_1 and η_2 to maximize the expected social welfare – partial-information social optimum. The name reflects the supposition that the choices can be conditioned on η_1 and η_2 but not on ε_1 and ε_2.

The partial-information socially optimal choices of inputs e_1 and e_2 are given by

$$e_i^{**}(\eta_i) = \frac{\theta_i - w_i + \mu_{i1} + \mu_{i2} E[\varepsilon_i | \eta_i]}{2\gamma_i} > 0 \text{ for } i = 1, 2. \tag{3.10}$$

The expected social welfare under the partial-information social optimum is equal to

$$\begin{aligned} U^{**} &= U^{nr} + \sum_{i=1}^{2} \frac{[E(V_{e_i})]^2 + Var(V_{e_i}) - E[Var(V_{e_i}|\eta_i)]}{4\gamma_i} \\ &= U^* - \sum_{i=1}^{2} \frac{E[Var(V_{e_i}|\eta_i)]}{4\gamma_i}. \end{aligned} \tag{3.11}$$

Relative to the expected social welfare (3.9) evaluated at the first-best, the expected social welfare is reduced proportionately to $E[Var(V_{e_i}|\eta_i)]$. The latter is inversely related to the value of the information in η_i about ε_i and, hence, about V_{e_i}. This benchmark illustrates that the first-best social optimum cannot be achieved so long as η_i provides a less than perfect signal of ε_i. Note also that like the first-best, the partial-information social optimum weakly dominates the regulatory instruments analyzed in this article.

3.5 Regulatory Instruments

Recognizing that the rancher's privately optimal choices of e_1 and e_2 do not fully take into account the effect of his activities on ecosystem health, the regulator contemplates introducing a regulatory mechanism. The regulator has three regulatory instruments at her disposal to improve upon the unregulated outcome; (1) *input mandate*, where the regulator fixes the rancher's use of the observable input, (2) *cost-sharing/taxation*, where the regulator subsidizes/taxes the rancher's use of the observable input, and (3) *performance regulation*, where the regulator pays/taxes the rancher based on the value of an observable performance measure related to the rancher's use of inputs e_1 and e_2.[15]

The regulator, however, does not have full flexibility when choosing a regulatory mechanism. First, she faces informational constraints. It is assumed that the rancher's private payoff, $F(e_1, e_2) - w_1 e_1 - w_2 e_2$, and use of input e_2 are not observable by the regulator and, as a result, cannot be a part of a regulatory mechanism. The regulator, however, has full knowledge of how e_1 and e_2 influence the rancher's private payoff and uses this knowledge when determining the optimal regulatory mechanism. Neither the regulator nor the rancher observe the value of the public good related to ecosystem health, $V(e_1, e_2, \varepsilon_1, \varepsilon_2)$, or the realization of the random variables ε_1 and ε_2, which determine the influence of the rancher's input choices on the public good.[16] The regulator does observe, however, a subset of relevant ecosystem health outcomes over which the rancher has influence. This subset of ecosystem health outcomes – henceforth termed *performance measure* – provides an imperfect signal of the impact of the rancher's choices of e_1 and e_2 on ecosystem health. This performance measure is in general different from the actual and expected values of V.

[15] The performance regulation examined in this article has many similarities with payments for ecosystem services (PES) systems. Apart from the involuntary nature of participation in the performance regulation conditional on the rancher's choice to operate on public land, the performance regulation satisfies the other parts of the definition of a PES system in Engel, Pagiola, and Wunder (2008). Specifically, the performance regulation is a transaction where (i) a well-defined environmental service (the performance measure) is (ii) being "bought" by a service buyer (the regulator) from (iii) a service provider (the rancher) (iv) if and only if the service provider secures service provision (payment dependent on realization of the performance measure $P(\cdot)$).

[16] Our findings would be unchanged if, alternatively, we assumed that $V(e_1, e_2, \varepsilon_1, \varepsilon_2)$, ε_1, and ε_2 were observable but non-contractible.

The rancher receives private signals, η_1 and η_2, that determine how his choices of e_1 and e_2, respectively, influence the performance measure. Each private signal η_i, $i = 1, 2$, is correlated with the random variable ε_i, so that the marginal products of e_i, $i = 1, 2$, on the performance measure and $V(e_1, e_2, \varepsilon_1, \varepsilon_2)$ are correlated. As with the rancher's private payoff, the regulator has knowledge of how e_1 and e_2 influence $V(e_1, e_2, \varepsilon_1, \varepsilon_2)$, contingent on ε_1 and ε_2, and the performance measure, contingent on η_1 and η_2. While η_1, η_2, ε_1, and ε_2 are not observed by the regulator, their distributions are known to the regulator. Thus, a regulatory instrument can only be conditioned on the realizations of e_1 and the performance measure.

Second, the regulator's choice of a regulatory mechanism is constrained by her available budget $B \geq 0$; the regulator cannot implement a regulatory mechanism for which the expected budget outlay exceeds B.[17] In contrast, the ex post payment to the rancher may exceed B. The assumption that the budget cannot be exceeded ex ante is reasonable in circumstances where the regulator is proposing a regulatory mechanism to a large number of ranchers, so that events where the regulator's ex post payment to a rancher exceeds B are balanced by events where it is smaller than B.

Third, the regulator's choice of a regulatory mechanism is constrained by the rancher's participation constraint. The participation constraint requires that the rancher's ex ante expected profits from ranching (i.e., the expected profits before η_1 and η_2 are learned by the rancher) exceed his ex ante outside option (which we

[17] An alternative approach to incorporating budgetary considerations into our model is offered by modelling the shadow cost of public funds as in Laffont and Tirole (1993). There is an important difference, however, between the approach of Laffont and Tirole (1993) that focuses on a (exogenously given) marginal cost of public funds and the explicit budget constraint approach taken in the present article. In particular, the shadow cost of public funds in Laffont and Tirole (1993) is constant while in our model the shadow cost associated with the regulator's budget constraint varies with the regulator's budget. As such, the level of budgetary outlays has different effects in the two approaches.

normalize to zero).[18] Under the ex ante participation constraint, the rancher's ex post profits for some realizations of η_1 and η_2 may be lower than his ex ante reservation utility. Our choice to model the ex ante participation constraint rather than the ex post participation constraint is sensible when, for example, (i) the rancher observes η_1 and η_2 only after he enters into an agreement with the regulator to operate on public rangeland, (ii) the rancher commits certain fixed costs in order to operate on public rangeland, and (iii) the rancher's profit net of these fixed costs exceeds his reservation utility for all realizations of η_1 and η_2 when the ex ante participation constraint is satisfied. This last condition guarantees that the rancher will remain in operation for all realizations of η_1 and η_2 whenever his ex ante participation constraint is satisfied.

In what follows, it is assumed that the monitoring costs do not differ between the regulatory instruments. In reality, while the monitoring costs may differ between the instruments, they are not likely to vary significantly with the level of regulation pursued by the FLMA. That is, it is likely that, for a given ranch, the cost of monitoring compliance with an input mandate (an AUM restriction) does not vary with the level of the input mandate. We also expect the same relationship to hold for the level of the cost-share/tax and the piece rate of the performance regulation. In other words, the monitoring costs of different regulatory instruments are akin to fixed costs. If monitoring costs are indeed similar to fixed cost, then only FLMA's exogenous budget to fund regulation would be affected by differences in monitoring costs across the regulatory instruments. Moreover, while differences in fixed monitoring costs between the regulatory instruments may be substantial, they will not affect our

[18] Note that the rancher must participate in the performance regulation if he chooses to graze his animals on public land. As such, the rancher will graze his animals on public land, and, as a result, participate in the performance regulation, if his ex ante expected profit from doing so is greater than his outside option. If instead participation in the performance regulation on public lands were voluntary, the rancher would participate if his ex ante expected profit was greater than his ex ante expected profit in the absence of the performance regulation.

results on how different aspects of ranching on public rangeland influence the relative efficiency of the three regulatory instruments considered in this article (when used either individually or in combinations). For this reason, it is further assumed that the fixed monitoring costs of all three instruments are the same.[19]

The analysis in this article focuses on three regulatory instruments that are currently used on public rangelands. For this reason, we do not consider alternative regulatory schemes that could improve welfare over the regulatory instruments used in practice. For example, we do not consider a message game between the rancher and the regulator that would condition the regulatory instruments on the rancher's report of his private information η_1 and η_2 together with the choice of e_1 and the performance measure. We leave an analysis of such mechanisms to future work.

The timing in our model is as follows. First, the regulator announces the regulatory mechanism, and the rancher learns the level of each regulatory instrument. Subsequently, the rancher learns the realization of uncertainty concerning the impact of his choices of e_1 and e_2 on the performance measure and selects e_1 and e_2. Then, the realization of the performance measure is observed by both parties and the regulator also learns the rancher's choice of e_1. Finally, the payments (if any) are made between the regulator and the rancher.

[19] There are significant practical difficulties with esimating the FLMAs' monitoring costs associated with the regulatory instruments considered in this paper. The estimates are difficult to obtain for two reasons. First, much of the monitoring performed by FLMAs would take place even without grazing on public rangelands. For example, monitoring activities such as vegetation surveys occur as part of general rangeland management, and should not be considered as a cost of monitoring ranching on public rangelands (GAO 2005). Second, the information on the monitoring costs of various regulatory instruments is not readily available in published FLMA documents. It may only be possible to glean such information through a thorough audit of FLMA monitoring procedures.

3.5.1 Input Mandate

Input regulation in the form of grazing restrictions is the most prominent form of regulation on public rangelands. As was discussed above, non-use provisions in federal grazing leases imply that grazing restrictions are de facto mandates for most ranchers operating on public rangeland. For this reason, we consider the scenario where the regulator mandates the rancher's choice of the observable input, denoted by e_1'. Under this regulatory instrument, the rancher has to pay a relatively high penalty if his use of e_1 differs from e_1' so that the rancher never finds it advantageous to violate the mandate. It is assumed that it is costless for the regulator to enforce the input mandate.

Because the rancher's private payoff function, $\pi(e_1, e_2)$, is separable in e_1 and e_2, mandating e_1' does not influence the rancher's choice of e_2. As such, provided the rancher uses the public rangeland he will set e_2 equal to the unregulated level e_2^{nr} given by (3.4). Hence, the regulator's optimization problem under the input mandate is given by

$$\max_{e_1',\alpha} E\left[\pi\left(e_1', e_2^{nr}\right) + V\left(e_1', e_2^{nr}, \varepsilon_1, \varepsilon_2\right)\right] \tag{3.12}$$

subject to the regulator's budget constraint

$$B \geq \alpha, \tag{3.13}$$

and the rancher's participation constraint

$$\pi\left(e_1', e_2^{nr}\right) + \alpha \geq 0, \tag{3.14}$$

where α is a lump-sum transfer between the rancher and the regulator.[20] The regulator will provide a lump-sum transfer only when the optimal input mandate is so

[20] Any action by the regulator that benefits the rancher and which is not tied to the rancher's choice of inputs (e.g., certain types of range improvements by the regulator) can be thought of as a lump-sum transfer. In this sense, allowing for lump-sum transfers is quite realistic. There is also a purely technical reason for introducing lump-sum transfers. By allowing for a lump-sum transfer, which

large as to make ranching unprofitable. This can occur only if the expected marginal product of e_1 on ecosystem health is positive $(E(V_{e_1}) > 0)$.

Combining the constraints (3.13) and (3.14), we can re-write this optimization problem as

$$\max_{e_1'} E\left[\pi\left(e_1', e_2^{nr}\right) + V\left(e_1', e_2^{nr}, \varepsilon_1, \varepsilon_2\right)\right] \tag{3.15}$$

$$\text{subject to } \pi\left(e_1', e_2^{nr}\right) + B \geq 0.$$

First, consider the case where the regulator's optimization problem (3.15) is unconstrained. The optimal input mandate in this case is given by

$$e_1^m = \frac{\theta_1 + \mu_{11} - w_1}{2\gamma_1} = E\left[e_1^*\left(\varepsilon_1\right)\right], \tag{3.16}$$

where the last expression is the expectation of the first-best level of input e_1 from (3.8). When the constraint binds, the regulator's optimal choice of e_1' is implicitly given by the binding constraint: $\pi\left(e_1', e_2^{nr}\right) + B = 0$. Solving this equation for e_1' we obtain the optimal input mandate under the binding constraint:

$$e_1^m(B) = e_1^{nr} + \sqrt{\frac{\pi\left(e_1^{nr}, e_2^{nr}\right) + B}{\gamma_1}}. \tag{3.17}$$

By concavity of the objective function in (3.15) and the structure of its constraint, the regulator's constraint in (3.15) is binding at the optimum if and only if $\mu_{11} > 0$ and $e_1^m \geq e_1^m(B)$. Using the expressions for e_1^m and $e_1^m(B)$, we obtain that $e_1^m \geq e_1^m(B)$ if and only if

$$B \leq B^m \equiv \frac{\mu_{11}^2}{4\gamma_1} - \pi\left(e_1^{nr}, e_2^{nr}\right). \tag{3.18}$$

Thus, the optimal input mandate is given by $e_1^m(B)$ if $\mu_{11} > 0$ and $B \leq B^m$ and by e_1^m, otherwise.

---

enters linearly in both the regulator's budget constraint and the rancher's participation constraint, we can immediately reduce the number of constraints by one. This considerably simplifies our analysis. If we ruled out the lump-sum transfer, the results would not change substantially, but their presentation would considerably lengthen an already long paper.

The expected social welfare under the optimal input mandate is given by

$$U^m = \begin{cases} U^{nr} + \frac{\left[E\left(V_{e_1}\right)\right]^2}{4\gamma_1} - \gamma_1 \left\{E\left[e_1^*\left(\varepsilon_1\right)\right] - e_1^m\left(B\right)\right\}^2, \text{ if } \mu_{11} > 0 \text{ and } B \leq B^m \\ U^{nr} + \frac{\left[E\left(V_{e_1}\right)\right]^2}{4\gamma_1}, \text{ otherwise} \end{cases} . \tag{3.19}$$

It follows immediately from this expression that the input mandate strictly improves upon the expected social welfare in the non-regulated case when the expected marginal product of the observable input e_1 on ecosystem health is non-zero, $E\left(V_{e_1}\right) \neq 0$. Because the regulator knows how e_1 influences both F and $E\left(V\right)$ (though ex post she cannot directly observe either F or V), she can set e_1 at the expected social optimum ($e_1^m = E\left[e_1^*\left(\varepsilon_1\right)\right]$) when the budget constraint is not binding. The input mandate, however, cannot restore the first-best outcome unless the constraint in (3.15) is non-binding, the effect of e_1 on ecosystem health is the same for all ranchers ($Var\left(V_{e_1}\right) = 0$), so that the one-size-fits-all input mandate is appropriate, and the unobservable input, e_2, does not influence ecosystem health.

When the budget constraint is binding, the loss in the expected social welfare relative to the non-binding constraint case is determined, among other factors, by the difference in the expected socially optimal usage of the observable input, $E\left[e_1^*\left(\varepsilon_1\right)\right]$, and the level dictated by the regulator's budget constraint, $e_1^m\left(B\right)$. This loss is decreasing in the regulator's budget B (since $e_1^m\left(B\right)$ is increasing in B while $E\left[e_1^*\left(\varepsilon_1\right)\right]$ is independent of B) and increasing in the expected marginal social value of the rancher's effort $E\left(V_{e_1}\right) = \mu_{11} > 0$ (since $E\left[e_1^*\left(\varepsilon_1\right)\right]$ is increasing in μ_{11} while $e_1^m\left(B\right)$ is independent of μ_{11}).

3.5.2 Cost-Sharing/Taxation

We now turn to examining the scenario where the regulator either subsidizes (cost-sharing) or taxes the rancher's use of the observable input e_1. In both cases, the cost of input e_1 incurred by the rancher is $(1-s)w_1e_1$, while the cost borne by the regulator is sw_1e_1. Thus, $s > 0$ corresponds to cost-sharing, while $s < 0$ represents taxation. As discussed above, taxation of forage use through grazing fees is an important element of FLMA policy on public rangelands. In addition to taxation, we consider cost-sharing because it has been suggested that for certain degraded conditions, it may be socially efficient to subsidize grazing above privately optimal levels as a means of noxious weed control and wildfire fuel reduction.

The rancher's choice of e_1 under cost-sharing/taxation is given by

$$e_1^c(s) = \begin{cases} 0, & \text{if } \theta_1 \leq (1-s)w_1 \\ \frac{\theta_1-(1-s)w_1}{2\gamma_1}, & \text{otherwise} \end{cases}. \tag{3.20}$$

As in the preceding section, the rancher's choice of e_2 under cost-sharing is equal to the unregulated level given by (3.4).

Given the rancher's choice of e_1^c as a function of s, the regulator's optimization problem can be written as

$$\max_{s,\alpha} E\left[\pi\left(e_1^c(s), e_2^{nr}\right) + V\left(e_1^c(s), e_2^{nr}, \varepsilon_1, \varepsilon_2\right)\right] \tag{3.21}$$

subject to the regulator's budget constraint

$$B \geq \alpha + sw_1e_1^c(s), \tag{3.22}$$

and the rancher's participation constraint

$$F\left(e_1^c(s), e_2^{nr}\right) - (1-s)\,w_1e_1^c(s) - w_2e_2^{nr} + \alpha \geq 0, \tag{3.23}$$

where α is a lump-sum transfer between the rancher and the regulator.

Combining the constraints (3.22) and (3.23), we can re-write this optimization problem as

$$\max_{s} E\left[\pi\left(e_1^c(s), e_2^{nr}\right) + V\left(e_1^c(s), e_2^{nr}, \varepsilon_1, \varepsilon_2\right)\right] \tag{3.24}$$

$$\text{subject to } \pi\left(e_1^c(s), e_2^{nr}\right) + B \geq 0.$$

Since maximization over s in (3.24) is equivalent to maximization over e_1, the optimization problem (3.24) is equivalent to choosing an optimal input mandate in (3.15). Hence, the elicited choices of the inputs and the resulting expected social welfare under the optimal cost-sharing/taxation mechanism coincide with those under the optimal input mandate. Specifically, the optimal cost-sharing variable s and the resulting choice of input e_1 by the rancher are given, respectively, by

$$s^c(B) = \frac{2\sqrt{\gamma_1 \left[\pi\left(e_1^{nr}, e_2^{nr}\right) + B\right]}}{w_1} \text{ and} \tag{3.25}$$

$$e_1^c(B) = e_1^m(B) = e_1^{nr} + \sqrt{\frac{\pi\left(e_1^{nr}, e_2^{nr}\right) + B}{\gamma_1}}, \text{ if } \mu_{11} > 0 \text{ and } B \leq B^m,$$

and

$$s^c = E\left(V_{e_1}\right)/w_1 = \mu_{11}/w_1 \text{ and} \tag{3.26}$$

$$e_1^c = e_1^m = E\left[e_1^*(\varepsilon_1)\right] = \frac{\theta_1 - w_1 + \mu_{11}}{2\gamma_1}, \text{ otherwise.}$$

Thus, when the constraint in (3.24) is binding, the optimal cost-sharing is positively affected by the reciprocal of the price of input e_1, the reverse of the rancher's profits under the unregulated outcome, and the regulator's budget B. Under a non-binding constraint, the regulator sets cost-sharing variable s equal to the ratio of the expected marginal externality associated with e_1 and its market price. The expected social welfare evaluated at the optimal cost-sharing/taxation arrangement is equal to

$$U^c = U^m = \begin{cases} U^{nr} + \frac{\mu_{11}^2}{4\gamma_1} - \gamma_1 \left\{E\left[e_1^*(\varepsilon_1)\right] - e_1^c(B)\right\}^2, & \text{if } \mu_{11} > 0 \text{ and } B \leq B^m \\ U^{nr} + \frac{\mu_{11}^2}{4\gamma_1}, & \text{otherwise} \end{cases}. \tag{3.27}$$

As a result of this equivalence between the two regulatory instruments, the comparative statics results under the optimal cost-sharing/taxation coincide with those under the optimal input mandate. As we discuss below, the equivalence between the optimal input mandate and optimal cost-sharing/taxation ceases to hold when the two regulatory instruments are used in combination with performance regulation.

3.5.3 Performance Regulation

Suppose that both the rancher and the regulator observe an imperfect measure of how the rancher's choices of e_1 and e_2 influence ecological health:

$$P(e_1, e_2, \eta_1, \eta_2) = (\phi_{11} + \phi_{12}\eta_1) e_1 + (\phi_{21} + \phi_{22}\eta_2) e_2, \tag{3.28}$$

where η_i $(i = 1, 2)$ is a random variable distributed over the interval $[-\eta_i', \eta_i']$ with $E(\eta_i) = 0$ and $Var(\eta_i) = \sigma_{\eta_i}^2$. The rancher observes the realizations of random variables η_1 and η_2 prior to choosing e_1 and e_2. In contrast, the regulator does not learn the realizations of these random variables. It is further assumed that $Corr(\varepsilon_i, \eta_i) > 0$ for $i = 1, 2$. Note that, similarly to the effect of input e_i on the social value of the public good, the variance of P is increasing in e_1 and e_2. It is assumed that the performance measure P is verifiable so that it can be a part of a regulatory mechanism. We consider a linear incentive contract of the form $\alpha + \beta P$, where α is a lump-sum transfer between the regulator and the firm (the base payment) and β is the piece rate per unit of the performance measure P (the power of the incentive contract).[21] We call this regulatory instrument *performance regulation*. We focus on the case where $\beta \geq 0$. Because there are no restrictions on the sign of $P(\cdot)$ and the units of the performance measure are chosen by the regulator, this assumption does not reduce the generality of our results.

Before turning to the formal analysis note that performance regulation has a key advantage over the other two instruments in that it provides the rancher with incentives to respond to private information about η_1 and η_2.[22] When random variables η_1 and η_2 are correlated with random variables ε_1 and ε_2, respectively, performance

[21] The use of a linear incentive contract has the following justification. First, for a principal-agent framework very similar to the one in the present paper, Edmans and Gabaix (2010) demonstrate that the optimal contract is linear. Second, the use of a linear incentive contract considerably simplifies the solution, comparative statics results, and their interpretations.

[22] It is reasonable to expect that in many instances the regulator will be better informed than

regulation gives the rancher the incentives and flexibility to adjust his choice of inputs based on his private knowledge of ranch-level ecological conditions and the effect of the inputs on ecosystem health. As we formally show below, the benefit of giving the rancher this flexibility is increasing in the accuracy of the performance measure (as measured by its distortion $Corr(\varepsilon_i, \eta_i)$, $i = 1, 2$) and the value of the rancher's private information $Var(V_{e_1})$ and $Var(V_{e_2})$ and decreasing in the noisiness of the performance measure $Var(P_{e_i})$, $(i = 1, 2)$.[23] Note that performance regulation influences the rancher's choices of both the observed input, e_1, and the unobserved input, e_2. This way, performance regulation in the model reflects the application of performance-based approaches on public rangelands, which are as much about how grazing is conducted (e.g., rotational grazing, keeping cattle out of riparian areas) as they are about influencing AUM decisions.

The rancher's optimization problem under performance regulation is given by

$$\max_{e_1,e_2} \left[\pi(e_1, e_2) + \alpha + \beta P(e_1, e_2, \eta_1, \eta_2)\right]. \tag{3.29}$$

the rancher about certain aspects of the underlying ecological processes on the allotment. While this may be true, the regulator has no incentive to keep this information from the rancher as both parties benefit from improved rangeland health. Indeed, information campaigns aimed at improving rancher understanding of rangeland ecology are an important component of FLMA policies. These campaigns often focus on invasive grasses, noxious weeds, soil erosion, and best practices in land management. In contrast, the rancher has no incentive to share his private information with the regulator.

[23]In reality, the rancher may also have private information about the opportunity costs of his unobserved effort. The implicit assumption in our analysis, therefore, is that the variation in the opportunity costs of unobserved effort is of secondary importance relative to the variation in how the rancher's input choices impact ecosystem health. Our justification of this assumption relies on the observation that the rangeland ecological conditions vary significantly while rancher production technology is relatively homogeneous across the western United States.

The rancher's optimal choice of inputs is given by

$$e_i^p(\beta,\eta_i) = \begin{cases} 0, & \text{if } \theta_i + \beta(\phi_{i1} + \phi_{i2}\eta_i) \leq w_i \\ \frac{\theta_i - w_i + \beta(\phi_{i1}+\phi_{i2}\eta_i)}{2\gamma_i}, & \text{otherwise} \end{cases} \quad \text{for } i = 1, 2. \quad (3.30)$$

Given the rancher's choices $e_1^p(\beta,\eta_1)$ and $e_2^p(\beta,\eta_2)$, the regulator's optimization problem can be written as

$$\max_{\alpha,\beta} E\left[\pi\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2)\right) + V\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2), \varepsilon_1, \varepsilon_2\right)\right] \quad (3.31)$$

subject to the regulator's budget constraint

$$B \geq \alpha + \beta E\left[P\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2)\right)\right] \quad (3.32)$$

and the rancher's participation constraint

$$E\left[\pi\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2)\right) + \beta P\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2)\right)\right] + \alpha \geq 0. \quad (3.33)$$

Combining these constraints, we can re-write this optimization problem as

$$\max_{\beta} E\left[\pi\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2)\right) + V\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2), \varepsilon_1, \varepsilon_2\right)\right] \quad (3.34)$$

$$\text{subject to } E\left[\pi\left(e_1^p(\beta,\eta_1), e_2^p(\beta,\eta_2)\right)\right] + B \geq 0.$$

In what follows, we focus on the scenarios where the regulator's optimal choice of β is such that the rancher uses positive quantities of e_1 and e_2 for all realizations of η_1 and η_2, i.e., $e_1^p(\beta,\eta_1) > 0$ for all $\eta_1 \in [-\eta_1', \eta_1']$ and $e_2^p(\beta,\eta_2) > 0$ for all $\eta_2 \in [-\eta_2', \eta_2']$.[24]

The algorithm for identifying the optimal solution to the optimization problem (3.34) is similar to that in the two preceding sections. Let β^p denote the solution to the unconstrained problem (3.34):

$$\beta^p = \frac{\sum_{i=1}^{2} \frac{\phi_{i1}\mu_{i1} + \phi_{i2}\mu_{i2}E(\varepsilon_i\eta_i)}{4\gamma_i}}{\sum_{i=1}^{2} \frac{\phi_{i1}^2 + \phi_{i2}^2\sigma_{\eta_i}^2}{4\gamma_i}} = \frac{\sum_{i=1}^{2} \frac{E(P_{e_i})E(V_{e_i}) + \sqrt{Var(V_{e_i})}\sqrt{Var(P_{e_i})}Corr(\varepsilon_i,\eta_i)}{4\gamma_i}}{\sum_{i=1}^{2} \frac{\left[E(P_{e_i})\right]^2 + Var(P_{e_i})}{4\gamma_i}}. \quad (3.35)$$

[24] From (3.30), condition $\beta < \min\left[\frac{-(\theta_1 - w_1)}{(\phi_{11} - \phi_{12}\eta_1')}, \frac{-(\theta_2 - w_2)}{(\phi_{21} - \phi_{22}\eta_2')}\right]$ is sufficient for $e_i^P(\beta,\eta_i) > 0$ $(i = 1, 2)$ for all $\eta_i \in [-\eta_i', \eta_i']$.

The solution to the binding constraint in (3.34) is denoted by $\beta^p(B)$ and is given by

$$\beta^p(B) = \sqrt{\frac{\sum_{i=1}^{2} \left[(\theta_i - w_i)^2/4\gamma_i\right] + B}{\sum_{i=1}^{2} \left[\left(\phi_{i1}^2 + \phi_{i2}^2\sigma_{\eta_i}^2\right)/4\gamma_i\right]}} = \sqrt{\frac{\pi\left(e_1^{nr}, e_2^{nr}\right) + B}{\sum_{i=1}^{2} \frac{\left[E\left(P_{e_i}\right)\right]^2 + Var\left(P_{e_i}\right)}{4\gamma_i}}}. \tag{3.36}$$

By concavity of the objective function in (3.34) and the form of the constraint in (3.34), the regulator's constraint is binding at the optimum if and only if $\beta^p \geq \beta^p(B)$. Using the expression for β^p and $\beta^p(B)$, we obtain that $\beta^p \geq \beta^p(B)$ if and only if

$$B \leq B^p \equiv \left\{\sum_{i=1}^{2} \frac{\left[E\left(P_{e_i}\right)\right]^2 + Var\left(P_{e_i}\right)}{4\gamma_i}\right\}\left(\beta^p\right)^2 - \pi\left(e_1^{nr}, e_2^{nr}\right). \tag{3.37}$$

The optimal performance bonus and the resulting choice of input e_i $(i = 1, 2)$ are thus given, respectively, by

$$\beta^p(B) \text{ and } e_i^p(\eta_i, B) = \frac{\theta_i - w_i + \beta^p(B)\left(\phi_{i1} + \phi_{i2}\eta_i\right)}{2\gamma_i}, \text{ if } B \leq B^p \tag{3.38}$$

and

$$\beta^p \text{ and } e_i^p(\eta_i) = \frac{\theta_i - w_i + \beta^p\left(\phi_{i1} + \phi_{i2}\eta_i\right)}{2\gamma_i}, \text{ otherwise.} \tag{3.39}$$

A number of standard results follow from the expression for β^p.[25] First, the power of the incentive scheme is determined by the expected marginal social value of the rancher's effort $|E(V_{e_i})|$ $(i = 1, 2)$, irrespective of whether input e_i is detrimental $(E(V_{e_i}) < 0)$ or beneficial $(E(V_{e_i}) > 0)$ to ecosystem health. Second, note that $Corr(P_{e_i}, V_{e_i}) = Corr(\varepsilon_i, \eta_i)$ captures the level of distortion in the performance measure. It follows from the expression for β^p that the less distorted the performance measure (the larger $Corr(\varepsilon_i, \eta_i)$) the more valuable it is in providing incentives to the rancher. As a result, the power of the incentive mechanism is increasing in both $Corr(\varepsilon_1, \eta_1)$ and $Corr(\varepsilon_2, \eta_2)$. Third, observe that $Var(P_{e_i})$ captures the noisiness

[25] The expression for the optimal performance regulation when the budget constraint is slack, β^p, is analogous to the expression for the optimal power of an incentive contract in Baker (1992).

of the performance measure. An increase in the noisiness of the performance measure decreases its value to the regulator for providing incentives to the rancher. As a result, the regulator will offer a lower-powered incentive scheme under a relatively noisy performance measure. Finally, the power of the incentive scheme is increasing in the variances of the expected marginal social values of the rancher's efforts, $Var\left(V_{e_1}\right)$ and $Var\left(V_{e_2}\right)$, which measure the value of the rancher's private information about the effect of his inputs on ecosystem health.

When the constraint in (3.34) is binding, the optimal piece rate per unit of the performance measure, $\beta^p(B)$, is determined by the regulator's ability to fund regulation, which is a function of her exogenous budget, B, and the rents from ranching without regulation, $F\left(e_1^{nr}, e_2^{nr}\right) - w_1 e_1^{nr} - w_2 e_2^{nr}$. When the expected bonus payment is positive, i.e., $E\left[\beta^p(B) P\left(e_1^p(\eta_1, B), e_2^p(\eta_2, B)\right)\right] > 0$, higher rents from ranching without regulation allow the regulator to fund a larger $\beta^p(B)$ by transferring less money to the rancher (smaller $\alpha > 0$) without violating the rancher's participation constraint. Similarly, when $E\left[\beta^p(B) P\left(e_1^p(\eta_1, B), e_2^p(\eta_2, B)\right)\right] < 0$, higher rents from ranching without regulation allow the regulator to impose a larger $\beta^p(B)$ without violating the rancher's participation constraint.

The expected social welfare under the optimal performance regulation is given by

$$U^p = \begin{cases} U^{nr} + \left\{ \sum_{i=1}^{2} \frac{\left[E\left(P_{e_i}\right)\right]^2 + Var\left(P_{e_i}\right)}{4\gamma_i} \right\} \left\{ \left(\beta^p\right)^2 - \left[\beta^p - \beta^p(B)\right]^2 \right\}, & \text{if } B \le B^p \\ U^{nr} + \left\{ \sum_{i=1}^{2} \frac{\left[E\left(P_{e_i}\right)\right]^2 + Var\left(P_{e_i}\right)}{4\gamma_i} \right\} \left(\beta^p\right)^2, & \text{otherwise} \end{cases} . \tag{3.40}$$

When the budget constraint is not binding, the extent to which the performance regulation improves welfare is determined by the same factors that determine the power of the incentive scheme; the expected social welfare is increasing in the expected marginal social values of the rancher's inputs $\left|E\left(V_{e_i}\right)\right|$, decreasing in the level of distortions of the performance measure (relatively small values of $Corr\left(\varepsilon_i, \eta_i\right)$), decreasing in the noisiness of the performance measure, $Var\left(P_{e_i}\right)$, and increasing in the variance of the expected marginal social values of the rancher's inputs, $Var\left(V_{e_i}\right)$. Importantly, the quality of the performance measure is a critical determinant of the benefits of performance regulation. When the performance measure provides a good signal of how the rancher's actions influence ecosystem health (low distortion and noisiness), the performance regulation can substantially increase the expected social welfare.

Equation (3.40) reveals that the loss in expected social welfare from the binding budget constraint is a function of the difference between the second-best optimal piece rate, β^p, and the level dictated by the regulator's budget constraint, $\beta^p(B)$. This loss is decreasing in the regulator's budget, B, and the rents from ranching without regulation, $\pi\left(e_1^{nr}, e_2^{nr}\right)$, which influence the regulator's ability to fund regulation, and, hence, $\beta^p(B)$, but do not affect β^p. In addition, this inefficiency is increasing in the expected marginal social values of the rancher's efforts, $\left|E_\varepsilon\left(V_{e_i}\right)\right| = \left|\mu_{i1}\right|$ $(i = 1, 2)$, which increase the expected social benefits from performance regulation, and, hence, the desired level of β^p, but do not affect $\beta^p(B)$, which is determined by the regulator's budget constraint. Note that the loss from the binding budget constraint is independent of whether input e_i $(i = 1, 2)$ is detrimental $\left(E\left(V_{e_i}\right) < 0\right)$ or beneficial $\left(E\left(V_{e_i}\right) > 0\right)$ to ecosystem health. Finally, the loss arising from the binding budget constraint is larger when the performance measure provides a less distorted (high $Corr\left(\varepsilon_i, \eta_i\right)$, $i = 1, 2$) and less noisy (low $Var\left(P_{e_i}\right)$, $i = 1, 2$) signal of the influence of the rancher's input choices on ecosystem health. This follows from the fact that the expected social benefits from performance regulation, and, conversely, the loss in the expected social welfare from the binding budget constraint, are greater when the performance measure provides a better quality signal.

Finally, when the performance measure is a perfect signal of the rancher's input choices $\left(V\left(e_1, e_2, \varepsilon_1, \varepsilon_2\right) = P\left(e_1, e_2, \eta_1, \eta_2\right)\right.$ for each e_1 and $\left.e_2\right)$ and the regulator's budget constraint does not bind $\left(B \geq B^p\right)$, the optimal performance regulation can achieve the social optimum given by (3.9). On the other hand, when the regulator's budget constraint is binding ($B < B^p$), the first-best outcome cannot be achieved even under perfect monitoring.

3.6 Pairwise Comparisons of Regulatory Instruments

We now turn to the analysis of the relative efficiency of the input-mandate and cost-sharing/taxation versus performance regulation. Recall that by construction of the model – the rancher has no private information about how e_1 and e_2 influence his private objective – the input mandate and cost-sharing/taxation are equivalent. We demonstrate in Appendix I that when $E(V_{e_1}) > 0$, the optimal input mandate (or, equivalently, the optimal cost-sharing/taxation) dominates the optimal performance regulation for all B if and only if

$$\frac{[E(V_{e_i})]^2}{4\gamma_1} \geq \left\{ \sum_{i=1}^{2} \frac{[E(P_{e_i})]^2 + Var(P_{e_i})}{4\gamma_i} \right\} (\beta^p)^2 . \tag{3.41}$$

It follows from this inequality that the efficiency of the performance regulation relative to the input mandate or cost-sharing/taxation is determined by the same factors that determine the power of the incentive scheme. In particular, the performance regulation will dominate the input mandate or cost-sharing/taxation when (i) the performance measure is less distorted (large $Corr(P_{e_i}, V_{e_i}) = Corr(\varepsilon_i, \eta_i)$, $i = 1, 2$), (ii) the performance measure is less noisy (small $Var(P_{e_i})$, $i = 1, 2$), and (iii) the rancher's information regarding the effect of his inputs on ecosystem health is more valuable (large $Var(V_{e_i})$, $i = 1, 2$). This finding also reveals that the regulator's budget constraint does not determine her choice between an input mandate or cost-sharing/taxation and performance regulation when $E(V_{e_1}) > 0$.

When $E(V_{e_1}) < 0$, so that the regulator's budget constraint is not relevant for the input mandate or cost-sharing/taxation, and condition (3.41) holds, the optimal input mandate or cost-sharing/taxation will dominate the optimal performance regulation. When $E(V_{e_1}) < 0$ and (3.41) does not hold, however, (3.19) and (3.40) imply that the optimal input mandate or cost-sharing/taxation dominates the optimal performance regulation if and only if

$$\frac{[E(V_{e_i})]^2}{4\gamma_1} \geq \left[(\beta^p)^2 - [\beta^p - \beta^p(B)]^2\right] \left\{ \sum_{i=1}^{2} \frac{[E(P_{e_i})]^2 + Var(P_{e_i})}{4\gamma_i} \right\} > 0. \tag{3.42}$$

Thus, in this case, the efficiency of the performance regulation relative to the input mandate or cost-sharing/taxation is determined by the factors that affect both β^p and $\beta^p(B)$, which are the the optimal piece rates under the non-binding and binding budget constraints, respectively. It follows from (3.42) that the optimal performance regulation is more likely to dominate the optimal input mandate or cost-sharing/taxation when the regulator has a greater ability to fund regulation, which, as was explained above, is determined by her exogenous budget, B, and the rents from ranching without regulation, $\pi(e_1^{nr}, e_2^{nr})$.

Using the expressions for the expected social welfare under the three regulatory instruments, we can determine how different parameters of the model affect the relative efficiency of different regulatory instruments. Take, for example, the effect of the expected marginal product of the unobservable input e_2 on ecosystem health $E(V_{e_2}) = \mu_{21}$. Using (3.19) and (3.40), we obtain the effect of μ_{21} on the efficiency of the performance regulation relative to the input-mandate and cost-sharing/taxation:

$$\frac{\partial(U^p - U^m)}{\partial\mu_{21}} = \frac{\partial(U^p - U^c)}{\partial\mu_{21}} = \begin{cases} \frac{1}{2\gamma_2}\phi_{21}\beta^p(B), \text{ if } B \leq B^p \\ \frac{1}{2\gamma_2}\phi_{21}\beta^p, \text{ otherwise} \end{cases}, \tag{3.43}$$

where β^p and $\beta^p(B)$ are given by (3.35) and (3.36), respectively. It follows from (3.43) that an increase in the magnitude of the expected marginal product of the unobserved input on ecosystem health, $|E(V_{e_2})| = |\mu_{21}|$, will increase the expected social welfare of the performance regulation relative to both the input mandate and cost-sharing/taxation. This follows from the fact that, out of the three regulatory instruments considered in this article, only the performance regulation influences the rancher's choice of e_2, and the regulator's ability to influence the rancher's choice of e_2 through performance regulation is more valuable when e_2 is expected to have a larger impact (in absolute terms) on ecosystem health.

3.7 Joint Use of Regulatory Instruments

3.7.1 Performance-Regulation and Input Mandate

When the input mandate and performance regulation are used jointly, the regulator mandates the rancher's choice of the observable input, denoted by e_1', and institutes a payment to/from the rancher based on the ex post realization of the verifiable performance measure, P. As before, this ex post payment is made according to the linear incentive contract $\alpha+\beta P$, where α is a lump-sum transfer between the regulator and the rancher and $\beta \geq 0$ is the piece rate per unit of P. It is also assumed that the rancher has to pay a relatively high penalty if his use of e_1 differs from e_1' and, as a result, the rancher never finds it advantageous to violate the mandate.

The rancher's optimization problem can be written as

$$\max_{e_2} \left[\pi\left(e_1', e_2\right) + \alpha + \beta P\left(e_1', e_2, \eta_1, \eta_2\right)\right], \tag{3.44}$$

which yields the rancher's optimal choice of e_2 :

$$e_2^{pm}\left(\beta, \eta_2\right) = \begin{cases} 0, & \text{if } \theta_2 + \beta\left(\phi_{21} + \phi_{22}\eta_2\right) \leq w_2 \\ \frac{\theta_2 - w_2 + \beta\left(\phi_{21}+\phi_{22}\eta_2\right)}{2\gamma_2}, & \text{otherwise} \end{cases}. \tag{3.45}$$

Given the rancher's choice $e_2^{pm}\left(\beta, \eta_2\right)$, the regulator's optimization problem can be written as:

$$\max_{e_1',\beta} E\left[\pi\left(e_1', e_2^{pm}\left(\beta,\eta_2\right)\right) + V\left(e_1', e_2^{pm}\left(\beta,\eta_2\right), \varepsilon_1, \varepsilon_2\right)\right] \tag{3.46}$$

$$\text{subject to } E\left[\pi\left(e_1', e_2^{pm}\left(\beta,\eta_2\right)\right)\right] + B \geq 0.$$

In what follows, we focus on the scenarios where the regulator's optimal choice of β is such that the rancher's choice of e_2 is positive for all realizations of η_2. Due to space considerations we only report the results for the non-binding budget constraint case.

The solution to the unconstrained problem (3.46), denoted by e_1^{pm} and β^{pm}, is given by

$$e_1^{pm} = E\left[e_1^*\left(\varepsilon_1\right)\right] = \frac{\theta_1 - w_1 + \mu_{11}}{2\gamma_1} \tag{3.47}$$

and

$$\beta^{pm} = \frac{\frac{\phi_{21}\mu_{21}+\phi_{22}\mu_{22}Cov(\eta,\varepsilon)}{4\gamma_2}}{\frac{\phi_{21}^2+\phi_{22}^2\sigma_\eta^2}{4\gamma_2}} = \frac{\frac{E\left(P_{e_2}\right)E\left(V_{e_2}\right)+\sqrt{Var\left(V_{e_2}\right)Var\left(P_{e_2}\right)}Corr(\varepsilon_2,\eta_2)}{4\gamma_2}}{\frac{\left[E\left(P_{e_2}\right)\right]^2+Var\left(P_{e_2}\right)}{4\gamma_2}}. \tag{3.48}$$

The expected social welfare under the joint use of the input mandate and performance regulation is given by

$$U^{pm} = U^{nr} + \frac{\mu_{11}^2}{4\gamma_1} + \left\{\frac{\left[E\left(P_{e_2}\right)\right]^2 + Var\left(P_{e_2}\right)}{4\gamma_2}\right\}\left(\beta^{pm}\right)^2 \tag{3.49}$$

It follows directly from (3.19) and (3.49) that using an input mandate and performance regulation in tandem is unambiguously preferred to the use of the input mandate in isolation. Similarly, it follows from (3.40) and (3.49) that the input mandate improves the efficiency of the performance regulation for large B (for $B \geq \max\{B^m, B^p\}$) if and only if

$$\frac{\mu_{11}^2}{4\gamma_1} + \frac{\left[\frac{E\left(P_{e_2}\right)E\left(V_{e_2}\right)+\sqrt{Var\left(V_{e_2}\right)Var\left(P_{e_2}\right)}Corr(\varepsilon_2,\eta_2)}{4\gamma_2}\right]^2}{\frac{\left[E\left(P_{e_2}\right)\right]^2+Var\left(P_{e_2}\right)}{4\gamma_2}} > \frac{\left[\sum_{i=1}^{2}\frac{E\left(P_{e_i}\right)E\left(V_{e_i}\right)+\sqrt{Var\left(V_{e_i}\right)Var\left(P_{e_i}\right)}Corr(\varepsilon_i,\eta_i)}{4\gamma_i}\right]^2}{\sum_{i=1}^{2}\frac{\left[E\left(P_{e_i}\right)\right]^2+Var\left(P_{e_i}\right)}{4\gamma_i}}. \tag{3.50}$$

where β^p is defined by (3.35). This inequality reveals that the input mandate will improve the efficiency of the performance regulation when the expected marginal product of the observable input, e_1, on ecosystem health, $|\mu_{11}| = |E(V_{e_1})|$, is relatively large, and when the performance measure provides a relatively poor signal of influence of e_1 on ecosystem health, i.e., high distortion (low $Corr(\varepsilon_1, \eta_1)$) and noisiness (high $Var(P_{e_1})$). Under these conditions, the input mandate improves welfare by limiting the rancher's ability to make socially inefficient choices of e_1, while performance regulation continues to motivate the rancher to pursue a more socially desirable choice of e_2.

The results in this section provide a rationale for the reliance of FLMAs on input mandates (grazing restrictions) in conjunction with penalties/bonuses based on observed performance. Our model predicts that this regulatory mix improves welfare relative to the performance regulation alone when monitoring provides a poor signal for how livestock grazing influences ecological health. Indeed, several studies have documented the difficulties associated with monitoring the relationship between livestock grazing and rangeland health. For example, Holechek (1988) describes the challenges to setting appropriate grazing restrictions for a public land allotment due to heterogeneity in vegetation, soil type, slope, and distance to water, all of which lead to non-uniform forage utilization by livestock across an allotment.

3.7.2 Performance Regulation and Cost-Sharing

When cost-sharing/taxation and performance regulation are used jointly, the regulator subsidizes (cost-shares) or taxes the rancher's use of the observable input, e_1, and mandates a payment based on the ex post realization of the verifiable performance measure, P. As before, the cost of input e_1 incurred by the rancher is $(1-s)w_1e_1$, while the cost borne by the regulator is sw_1e_1. The ex post payment is made according to the linear incentive contract $\alpha + \beta P$, where α is a lump-sum transfer between the regulator and the firm and $\beta \geq 0$ is the piece rate per unit of P.

The rancher's optimization problem can be written as

$$\max_{e_1,e_2} \left[F\left(e_1,e_2\right)-(1-s)w_1e_1-w_2e_2+\alpha+\beta P\left(e_1,e_2,\eta_1,\eta_2\right)\right] \tag{3.51}$$

The rancher's optimal choice of inputs is given by

$$e_1^{pc}\left(s,\beta,\eta_1\right)=\begin{cases} 0, & \text{if } \theta_1+\beta\left(\phi_{11}+\phi_{12}\eta_1\right)\leq(1-s)w_1 \\ \frac{\theta_1-(1-s)w_1+\beta(\phi_{11}+\phi_{12}\eta_1)}{2\gamma_1}, & \text{otherwise} \end{cases}, \tag{3.52}$$

$$e_2^{pc}\left(\beta,\eta_2\right)=\begin{cases} 0, & \text{if } \theta_2+\beta\left(\phi_{21}+\phi_{22}\eta_2\right)\leq w_2 \\ \frac{\theta_2-w_2+\beta(\phi_{21}+\phi_{22}\eta_2)}{2\gamma_2}, & \text{otherwise} \end{cases}.$$

Given the rancher's choices $e_1^{pc}\left(s,\beta,\eta_1\right)$ and $e_2^{pc}\left(\beta,\eta_2\right)$, the regulator's optimization problem can be written as:

$$\max_{s,\beta} E\left[\pi\left(e_1^{pc}\left(s,\beta,\eta_1\right),e_2^{pc}\left(\beta,\eta_2\right)\right)+V\left(e_1^{pc}\left(s,\beta,\eta_1\right),e_2^{pc}\left(\beta,\eta_2\right),\varepsilon_1,\varepsilon_2\right)\right] \tag{3.53}$$

$$\text{subject to } B+E\left[\begin{array}{c} F\left(e_1^{pc}\left(s,\beta,\eta_1\right),e_2^{pc}\left(\beta,\eta_2\right)\right) \\ -\left(1-s\right)w_1e_1^{pc}\left(s,\beta,\eta_1\right)-w_2e_2^{pc}\left(\beta,\eta_2\right) \end{array}\right]\geq 0.$$

In what follows, we consider the scenario where the regulator's optimal choices of s and β are such that $\theta_1+\beta\left(\phi_{11}+\phi_{12}\eta_1\right)>(1-c)w_1$ for all $\eta_1\in[-\eta_1',\eta_1']$ and $\theta_2+\beta\left(\phi_{21}+\phi_{22}\eta_2\right)>w_2$ for all $\eta_2\in[-\eta_2',\eta_2']$. These assumptions ensure that $e_1^{pc}\left(s,\beta,\eta_1\right)>0$ for all $\eta_1\in[-\eta_1',\eta_1']$ and $e_2^{pc}\left(\beta,\eta_2\right)>0$ for all $\eta_2\in[-\eta_2',\eta_2']$. We also restrict our focus in the remainder of this section to the case where the regulator's budget constraint does not bind.

The solution to the unconstrained problem (3.53), denoted by s^{pc} and β^{pc}, is given by

$$s^{pc}=\frac{1}{w_1}\left[E\left(V_{e_1}\right)-\beta^{pc}E\left(P_{e_1}\right)\right]=\frac{1}{w_1}\left(\mu_{11}-\beta^{pc}\phi_{11}\right) \tag{3.54}$$

and

$$\beta^{pc}=\frac{\frac{\phi_{12}\mu_{12}E(\varepsilon_1\eta_1)}{4\gamma_1}+\frac{\phi_{21}\mu_{21}+\phi_{22}\mu_{22}E(\varepsilon_2\eta_2)}{4\gamma_2}}{\frac{\phi_{12}^2\sigma_{\eta_1}^2}{4\gamma_1}+\frac{\phi_{21}^2+\phi_{22}^2\sigma_{\eta_2}^2}{4\gamma_2}}. \tag{3.55}$$

The expected social welfare under the optimal joint use of cost-sharing/taxation and performance regulation is given by

$$U^{pc} = U^{nr} + \frac{\left[E\left(V_{e_1}\right)\right]^2}{4\gamma_1} + \left\{ \frac{Var\left(P_{e_1}\right)}{4\gamma_1} + \frac{\left[E\left(P_{e_2}\right)\right]^2 + Var\left(P_{e_2}\right)}{4\gamma_2} \right\} \left(\hat{\beta}^{pc}\right)^2. \quad (3.56)$$

As expected, the optimal joint use of cost-sharing/taxation and performance regulation unambiguously improves welfare over both cost-sharing/taxation and performance regulation in isolation. The improvement in welfare stems from the fact that cost-sharing/taxation improves welfare by bringing the use of the observable input, e_1, in line with its marginal social costs/marginal social benefits while still giving the rancher flexibility to use his private knowledge of the ranch-level ecological conditions when making his input choices. Under performance regulation alone, the regulator is restricted to offer the same incentive β for both inputs since there is only a single performance measure. An addition of cost-sharing to performance regulation allows the regulator to offer differing incentives for the use of the two inputs.

From (3.49) and (3.56), the joint use of cost-sharing/taxation and performance regulation will dominate the joint use of an input mandate and performance regulation when the performance measure provides a relatively good signal of the influence of the observable input, e_1, on ecosystem health (i.e., low distortion (high $Corr\left(\varepsilon_1, \eta_1\right)$) and noisiness (low $Var\left(P_{e_i}\right)$)). The joint use of cost-sharing/taxation and performance regulation has the advantage of providing the rancher with incentives to use his private knowledge of ranch-level ecological conditions when choosing e_1. This advantage, however, only improves welfare relative to the joint use of an input mandate and performance regulation when the performance measure provides a good signal of e_1's influence on ecosystem health. When the performance measure provides a poor signal, the joint use of an input mandate and performance regulation dominates by fixing the rancher's choice of e_1 at the ex ante social optimum, $E\left[e_1^*\left(\varepsilon_1\right)\right]$.

3.8 Discussion

We have developed a theoretical model of optimal regulation by a budget-constrained regulator under asymmetric information, moral hazard, and imperfect monitoring. We used the model to evaluate the relative efficiency of three prominent regulatory instruments on public rangelands: input mandates, cost-sharing/taxation, and performance regulation. The model extends the received literature by presenting an informational and institutional environment that closely resembles the actual regulation of ranching on public rangeland and allows for the efficiency of these three regulatory instruments to be evaluated and compared.

Our analysis yields a number of results about the relative efficiency of the three regulatory instruments. First, we find that both an input mandate and cost-sharing/taxation improve welfare by bringing the use of observable inputs in line with their marginal social costs/social benefits. In addition, given the information structure of our model – the rancher has no private information about how his input choices influence ranch profits – the input mandate and cost-sharing/taxation are equivalent when they are not used in combination with performance regulation.

Second, we find that performance regulation improves social welfare when the performance measure is a sufficiently accurate signal of how the rancher's activities influence ecosystem health. As one would expect, the optimal performance regulation under perfect monitoring achieves the first-best outcome as long as the budget constraint is non-binding. We find that performance regulation has two main advantages over an input mandate and cost-sharing/taxation. First, in contrast to an input mandate or cost-sharing/taxation, performance regulation allows the regulator to influence the rancher's unobservable inputs (e.g., grazing management to reduce the ecological harm of livestock grazing). This advantage of performance regulation is greater when the unobserved inputs have a larger influence, either positive or negative, on rangeland ecological health. Second, also in contrast to an input mandate or cost-sharing/taxation, performance regulation gives the rancher the incentives and flexibility to use his private knowledge of ranch-level ecological conditions when making his input choices. Given the potential for considerable heterogeneity of range conditions even within small geographic areas, this is an important benefit for the regulation of ranching on public land.

Third, we identify two scenarios where an optimal input mandate or cost-sharing/taxation dominates performance regulation. First, due to the difficulties of monitoring changes in rangeland health on each ranch, FLMAs often have to base regulation on a very distorted or noisy signal of how the rancher's activities have influenced ecosystem health. Under these circumstances, we find that performance regulation can be dominated by an optimal input mandate or cost-sharing/taxation. Second, an optimal input mandate or cost-sharing/taxation can dominate performance regulation when the regulator faces a relatively strict budget constraint so that she is limited in her ability to elicit optimal input use through performance regulation.

Finally, we find that when the performance measure is relatively uninformative, performance regulation in conjunction with an input mandate can dominate either performance regulation alone or performance regulation in combination with cost-sharing. This result provides a rationale for the reliance of budget-constrained FLMAs on input mandates (i.e., grazing restrictions) in conjunction with penalties/bonuses based on observed performance on public rangelands.

An immediate and important extension of the analysis in this article is a determination of the optimal regulatory scheme based on the observable input and the performance measure in a framework where non-linear incentive contracts are allowed. Such a regulatory scheme would undoubtedly improve welfare over the regulatory instruments used in practice and considered in this article, and would provide insight into how existing regulations on public rangelands could be improved given the informational and institutional constraints on efficient regulation faced by FLMAs.

Appendix A

Consumption Spreading Across Household Members

This appendix demonstrates that given the assumptions on household preferences in both the model of supply-side NBG presented in Chapter 1.4 and in the model of demand-side NBG presented in Chapter 1.7.1, household j maximizes utility by spreading consumption evenly across all of its members.

A.1 Homothetic Preferences

This section of the appendix demonstrates that given the assumptions on household preferences in the model presented in Chapter 1.4, household j maximizes utility by spreading consumption evenly across all of its members.

To begin, the utility of household member h at time t is given by

$$u\left(Y_{hjM}(t), Y_{hjS}(t)\right) = \ln\left[\mu Y_{hjM}(t)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\mu)\, Y_{hjS}(t)^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}},$$

where $Y_{hjM}(t)$ and $Y_{hjS}(t)$ are household member h's consumption of M- and S-sector output at time t. For given levels of total household consumption of M- and S-sector output at time t, $\hat{Y}_{jM}(t)$ and $\hat{Y}_{jS}(t)$, household j chooses the consumption of M- and S-sector output for each of its members to maximizes total household utility at time

t (which is the the summation of the utilities of all members),

$$\sum_{h \in \mathbb{L}_j(t)} u\left(Y_{hjM}(t), Y_{hjS}(t)\right),$$

where $\mathbb{L}_j\left(t\right)$ is the set of members of household j at time t, subject to the constraints,

$$\hat{Y}_{jM}(t) - \sum_{h \in \mathbb{L}_j(t)} Y_{hjM}(t) \geq 0,\ \hat{Y}_{jS}(t) - \sum_{h \in \mathbb{L}_j(t)} Y_{hjS}(t) \geq 0,$$

and

$$Y_{hjM}(t) \geq 0, Y_{hjS}(t) \geq 0 \text{ for all } h \in \mathbb{L}_j\left(t\right).$$

The first two constraint must bind at an optimum because the utility of every member of household j is strictly increasing in both $Y_{hjM}(t)$ and $Y_{hjS}(t)$. The second and third constraints will be slack at an optimum because $\lim_{Y_{hjP}(t) \to 0} \frac{\partial u\left(Y_{hjM}(t), Y_{hjS}(t)\right)}{\partial Y_{hjP}(t)} = \infty$ for $P = M, S$ and for all $h \in \mathbb{L}_j\left(t\right)$. Given these observations, the Lagrangean is

$$\mathcal{L} = \sum_{h \in \mathbb{L}_j(t)} u\left(Y_{hjM}(t), Y_{hjS}(t)\right) + \xi_{jM}(t)\left(\hat{Y}_{jM}(t) - \sum_{h \in \mathbb{L}_j(t)} Y_{hjM}(t)\right) + \xi_{jS}(t)\left(\hat{Y}_{jS}(t) - \sum_{h \in \mathbb{L}_j(t)} Y_{hjS}(t)\right)$$

where $\xi_{jM}(t)$ and $\xi_{jS}(t)$ are Lagrangean multipliers. The first order conditions are

$$\frac{\mu Y_{hjM}(t)^{-\frac{1}{\varepsilon}}}{\left[\mu Y_{hjM}(t)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\mu) Y_{hjS}(t)^{\frac{\varepsilon-1}{\varepsilon}}\right]} = \xi_{jM}(t) \text{ for } h \in \mathbb{L}_j\left(t\right), \tag{A.1}$$

$$\frac{(1-\mu) Y_{hjS}(t)^{-\frac{1}{\varepsilon}}}{\left[\mu Y_{hjM}(t)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\mu) Y_{hjS}(t)^{\frac{\varepsilon-1}{\varepsilon}}\right]} = \xi_{jS}(t) \text{ for } h \in \mathbb{L}_j\left(t\right), \tag{A.2}$$

and

$$\hat{Y}_{jM}(t) = \sum_{h \in \mathbb{L}_j(t)} Y_{hjM}(t) \text{ and } \hat{Y}_{jS}(t) = \sum_{h \in \mathbb{L}_j(t)} Y_{hjS}(t). \tag{A.3}$$

Because the objective function is a strictly concave and continuous function mapping $\mathbb{R}^{2 \times l_j(t)}$ (an open convex set) into $\mathbb{R}$ and the constraint functions are both concave and continuos functions mapping $\mathbb{R}^{l_j(t)}$ (an open convex set) into $\mathbb{R}$, and *Slater's condition* is satisfied, the first order conditions are both necessary and sufficient to identify an unique optimal solution (Sundaram 1996, p. 187-188).[1]

The first order conditions imply that each household member consumes the same ratio of M-sector and S-sector output, i.e.,

$$\frac{Y_{hjM}(t)}{Y_{hjS}(t)} = \frac{Y_{h'jM}(t)}{Y_{h'jS}(t)} \text{ for } h, h' \in \mathbb{L}_j(t). \tag{A.4}$$

Together, (A.3) and (A.4) imply that

$$\frac{\hat{Y}_{jM}(t)}{\hat{Y}_{jS}(t)} = \frac{\sum\limits_{h \in \mathbb{L}_j(t)} Y_{hjM}(t)}{\hat{Y}_{jS}(t)} = \frac{\sum\limits_{h \in \mathbb{L}_j(t)} \left(\frac{Y_{h'jM}(t)}{Y_{h'jS}(t)}\right) Y_{hjS}(t)}{\hat{Y}_{jS}(t)} = \frac{\left(\frac{Y_{h'jM}(t)}{Y_{h'jS}(t)}\right) \sum\limits_{h \in \mathbb{L}_j(t)} Y_{hjS}(t)}{\hat{Y}_{jS}(t)} = \frac{Y_{h'jM}(t)}{Y_{h'jS}(t)}$$

for all $h' \in \mathbb{L}_j(t)$. That is, (A.3) and (A.4) imply that the ratio of M-sector and S-sector output that is the same for each household member at time t and equal to the ratio of M-sector and S-sector output for the entire household at time t.

These results imply that the first order conditions ((A.1) and (A.2)) can be re-expressed as

$$Y_{hjM}(t) = \frac{\xi_{jM}(t)}{\mu} \left[\mu + (1-\mu)\left(\frac{\hat{Y}_{jS}(t)}{\hat{Y}_{jM}(t)}\right)^{\frac{\varepsilon-1}{\varepsilon}}\right] \text{ for all } h \in \mathbb{L}_j(t).$$

and

$$Y_{hjS}(t) = \frac{\xi_{jS}(t)}{\mu} \left[\mu \left(\frac{\hat{Y}_{jM}(t)}{\hat{Y}_{jS}(t)}\right)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\mu)\right] \text{ for all } h \in \mathbb{L}_j(t).$$

These two equations demonstrated that for given levels of total household consumption of M- and S-sector output, household j's maximizes utility at time t by choosing identical levels of $Y_{hjM}(t)$ and $Y_{hjS}(t)$ for each of its members. That is, household j maximizes utility at time t by spreading consumption evenly across all of its members.

[1] Slater's condition is satisfied becuase there exists a vector of consumption of M- and S-sector output for each member for household j in $\mathbb{R}^{2\times l_j(t)}$ such that $\hat{Y}_{jM}(t) - \sum\limits_{i \in \mathbb{L}_j(t)} Y_{ijM}(t) > 0$ and $\hat{Y}_{jS}(t) - \sum\limits_{i \in \mathbb{L}_j(t)} Y_{ijS}(t) > 0$ (Sundaram 1996, p. 188-89). For example, if $Y_{ijM}(t) = 0$ and $Y_{ijS}(t) = 0$ for all $i \in \mathbb{L}_j(t)$ the Slater's condition is satisfied.

A.2 Non-Homothetic Preferences

This section of the appendix demonstrates that given the assumptions on household preferences in Chapter 1.7.1, household j maximizes utility by spreading consumption evenly across all of its members.

To begin, the utility of household member h at time t is given by

$$u\left(Y_{hjM}(t), Y_{hjS}(t)\right) = \ln\left[\left(Y_{hjM}(t) - \gamma\right)^{\mu} Y_{hjS}(t)^{1-\mu}\right],$$

where $Y_{hjM}(t)$ and $Y_{hjS}(t)$ are household member h's consumption of M- and S-sector output at time t. For given levels of total household consumption of M- and S-sector output at time t, $\hat{Y}_{jM}(t)$ and $\hat{Y}_{jS}(t)$, household j chooses the consumption of M- and S-sector output for each of its members to maximizes total household utility at time t (which is the the summation of the utilities of all members),

$$\sum_{h\in\mathbb{L}_j(t)} u\left(Y_{hjM}(t), Y_{hjS}(t)\right),$$

where $\mathbb{L}_j\left(t\right)$ is the set of members of household j at time t, subject to the constraints,

$$\hat{Y}_{jM}(t) - \sum_{h\in\mathbb{L}_j(t)} Y_{hjM}(t) \geq 0,\ \hat{Y}_{jS}(t) - \sum_{h\in\mathbb{L}_j(t)} Y_{hjS}(t) \geq 0,$$

and

$$Y_{hjM}(t) \geq 0, Y_{hjS}(t) \geq 0 \text{ for all } h \in \mathbb{L}_j\left(t\right).$$

The first two constraint must bind at an optimum because the utility of every member of household j is strictly increasing in both $Y_{hjM}(t)$ and $Y_{hjS}(t)$. The second and third constraints will be slack at an optimum because $\lim_{Y_{hjP}(t)\to 0} \frac{\partial u\left(Y_{hjM}(t), Y_{hjS}(t)\right)}{\partial Y_{hjP}(t)} = \infty$ for $P = M, S$ and for all $h \in \mathbb{L}_j\left(t\right)$ (i.e., because the Inada conditions hold for every member of household j). Given these observations, the Lagrangean is

$$\mathcal{L} = \sum_{h\in\mathbb{L}_j(t)} u\left(Y_{hjM}(t), Y_{hjS}(t)\right) + \xi_{jM}(t)\left(\hat{Y}_{jM}(t) - \sum_{h\in\mathbb{L}_j(t)} Y_{hjM}(t)\right) + \xi_{jS}(t)\left(\hat{Y}_{jS}(t) - \sum_{h\in\mathbb{L}_j(t)} Y_{hjS}(t)\right)$$

where $\xi_{jM}(t)$ and $\xi_{jS}(t)$ are Lagrangean multipliers. The first order conditions are

$$\begin{aligned}\frac{\partial u\left(Y_{hjM}(t), Y_{hjS}(t)\right)}{\partial Y_{hjM}(t)} &= \frac{\mu}{Y_{hjM}(t)-\gamma} = \xi_{jM}(t) \text{ for } h \in \mathbb{L}_j\left(t\right), \\ \frac{\partial u\left(Y_{hjM}(t), Y_{hjS}(t)\right)}{\partial Y_{hjS}(t)} &= \frac{1-\mu}{Y_{hjS}(t)} = \xi_{jS}(t) \text{ for } h \in \mathbb{L}_j\left(t\right),\end{aligned}$$

and

$$\hat{Y}_{jM}(t) = \sum_{h\in\mathbb{L}_j(t)} Y_{hjM}(t) \text{ and } \hat{Y}_{jS}(t) = \sum_{h\in\mathbb{L}_j(t)} Y_{hjS}(t).$$

Because the objective function is a strictly concave and continuous function mapping $\mathbb{R}^{2\times l_j(t)}$ (an open convex set) into $\mathbb{R}$ and the constraint functions are both concave and continuos functions mapping $\mathbb{R}^{l_j(t)}$ (an open convex set) into $\mathbb{R}$, and, as above, *Slater's condition* is satisfied, the first order conditions are both necessary and sufficient to identify an unique optimal solution (Sundaram 1996, p. 187-188).[2]

The first order conditions imply that

$$Y_{hjM}(t) = \frac{\mu}{\xi_{jM}(t)} + \gamma \text{ and } Y_{hjS}(t) = \frac{1-\mu}{\xi_{jS}(t)} \text{ for } h \in \mathbb{L}_j\left(t\right).$$

That is, the first order conditions imply that given levels of total household consumption of M- and S-sector output, household j's maximizes utility at time t by choosing identical levels of $Y_{hjM}(t)$ and $Y_{hjS}(t)$ for each of its members. That is, household j maximizes utility at time t by spreading consumption evenly across all of its members.

[2] As in the first section of this Appendix, Slater's condition is satisfied becuase there exists a vector of consumption of M- and S-sector output for each member for household j in $\mathbb{R}^{2\times l_j(t)}$ such that $\hat{Y}_{jM}(t) - \sum_{i\in\mathbb{L}_j(t)} Y_{ijM}(t) > 0$ and $\hat{Y}_{jS}(t) - \sum_{i\in\mathbb{L}_j(t)} Y_{ijS}(t) > 0$ (Sundaram 1996, p. 188-89).

Appendix B

Household j's Utility Maximization

B.1 Household j's Utility Maximization Problem: Supply-Side Non-Balanced Growth

This appendix presents household j's utility maximization problem at time t and derives (1.4) and (1.5). Household j's utility maximization problem at time t is

$$\max_{Y_{jM}(t),Y_{jS}(t)} U\left(Y_{jM}(t), Y_{jS}(t); l_j\left(t\right)\right) \tag{B.1}$$

$$s.t.\ E_j(t) - Y_{jM}(t) - p(t)Y_{jS}(t) \geq 0,\ Y_{jM}(t) \geq 0, \text{ and } Y_{jS}(t) \geq 0,$$

where $E_j(t)$ is household j's expenditure at time t, $p\left(t\right)$ is the price of S-sector output at time t with the price of M-sector output normalized to 1. The first constraint in (B.1) must bind at an optimum because household j's utility is strictly increasing in both $Y_{jM}(t)$ and $Y_{jS}(t)$ on $\mathbb{R}^2_+$. The second and third constraints will be slack at an optimum because $\lim_{Y_{jP}(t)\to 0} \frac{\partial U\left[Y_{jM}(t),Y_{jS}(t),l_j(t)\right]}{\partial Y_{jP}(t)} = \infty$ for $P = M, S$. As such, household j's utility maximization problem at time t can be re-expressed as

$$\max_{Y_{jM}(t)>0} l_j\left(t\right) \ln \left\{ \mu \left[\frac{Y_{jM}(t)}{l_j\left(t\right)}\right]^{\frac{\varepsilon-1}{\varepsilon}} + (1-\mu) \left[\frac{E_j(t) - Y_{jM}(t)}{l_j\left(t\right) p(t)}\right]^{\frac{\varepsilon-1}{\varepsilon}} \right\}^{\frac{\varepsilon}{\varepsilon-1}}. \tag{B.2}$$

The first-order condition for utility maximization is

$$\frac{l_j(t)\frac{\varepsilon}{\varepsilon-1}\left\{\frac{\varepsilon-1}{\varepsilon}\frac{\mu}{l_j(t)}\left[\frac{Y_{jM}(t)}{l_j(t)}\right]^{-\frac{1}{\varepsilon}}+\frac{\varepsilon-1}{\varepsilon}\frac{(1-\mu)}{l_j(t)p(t)}\left[\frac{E_j(t)-Y_{jM}(t)}{l_j(t)p(t)}\right]^{-\frac{1}{\varepsilon}}\right\}}{\left\{\mu\left[\frac{Y_{jM}(t)}{l_j(t)}\right]^{\frac{\varepsilon-1}{\varepsilon}}+(1-\mu)\left[\frac{E_j(t)-Y_{jM}(t)}{l_j(t)p(t)}\right]^{\frac{\varepsilon-1}{\varepsilon}}\right\}}=0. \tag{B.3}$$

Because (B.2) is strictly concave in $Y_{jM}(t)$ on $\mathbb{R}_{++}$ (a non-empty convex set), the unique solution to (B.3) is the unique optimal solution to household j's utility maximization problem at time t (Sundaram 1996, p. 186).

The first-order condition implies that household j's utility-maximizing demands for M- and S-sector output at time t are

$$Y_{jM}(t)=\frac{E_j(t)}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p(t)^{1-\varepsilon}} \quad \text{and} \quad Y_{jS}(t)=\frac{E_j(t)}{p(t)+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p(t)^{\varepsilon}},$$

which correspond to the expressions in (1.5)and that household j's indirect utility function at time t is

$$\mathbb{V}_j(t)=l_j(t)\ln\left[\frac{E_j(t)}{l_j(t)}\left[\mu^{\varepsilon}+(1-\mu)^{\varepsilon}p(t)^{1-\varepsilon}\right]^{\frac{1}{\varepsilon-1}}\right],$$

which is the expression in (1.4).

B.2 Household j's Utility Maximization Problem: Non-Homothetic Preferences

This appendix presents household j's utility maximization problem at time t when household preferences are non-homothetic, i.e., for the case of demand-side NBG, and derives (1.79) and (1.78) from Chapter 1.7.1. Given the assumptions that $Y_{jM}(t)>0$ and $Y_{jS}(t)>0$ for all t, household j's utility maximization problem at time t is

$$\max_{Y_{jM}(t),Y_{jS}(t)} U\left(Y_{jM}(t),Y_{jS}(t),l_j(t)\right) \tag{B.4}$$

$$s.t.\ E_j(t)-Y_{jM}(t)-p(t)Y_{jS}(t)\geq 0.$$

As in Chapter 1, the first constraint in Equation B.4 must bind at an optimum because household j's utility is strictly increasing in both $Y_{jM}(t)$ and $Y_{jS}(t)$ on $\mathbb{R}^2_+$. As such, household j's utility maximization problem at time t can be re-expressed as

$$\max_{Y_{jM}(t)} l_j(t) \ln \left\{ \left[\frac{Y_{jM}(t)}{l_j(t)} - \gamma \right]^{\mu} \left[\frac{E_j(t) - Y_{jM}(t)}{p(t) l_j(t)} \right]^{1-\mu} \right\} \tag{B.5}$$

The first-order condition for utility maximization is

$$\frac{\mu \left[\frac{E_j(t) - Y_{jM}(t)}{l_j(t)} \right] - (1-\mu) \left[\frac{Y_{jM}(t) - \gamma l_j(t)}{l_j(t)} \right]}{\left[\frac{Y_{jM}(t) - \gamma l_j(t)}{l_j(t)} \right] \left[\frac{E_j(t) - Y_{jM}(t)}{p(t) l_j(t)} \right]} = 0. \tag{B.6}$$

Because Equation B.5 is strictly concave in $Y_{jM}(t)$ on $\mathbb{R}_{++}$ (a non-empty convex set), the unique solution to Equation B.6 is the unique optimal solution to household j's utility maximization problem at time t (Sundaram 1996, p. 186).

The first-order condition implies that household j's utility maximizing demands for M-sector output and S-sector output at time t are,

$$Y_{jM}(t) = \mu E_j(t) + (1-\mu)\gamma l_j(t) \quad \text{and} \quad Y_{jS}(t) = (E_j(t) - \gamma l_j(t))(1-\mu) p(t)^{-1},$$

which is (1.79) from Chapter 1.7.1, and that household j's indirect utility function at time t is

$$\mathbb{V}_j(t) = l_j(t) \ln \left[\mu^{\mu} (1-\mu)^{(1-\mu)} (E_j(t) - \gamma l_j(t)) p(t)^{\mu - 1} \right],$$

which is (1.78) from Chapter 1.7.1.

Appendix C

Aggregation of Constant Returns to Scale Production Functions

This appendix provides a formal proof that when all firms in a sector operate with identical constant returns to scale production technologies, sectoral production can be represented by an aggregate production function. This appendix uses results from Green (1964).

Green (1964, Theorem 10) proves that when the necessary condition for an optimal distribution of factors among firms that the marginal rate of substitution between any two inputs must be the same for any two firms is satisfied, a sufficient condition for consistent aggregation of individual production functions is that:

1. For each firm, expansion paths are straight lines through their origins.

2. For a given set of marginal rates of substitution, the straight line expansion paths for all firms are parallel.

It is left to demonstrate that these conditions hold when all firms operate with identical constant returns to scale production technologies.

For the case where all firms have identical constant return to scale production technologies, the necessary condition for an optimal distribution of factors among firms implies that

$$\frac{\frac{\partial f(x_{i1},x_{i2},...,x_{in})}{\partial x_{is}}}{\frac{\partial f(x_{i1},x_{i2},...,x_{in})}{\partial x_{ir}}} = \frac{\frac{\partial f(x_{j1},x_{j2},...,x_{jn})}{\partial x_{js}}}{\frac{\partial f(x_{j1},x_{j2},...,x_{jn})}{\partial x_{jr}}},$$

for all firms i, j and all inputs s, r, where x_{is} is firm i's s^{th} input, $i = 1, .., m$ and $s = 1, ..., \not n$, and $f(.)$ is the common constant returns to scale production function shared by all firms. Given that $f(.)$ is homogeneous of degree one, $\frac{\partial f(.)}{\partial x_{is}}$, $j = 1, 2, ..., n$, is homogeneous of degree zero. This means that the above equation can be re-expressed as

$$\frac{\frac{\partial f\left(\frac{x_{i1}}{x_{is}},.,1,.,\frac{x_{in}}{x_{is}}\right)}{\partial x_{is}}}{\frac{\partial f\left(\frac{x_{i1}}{x_{is}},.,1,.,\frac{x_{in}}{x_{is}}\right)}{\partial x_{ir}}} = \frac{\frac{\partial f\left(\frac{x_{i1}}{x_{is}},.,1,.,\frac{x_{in}}{x_{is}}\right)}{\partial x_{js}}}{\frac{\partial f\left(\frac{x_{i1}}{x_{is}},.,1,.,\frac{x_{in}}{x_{is}}\right)}{\partial x_{jr}}}, \tag{C.1}$$

From (C.1), when all firms have identical constant return to scale production technologies, the marginal rate of substitution between inputs depends only on the ratios between inputs. This implies that the ratio of inputs will be the same for all levels of firm output; i.e., that each firm's expansion paths are straight lines. Moreover, homogeneity of degree one implies that all expansion paths go through the origin, i.e., $f(0, 0, ..., 0) = 0 \times f(x_{i1}, x_{i2}, ..., x_{in}) = 0$.

In addition, given that the marginal rate of substitution between any two inputs must be the same for any two firms, these straight line expansion paths for all firms are parallel. This means that the optimal ratios of inputs will be the same for all firms and equal to the ratios of the totals in the sector.

Hence, when all firms have identical constant return to scale production technologies, the expansion paths for all firms are parallel straight lines through their origins, so that consistent aggregation is possible and sectoral production can be represented by an aggregate production function.

To see that it is indeed the case that sectoral production can be represented by an aggregate production function when all firms have identical constant return to scale production technologies, note that the condition the optimal ratios of inputs will be the same for all firms and equal to the ratios of the totals in the economy implies that firm i's use of input x_{is} depends only on firm i's output, y_i, and the ratios of the totals in the sector, $\frac{x_1}{x_s}, ., 1, ., \frac{x_n}{x_s}$, where $x_s = \sum_{i=1}^{m} x_{is}$ is the total use of input s in the sector; i.e.,

$$y_i = f\left(x_{i1}, x_{i2}, ..., x_{in}\right) = x_{is} f\left(\frac{x_{i1}}{x_{is}}, ., 1, ., \frac{x_{in}}{x_{is}}\right) = x_{is} f\left(\frac{x_1}{x_s}, ., 1, ., \frac{x_n}{x_s}\right). \qquad \text{(C.2)}$$

(C.2) implies that the total output in the sector can be expressed in terms of the aggregate production function,

$$y = \sum_{i=1}^{m} y_i = \left(\sum_{i=1}^{m} x_{is}\right) f\left(\frac{x_1}{x_s}, ., 1, ., \frac{x_n}{x_s}\right) = x_s f\left(\frac{x_1}{x_s}, ., 1, ., \frac{x_n}{x_s}\right) = f\left(x_1, x_2, ..., x_n\right).$$

Appendix D

M-sector Cost-Minimization

This appendix describes how the cost-minimizing input choices for M-sector firms in region k at time t are derived in two stages.[1] In the first stage, M-sector firms in region k minimize the cost of purchasing a given quantity of $I_k(t)$:

$$\min_{[q_i(t)]_{i\in\mathbb{M}(t)}} \sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t) \tag{D.1}$$

$$s.t. \left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k \geq 0 \text{ and } q_i(t) \geq 0,\ i \in \mathbb{M}(t)\,.$$

The first constraint in (D.1) will bind at an optimum because the cost of purchasing a given quantity of $I_k(t)$ is strictly increasing in $q_{ik}(t)$ for all $i = 1, ..., N(t)$ (i.e., because $p_{ik}(t) > 0$ for all $i = 1, ..., N(t)$). The remaining $N(t)$ constraints will be slack at an optimum because $\lim\limits_{q_k(i)\to 0} \frac{\partial I_k(t)}{\partial q_{ik}(t)} = \lim\limits_{q_k(i)\to 0} \left(\frac{I_k(t)}{q_{ik}(t)}\right)^{1-\alpha} = \infty$ for all $i = 1, ..., N(t)$. Given these observations, the Lagrangean for this problem is

$$\mathcal{L} = \sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t) - \iota(t) \left\{ \left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k \right\}$$

---

[1] Green (1964, Theorem 4) proves that a necessary and sufficient condition for (i) the consistency of the two-stage maximization (minimization) procedure and (ii) for there to be a price index, $P_k(t)$, such that $P_k(t)\, I_k(t) = \sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t)$, is for each quantity index – i.e., $I_k(t)$ and $L_{kM}(t)$ – to be homogeneous of degree one, which is the case for M-sector firms in region k at time t.

where $\iota(t)$ is the Lagrangean multiplier. The first-order conditions are

$$p_{ik}(t) - \iota(t)\left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}-1} q_{ik}(t)^{\alpha-1} = 0 \text{ for all } i \in \mathbb{M}(t) \tag{D.2}$$

and

$$\left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k = 0. \tag{D.3}$$

Because the objective function, $\sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t)$, is a convex and continuous function mapping $\mathbb{R}^{N(t)}$ (an open convex set) into $\mathbb{R}$, the constraint function, $\left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k$, is a concave and continuous function mapping $\mathbb{R}^{N(t)}$ (an open convex set) into $\mathbb{R}$, and *Slater's condition* is satisfied, the first-order conditions in (D.2) and (D.3) are both necessary and sufficient to identify the optimal solution to (D.1) (Sundaram, 1996, p. 187-188).[2,3]

(D.2) implies that the cost-minimizing demands for intermediate inputs by M-

---

[2]Slater's condition is satisfied becuase there exists a vector of intermediate inputs in $\mathbb{R}^{N(t)}$ such that $\left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k > 0$ (Sundaram 1996, p. 188-89). For example, if $q_{ik}(t) = \left(\bar{I}_k + 1\right) N(t)^{-\frac{1}{\alpha}}$ for all $i \in \mathbb{M}(t)$ then $\left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k = 1 > 0$.

[3]Given that $\sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t)$ is a continuous function mapping $\mathbb{R}^{N(t)}$ (an open convex set) into $\mathbb{R}$, a sufficient condition for convexity is for the Hessian of $\sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t)$ to be a symmetric positive semi-definite matrix for every vector of intermediate inputs in $\mathbb{R}^{N(t)}$ (Sundaram 1996, p. 184). The Hessian of $\sum_{i\in\mathbb{M}(t)} p_{ik}(t)\, q_{ik}(t)$ is an $N(t) \times N(t)$ matrix of zeros for every vector of intermediate inputs in $\mathbb{R}^{N(t)}$, which is symmetric and positive semi-definite. As was mentioned in Chapter 1.4.2, the assumption that $\alpha < 1$ implies that $\left[\sum_{i\in\mathbb{M}(t)} q_{ik}(t)^{\alpha}\right]^{\frac{1}{\alpha}} - \bar{I}_k$ is a concave function (see Dixit and Stiglitz 1977).

sector firms in region k satisfy

$$q_{ik}(t) = q_{i'k}(t) \left(\frac{p_{ik}(t)}{p_{i'k}(t)} \right)^{\frac{1}{\alpha-1}} \text{ for all } i, i' \in \mathbb{M}(t).$$

Substituting this expression for $q_{ik}(t)$ into (D.3) and rearranging terms gives the following expression for the cost-minimizing demand for firm i''s variety of intermediate input by M-sector firms in region k:

$$q_{i'k}(t) = p_{i'k}(t)^{\frac{1}{\alpha-1}} \bar{I}_k \left[\sum_{i \in \mathbb{M}(t)} p_{ik}(t)^{\frac{\alpha}{\alpha-1}} \right]^{\frac{-1}{\alpha}}. \tag{D.4}$$

Substituting this expression into (D.1) implies that the minimum cost of purchasing $\bar{I}_k$ for M-sector firms in region k at time t can be expressed as

$$\begin{aligned} \sum_{i' \in \mathbb{M}(t)} p_{i'k}(t) q_{i'k}(t) &= \sum_{i' \in \mathbb{M}(t)} p_{i'k}(t) p_{i'k}(t)^{\frac{1}{\alpha-1}} \bar{I}_k \left[\sum_{i \in \mathbb{M}(t)} p_{ik}(t)^{\frac{\alpha}{\alpha-1}} \right]^{\frac{-1}{\alpha}} \\ &= \bar{I}_k \sum_{i' \in \mathbb{M}(t)} p_{i'k}(t)^{\frac{\alpha}{\alpha-1}} \left[\sum_{i \in \mathbb{M}(t)} p_{ik}(t)^{\frac{\alpha}{\alpha-1}} \right]^{\frac{-1}{\alpha}} \\ &= \bar{I}_k \left[\sum_{i \in \mathbb{M}(t)} p_{ik}(t)^{\frac{\alpha}{\alpha-1}} \right]^{\frac{\alpha-1}{\alpha}} = \bar{I}_k P_k(t). \end{aligned} \tag{D.5}$$

From (D.5), $P_k(t) \equiv \left[\sum_{i \in \mathbb{M}(t)} p_{ik}(t)^{\frac{\alpha}{\alpha-1}} \right]^{\frac{\alpha-1}{\alpha}}$ is a price index that gives the minimum cost for M-sector firms in region k of purchasing a unit of $I_k(t)$ at time t.

In the second stage of the cost-minimization problem, M-sector firms in region k choose $I_k(t)$ and $L_{kM}(t)$ to minimize the cost of producing a given level of output. That is, the cost minimization problem for M-sector firms in region k at time t becomes

$$\min_{I_k(t), L_{kM}(t)} P_k(t) I_k(t) + w_k(t) L_{kM}(t) \tag{D.6}$$
$$s.t. \;\; I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM} \geq 0, \; I_k(t) \geq 0, \text{ and } L_{kM}(t) \geq 0.$$

The first constraint in (D.1) will bind at an optimum because the cost of purchasing a given quantity of $Y_{kM}(t)$ is strictly increasing in $I_k(t)$ and $L_{kM}(t)$ (that is, because

$P_k(t) > 0$ and $w_k(t) > 0$). The second and third constraints will be slack at an optimum because $\lim_{I_k(t)\to 0} \partial Y_{kM}(t) / \partial I_k(t) = \infty$ and $\lim_{L_{kM}(t)\to 0} \partial Y_{kM}(t) / \partial L_{kM}(t) = \infty$. As such, the Lagrangean for the second stage of the cost-minimization problem is

$$\mathcal{L} = P_k(t) I_k(t) + w_k(t) L_{kM}(t) - \varsigma(t) \left[I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM} \right],$$

where $\varsigma(t)$ is the Lagrangean multiplier. The first-order conditions are

$$P_k(t) - \varsigma(t) \alpha I_k(t)^{\alpha-1} L_{kM}(t)^{1-\alpha} = 0 \quad \text{(D.7)}$$

$$w_k(t) - \varsigma(t) (1-\alpha) I_k(t)^{\alpha} L_{kM}(t)^{-\alpha} = 0 \quad \text{(D.8)}$$

and

$$I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM} = 0. \quad \text{(D.9)}$$

Because the objective function, $P_k(t) I_k(t) + w_k(t) L_{kM}(t)$, is a convex and continuous function mapping $\mathbb{R}^2$ (an open convex set) into $\mathbb{R}$, the constraint function, $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM}$, is a concave and continuos function mapping $\mathbb{R}^2$ (an open convex set) into $\mathbb{R}$, and *Slater's condition* is satisfied, the first-order conditions in (D.7) to (D.9) are both necessary and sufficient to identify the optimal solution to (D.6) (Sundaram, 1996, p. 187-188).[4,5]

---

[4]Slater's condition is satisfied becuase there exists a vector $[I_k(t), L_{kM}(t)]$ in $\mathbb{R}^2$ such that $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM} > 0$ (Sundaram 1996, p. 188-89). For example, if $I_k(t) = \bar{Y}_{kM} + 1$ and $L_{kM}(t) = \bar{Y}_{kM} + 1$ then $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM} = 1 > 0$.

[5]Given that $P_k(t) I_k(t) + w_k(t) L_{kM}(t)$ is a continuous function mapping $\mathbb{R}^2$ (an open convex set) into $\mathbb{R}$, a sufficient condition for convexity is for the Hessian of $P_k(t) I_k(t) + w_k(t) L_{kM}(t)$ to be positive semi-definite for every vector of $I_k(t)$ and $L_{kM}(t)$ in $\mathbb{R}^2$ (Sundaram 1996, p. 184). The Hessian of $P_k(t) I_k(t) + w_k(t) L_{kM}(t)$ is an 2×2 matrix of zeros for every vector of $I_k(t)$ and $L_{kM}(t)$ in $\mathbb{R}^2$, which is positive semi-definite.

Similarly, given that $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM}$ is a continuous function mapping $\mathbb{R}^2$ (an open convex set) into $\mathbb{R}$, a sufficient condition for concavity is for the Hessian of $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM}$ to be negative semi-definite for every vector of $I_k(t)$ and $L_{kM}(t)$ in $\mathbb{R}^2$ (Sundaram 1996, p. 184).

The first-order conditions in (D.7) to (D.9) give the following expressions for the cost-minimizing input demands:

$$\begin{aligned} I_k(t) &= Y_{kM} \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \left(\frac{w_k(t)}{P_k(t)}\right)^{1-\alpha} \\ L_{kM}(t) &= Y_{kM} \left(\frac{\alpha}{1-\alpha}\right)^{-\alpha} \left(\frac{w_k(t)}{P_k(t)}\right)^{-\alpha}, \end{aligned} \tag{D.10}$$

and for the cost function for the M-sector in region k at time t:

$$C_{kM}(t) = Y_{kM} \left(\frac{P_k(t)}{\alpha}\right)^{\alpha} \left(\frac{w_k(t)}{1-\alpha}\right)^{1-\alpha}. \tag{D.11}$$

From (D.4), (D.5), and (D.10), the demands for firm i_k's variety of intermediate input by M-sector firms in region k and region l at time t are

$$q_{i_k k}(t) = p_{i_k k}(t)^{\frac{1}{\alpha-1}} \left(\frac{\alpha}{1-\alpha}\right) w_k(t) L_{kM}(t) P_k(t)^{\frac{-\alpha}{\alpha-1}} \tag{D.12}$$

and

$$q_{i_k l}(t) = p_{i_k l}(t)^{\frac{1}{\alpha-1}} \left(\frac{\alpha}{1-\alpha}\right) w_l(t) L_{lM}(t) P_l(t)^{\frac{-\alpha}{\alpha-1}}.$$

(D.10), (D.11), and (D.12) are (1.17), (1.18), and (1.21) in Section 1.4.2.

The Hessian of $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM}$ is a symmetric 2×2 matrix where the first naturally ordered principal minors are $\alpha(\alpha-1) I_k(t)^{\alpha-2} L_{kM}(t)^{1-\alpha} \leq 0$ and $\alpha(\alpha-1) I_k(t)^{\alpha} L_{kM}(t)^{-\alpha-1} \leq 0$, and the determinant of the second naturally ordered principal minor equals 0 for every $I_k(t) \geq 0$ and $L_{kM}(t) \geq 0$. Hence, the Hessian of $I_k(t)^{\alpha} L_{kM}(t)^{1-\alpha} - \bar{Y}_{kM}$ is negative semi-definite.

Appendix E

Existence of a Walrasian Equilibrium

In this appendix, the existence of a Walrasian equilibrium for the economies in Chapters 1.4 and 1.7.1 is established in two steps. First, it is demonstrated that a Walrasian quasiequilibrium exists for the economies in Chapters 1.4 and 1.7.1. Second, it is demonstrated that the "cheaper consumption" condition is satisfied in any Walrasian quasiequilibrium for the economies in Chapters 1.4 and 1.7.1. When the cheaper consumption condition is satisfied, a Walrasian quasiequilibrium is also a Walrasian equilibrium. Hence, demonstrating that a Walrasian quasiequilibrium exists and that the cheaper consumption condition is satisfied establishes that a Walrasian equilibrium exists for the economies in Chapters 1.4 and 1.7.1. This method for establishing the existence of a Walrasian equilibrium follows Propositions 17.BB.1 and 17.BB.2 from Mas-Colell, Whinston, and Green (1995).

E.1 Definitions

Before getting started, it is necessary to define the consumption and production sets for the economies in Chapters 1.4 and 1.7.1 and give a formal definition for a Walrasian equilibrium. The consumption and production sets are identical for the economies in both chapters. The proof of the existence of an equilibrium offered in this appendix therefore holds for both chapters.

The consumption set for household j at time t is

$$X_j(t) = \mathbb{R}^2_+ = \left\{ (Y_{jM}(t), Y_{jS}(t)) \in \mathbb{R}^2 \middle| Y_{jM}(t) \geq 0 \text{ and } Y_{jS}(t) \geq 0 \right\}.$$

The production sets for the economy in Chapters 1.4 and 1.7.1 are as follows. When defining production sets, the convention that positive numbers denote outputs and negative numbers denote inputs is used. The production set for the M-sector in region $k = a, b$ at time t is

$$\mathcal{Y}_{kM}(t) = \left\{ \begin{array}{c|c} \left(Y_{kM}(t), -L_{kM}(t), [-q_{ik}(t)]_{i=1}^{N(t)}\right) & \begin{array}{r} \left[\sum_{i \in \mathbb{M}(t)} q_{ik}(t)^{\alpha}\right] L_{kM}(t)^{1-\alpha} - Y_{kM}(t) \geq 0, \\ L_{kM}(t) \geq 0, \text{ and } q_i(t) \geq 0, \; i = 1, ..., N(t) \end{array} \end{array} \right\}.$$

The production set for the S-sector in region k at time t for $F_k > 0$ is

$$\mathcal{Y}_{kS}(t) = \left\{ (Y_{kS}(t), -L_{kS}(t)) \middle| L_{kS}(t)^{\beta} F_k^{1-\beta} - Y_{kS}(t) \geq 0 \text{ and } L_{kS}(t) \geq 0 \right\}.$$

The production set for I-sector firm $i_k \in \mathbb{M}_k(t)$ at time t is

$$\mathcal{Y}_{i_k}(t) = \{ (Y_{i_k}(t), -Y_{i_k M}(t)) | \, Y_{i_k M}(t) - Y_{i_k}(t) \geq 0 \text{ and } Y_{i_k M}(t) \geq 0 \},$$

where, for the purposes of defining the production set, $Y_{i_k}(t)$ is the quantity of intermediate input produced by firm i_k at time t, and $Y_{i_k M}(t) \geq 0$ is the quantity of M-sector output used in production by firm i_k at time t. Finally, the production set for I-sector innovation in region k at time t is

$$\mathcal{Y}_{kI}(t) = \left\{ \left(\dot{N}_k(t), -L_{kI}(t)\right) \middle| \Omega_k(t) L_{kI}(t) - \dot{N}_k(t) \geq 0 \text{ and } L_{kI}(t) \geq 0 \right\}.$$

Given these definitions of the consumption and production sets, an allocation for the economy at time t is a specification of a consumption vector for each household $j = 1, ..., H$, $(Y_{jM}(t), Y_{jS}(t)) \in X_j(t)$, and a production vector for each firm/sector, $\left(Y_{kM}(t), -L_{kM}, [-q_{ik}(t)]_{i=1}^{N(t)}\right) \in \mathcal{Y}_{kM}(t)$, $(Y_{kS}(t), -L_{kS}) \in \mathcal{Y}_{kS}(t)$, $(Y_{i_k}(t), -Y_{i_k M}(t)) \in \mathcal{Y}_{i_k}(t)$, $i_k = 1, ..., N_k(t)$, and $\left(\dot{N}_k(t), -L_{kI}(t)\right) \in \mathcal{Y}_{kI}(t)$ for $k = a, b$. An allocation for the economy at time t is feasible if the market-clearing conditions for labor, $Y_M(t)$, $Y_S(t)$, and $q_i(t)$, $i = 1, ..., N(t)$, $k = a, b$, given in Chapter 1.4.2 hold.

An allocation and a price vector,

$$\left(p(t), r(t), w_b(t), w_a(t), w_a^F(t), w_b^F(t), [p_a(i,t)]_{i=1}^{N(t)}, [p_b(i,t)]_{i=1}^{N(t)}\right),$$

is a Walrasian equilibrium if:

1. The production vectors maximize profits for each firm/sector.

2. Each household's consumption vector maximizes their utility (i.e., is maximal given for their preferences, $\succeq_j$) on their budget set

$$\{(Y_{jM}(t), Y_{jS}(t)) \in X_j(t) | E(j,t) - Y_{jM}(t) - p(t) Y_{jS}(t) \geq 0\}.$$

3. The market-clearing conditions given in Chapter 1.4.2 hold.

It is clear from this definition that together, the static equilibrium conditions at time t given in Chapter 1.5 defines a Walrasian equilibrium. That is, when the static equilibrium conditions hold at time t, the allocation and price vector for the economy ensure that firms maximize profits given their technologies, households maximize utility given their budget constraints, and all good and factor markets clear.

E.2 Existence of a Walrasian Quasiequilibrium

This appendix will use the sufficient conditions for a Walrasian quasiequilibrium to exist for an economy given by Mas-Colell, Whinston, and Green (1995) in Proposition 17.BB.2. That is, it will be shown that a Walrasian quasiequilibrium exists for the economies in Chapters 1.4 and 1.7.1 because the following conditions hold:

1. For every household j in the economy for all t:

(a) its consumption set, $X_j(t)$, is closed and convex. (The consumption set for household j is $X_j(t) = \mathbb{R}^2_+$ for all t, which is a convex and closed set.)

(b) its preference relation, $\succeq_j$, is rational, continuous, locally non-satiated, and convex preference relation defined on $X_j(t)$.

(c) Each household's initial endowment of $Y_{jM}(t)$ and $Y_{jS}(t)$ is greater than or equal to some element in its consumption set. (This condition is satisfied for all households in Chapters 1.4 and 1.7.1 because each households initial endowment of both $Y_{jM}(t)$ and $Y_{jS}(t)$ at time t is zero and $(0,0) \in X_j(t)$ for all $j = 1, ..., H$. Recall that $Y_{jM}(t)$ and $Y_{jS}(t)$ are perishable.)

2. Every production set in the economy is closed, convex, includes the origin, and satisfies the free-disposal property for all t.

3. The set of feasible allocations for the economy is compact for all t.

E.2.1 Preferences

In this section, it will be shown that household j's preference relation is rational (complete and transitive), continuous, locally non-satiated, and convex. That is, it will be shown that household j's preference relation satisfies 1.b. The existence of household j's utility function at time t, $U(Y_{jM}(t), Y_{jS}(t), l_j(t))$, implies that household j's preferences relation satisfies

$$x \succeq_j y \Leftrightarrow U(x_1, x_2, l_j(t)) \geq U(y_1, y_2, l_j(t)),$$

where $x = (x_1, x_2)$ and $y = (y_1, y_2)$ are bundles in household j's consumption set, i.e., $x, y \in X_j(t)$.

To begin, it will be established that household j's preferences are complete and transitive. First, that household j's preferences are complete follows from the fact that $U(Y_{jM}(t), Y_{jS}(t), l_j(t))$ is continuous on $X_j(t)$, so that $U(x_1, x_2, l_j(t)) \geq U(y_1, y_2, l_j(t))$

or $U(y_1, y_2, l_j(t)) \geq U(x_1, x_2, l_j(t))$ (or both) for all $x, y \in X_j(t)$. This implies that $x \succeq_j y$ or $y \succeq_j x$ (or both) for all $x, y \in X_j(t)$, i.e., that household j's preferences are complete. Second, household j's preferences are transitive for all $x, y, z \in X_j(t)$ if $x \succeq_j y$ and $y \succeq_j z$ implies that $x \succeq_j z$. For household j, $x \succeq_j y$ and $y \succeq_j z$ implies that $U(x_1, x_2, l_j(t)) \geq U(y_1, y_2, l_j(t)) \geq U(z_1, z_2, l_j(t))$, which in turn implies that $x \succeq_j z$. This establishes that household j's preferences are transitive.

To establish that household j's preference are continuous, it is sufficient to demonstrate that for all $x \in X_j(t)$, the upper contour set $\{y \in X_j(t) | \, y \succeq x\}$ and the lower contour set $\{y \in X_j(t) | \, x \succeq y\}$ are closed sets, i.e., include their boundaries. The upper contour set can be re-expressed as $\{y \in X_j(t) | \, U(y_1, y_2, l_j(t)) \geq U(x_1, x_2, l_j(t))\}$, and the lower contour set can be re-expressed as $\{y \in X_j(t) | \, U(x_1, x_2, l_j(t)) \geq U(y_1, y_2, l_j(t))\}$. That the upper and lower contour sets are closed for all $x \in X_j(t)$ follows from the continuity of $U(Y_{jM}(t), Y_{jS}(t), l_j(t))$ and from the fact that the upper and lower contour sets are defined by weak inequalities.

In order to establish that household j's preferences are locally non-satiated, it is sufficient to prove the stronger property that household j's preferences are strongly monotone. That household j's preferences are strongly monotone follows from the fact that $U(Y_{jM}(t), Y_{jS}(t), l_j(t))$ is continuous and strictly increasing in both $Y_{jM}(t)$ and $Y_{jS}(t)$ over $X_j(t)$. These properties of $U(Y_{jM}(t), Y_{jS}(t), l_j(t))$ imply that if $x \geq y$ and $x \neq y$ (i.e., $x_1 > y_1$ or $x_2 > y_2$ or both), $U(x_1, x_2, l_j(t)) > U(y_1, y_2, l_j(t))$ and $x \succ_j y$. This establishes that household j's preferences are strongly monotone and, hence, locally non-satiated.

Finally, household j's preference are convex if for every $x \in X_j(t)$ the upper contour set $\{y \in X_j(t) | \, y \succeq x\}$ is convex. This upper contour set can be re-expressed as $\{y \in X_j(t) | \, U(y_1, y_2, l_j(t)) \geq U(x_1, x_2, l_j(t))\}$, which is a convex set by the concavity of $U(Y_{jM}(t), Y_{jS}(t), l_j(t))$.

E.2.2 Production Sets

Now it must be shown that every production set in the economies in Chapters 1.4 and 1.7.1 is closed, convex, includes the origin, and satisfies the free-disposal property. To begin, the production sets are closed because they are defined by weak inequalities. The production sets are convex because they represent the intersection of upper contour sets of quasi-concave functions, which are convex. It is also clear that the production sets include the origin (for example, $(0,0,0) \in Y_{kM}(t)$) and satisfy the free-disposal property (for example, if $y = (y_1, y_2, y_3) \in Y_{kM}(t)$ and $y' = (y'_1, y'_2, y'_3) \leq y$, then $y' \in Y_{kM}(t)$).

In addition, these production sets satisfy the no-free-lunch property (for example, if $y = (y_1, y_2, y_3) \in Y_{kM}(t)$ and $y \geq 0$, then $y = (0,0,0)$) and the irreversibility property (for example, if $y = (y_1, y_2, y_3) \in Y_{kM}(t)$ and $y \neq (0,0,0)$, then $-y \notin Y_{kM}(t)$). The no-free-lunch and irreversibility properties are important for the demonstration that the set of feasible allocations is compact.

E.2.3 Set of Feasible Allocations is Compact

It is left to demonstrate that the set of feasible allocations is compact. To establish this result, this appendix will use the sufficient condition for the set of feasible allocations to be compact given by Mas-Colell, Whinston, and Green (1995) in Proposition 16.AA.1. That is, the set of feasible allocations is compact if the following hold:

1. Every $X_j(t)$ for all t

   (a) is closed (established above).

   (b) is bounded below, i.e., no consumer can supply the market with an arbitrary large amount of any good. (This condition holds in Chapters 1.4 and 1.7.1 because the initial endowments of $Y_{jM}(t)$ and $Y_{jS}(t)$ are zero for every household and for all t.)

2. Every production set is closed (established above). Moreover, the aggregate production set

$$\mathcal{Y}(t) = \mathcal{Y}_{aM}(t)+\mathcal{Y}_{bM}(t)+\mathcal{Y}_{aS}(t)+\mathcal{Y}_{bS}(t)+\mathcal{Y}_{aI}(t)+\mathcal{Y}_{bI}(t)+\sum_{i_a\in\mathbb{M}_a(t)}\mathcal{Y}(i_a,t)+\sum_{i_b\in\mathbb{M}_b(t)}\mathcal{Y}(i_b,t)$$

   (a) is convex. (The aggregate production set is convex as it is the sum of convex sets.)

   (b) admits the possibility of inaction. (This property holds for the aggregate production set because it holds for each production set in the economy.)

   (c) satisfies the no-free-lunch property. (This property also holds for the aggregate production set because it holds for each production set in the economy.)

   (d) is irreversible. (Again, this property holds for the aggregate production set because it holds for each production set in the economy.)

E.3 Cheaper Consumption

The previous section established the existence of a Walrasian quasiequilibrium for the economies in Chapters 1.4 and 1.7.1. To establish that Walrasian quasiequilibria are also Walrasian equilibria it is necessary to demonstrate that the "cheaper consumption" condition is satisfied in any Walrasian quasiequilibrium. The cheaper consumption condition for household j at time t is satisfied if there is $x \in X_j(t)$ such that $x_1 + p(t)\, x_2 < E(j,t)$, where $E(j,t)$ is household j's expenditure at time t. The cheaper consumption condition must hold for all households in the economy in any Walrasian quasiequilibrium, because $\lim_{y_p(j,t)\to 0} \frac{\partial U\left(Y_{jM}(t),Y_{jS}(t),l_j(t)\right)}{\partial Y_{jP}(t)} = \infty$ for $P = M, S$ implies that $Y_{jM}(t) > 0$ and $Y_{jS}(t) > 0$, and hence $E(j,t) > 0$, for all t. That households can finance $E(j,t) > 0$ for all t follows from the fact that households have positive labor income, $w(t)\, l_j(t)$, for all t.

Appendix F

Proofs of Proposition 1 and Theorems 1 and 2

F.1 Proof of Proposition 1

Proof. To being, (1.35) from Proposition 1 (i.e., that $p_{i_k k}(t) = p_{kk}(t) = 1/\alpha$ and $p_{i_k l}(t) = p_{kl}(t) = \Gamma/\alpha$ for all $i_k \in \mathbb{M}_k(t)$ and for all t) is a direct implication of the conditions for profit-maximization for I-sector firms presented in Chapter 1.4.2. Substituting these expressions for $p_{kk}(t)$ and $p_{kl}(t)$ into the expression for $P_k(t)$ yields (1.36) from Proposition 1:

$$P_k(t) = \left[\sum_{i_k \in \mathbb{M}_k(t)} \left(\frac{1}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} + \sum_{i_l \in \mathbb{M}_l(t)} \left(\frac{\Gamma}{\alpha} \right)^{\frac{\alpha}{\alpha-1}} \right]^{\frac{\alpha-1}{\alpha}} = \alpha^{-1} \left(N_k(t) + \Gamma^{\frac{\alpha}{\alpha-1}} N_l(t) \right)^{\frac{\alpha-1}{\alpha}}.$$

Furthermore, substituting this expression for $P_k(t)$ into the expression for $w_k(t)$ from (1.20) implies that when the M-sector is in production in region k at time t, the equilibrium wage in region k is (1.37) from Proposition 1

$$w_k(t) = (1-\alpha)\,\alpha^{\frac{-\alpha}{\alpha-1}} P_k(t)^{\frac{\alpha}{\alpha-1}} = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}} \left(N_k(t) + \Gamma^{\frac{\alpha}{\alpha-1}} N_l(t) \right).$$

The owners of the fixed factor in region k will set the rental rate for the fixed factor at time t, $w_k^F(t)$, so that S-sector firms have just enough incentive to remain in operation, i.e., so that there are zero profits in the S-sector. If there were positive profits in the S-sector, the owners of the fixed factor in region k could costlessly increase their income by increasing $w_k^F(t)$. Substituting the expression for $L_{kS}(t)$ from (1.26) into expression for S-sector profits in region k at time t from (1.24), and setting the resulting expression equal to zero yields

$$p(t)\left(\frac{\beta p(t)}{w_k(t)}\right)^{\frac{\beta}{1-\beta}} {F_k}^{\beta} {F_k}^{1-\beta} - w_k(t)\left(\frac{\beta p(t)}{w_k(t)}\right)^{\frac{1}{1-\beta}} F_k - w_k^F(t) F_k = 0. \tag{F.1}$$

Rearranging (F.1) gives the expression for the rental rate for the fixed factor in region k at time t in (1.38) from Proposition 1:

$$w_k^F(t) = (1-\beta) p(t)\left(\frac{\beta p(t)}{w_k(t)}\right)^{\frac{\beta}{1-\beta}}.$$

As each $q_{ik}(t)$ enters M-sector production symmetrically and the price of each $q_{ik}(t)$ is fixed, M-sector firms in region k demand equal quantities of each intermediate input produced in region k and demand equal, but different, quantities of each intermediate input produced in region l. This means that all I-sector firms in region k face the same demand for their output, i.e., $q_{i_k k}(t) = q_{kk}(t)$ and $q_{i_k l}(t) = q_{kl}(t)$ for all $i_k \in \mathbb{M}_k(t)$. Using (1.21), (1.36), and (1.37),

$$q_{kk}(t) = \alpha^{\frac{-2}{\alpha-1}} L_{kM}(t) \text{ and } q_{kl}(t) = \Gamma^{\frac{1}{\alpha-1}} \alpha^{\frac{-2}{\alpha-1}} L_{lM}(t) \text{ for all } i_k \in \mathbb{M}_k(t). \tag{F.2}$$

These expressions for $q_{kk}(t)$ and $q_{kl}(t)$ imply that at time t, all I-sector firms in region k produce

$$Y_{i_k}(t) = Y_{ki}(t) = q_{kk}(t) + \Gamma q_{kl}(t) = \alpha^{\frac{-2}{\alpha-1}}\left(L_{kM}(t) + \Gamma^{\frac{\alpha}{a-1}} L_{lM}(t)\right) \text{ for all } i_k \in \mathbb{M}_k(t), \tag{F.3}$$

and, make profits equal to

$$\pi_{i_k k}(t) = \pi_{ki}(t) = \left(\frac{1}{\alpha} - 1\right) Y_{ki}(t) = (1-\alpha)\, \alpha^{\frac{-\alpha-1}{\alpha-1}}\left(L_{kM}(t) + \Gamma^{\frac{\alpha}{a-1}} L_{lM}(t)\right) \text{ for all } i_k \in \mathbb{M}_k(t). \tag{F.4}$$

(F.3) and (F.4) are (1.41) and (1.42) from Proposition 1.

In order to obtain an expression for the relative price of S-sector output, $p(t)$, first note that from (1.5), the total demand for S-sector output at time t is

$$\sum_{j=1}^{H} Y_{jS}(t) = E(t) \left[p(t) + \left(\frac{\mu}{1-\mu} \right)^{\varepsilon} p(t)^{\varepsilon} \right]^{-1}.$$

Substituting this expression for the total demand for S-sector output and the expressions for S-sector employment in regions a and b from (1.26) into the S-sector market-clearing condition from (1.33) yields (1.39) from Proposition 1:

$$\frac{E(t)}{p(t) + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p(t)^{\varepsilon}} = L_{aS}(t)^{\beta} F_a(t)^{1-\beta} + L_{aS}(t)^{\beta} F_a(t)^{1-\beta} = \left(\frac{\beta p(t)}{w_a(t)} \right)^{\frac{\beta}{1-\beta}} F_a + \left(\frac{\beta p(t)}{w_a(t)} \right)^{\frac{\beta}{1-\beta}} F_b.$$

This expression defines the relative price of S-sector output, $p(t)$, implicitly in terms of $E(t)$, $w_a(t)$, and $w_b(t)$ when the economy is in static equilibrium.

The expression for the value of an I-sector firm in region k at time t (provide I-sector innovation takes place in region k at time t) in (1.40) from Proposition 1 follows directly from the zero-profit condition in the I-sector, (1.30), and uses the expressions for $\Omega_k(t)$ from (1.27) and for $w_k(t)$ from (1.37).

Substituting the expressions for $q_{kk}(t)$ and $q_{kl}(t)$ from (F.2) into the expression for M-sector production in region k at time t, (1.16), yields (1.46) from Proposition 1:

$$Y_{kM}(t) = \left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t) \right) q_{kk}(t)^{\alpha} L_{kM}(t)^{1-\alpha} = \alpha^{\frac{2\alpha}{1-\alpha}} L_{kM}(t) \left(N_k(t) + \Gamma^{\frac{\alpha}{a-1}} N_l(t) \right). \tag{F.5}$$

The expression for M-sector employment in (1.43) from Proposition 1 can be obtained from the M-sector market-clearing condition, (1.32), by using the expression for M-sector production in region k at time t from (F.5) and the fact that, from (1.5), the total demand for M-sector output at time t is

$$\sum_{j=1}^{H} Y_{jM}(t) = E(t) \left[1 + \left(\frac{1-\mu}{\mu} \right)^{\varepsilon} p(t)^{1-\varepsilon} \right]^{-1}, \tag{F.6}$$

and, from the expressions for the quantity of intermediate inputs produced in regions k and l from (1.41),

$$\sum_{i_k \in \mathbb{M}_k(t)} Y_{i_k}(t) = \sum_{i_k \in \mathbb{M}_k(t)} Y_{ki}(t) = N_k(t) Y_{ki}(t) = N_k(t) \alpha^{\frac{2}{1-\alpha}} \left(L_{kM}(t) + \Gamma^{\frac{\alpha}{a-1}} L_{lM}(t) \right), \tag{F.7}$$

for all $i_k \in \mathbb{M}_k(t)$. Substituting (F.5), (F.6), and (F.7) into (1.32) yields

$$\begin{aligned} \frac{E(t)}{1 + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p(t)^{1-\varepsilon}} &= \alpha^{\frac{2\alpha}{1-\alpha}} \left(1 - \alpha^2\right) \left[\begin{array}{c} \left(N_a(t) + \Gamma^{\frac{\alpha}{a-1}} N_b(t)\right) L_{aM}(t) \\ + \left(\Gamma^{\frac{\alpha}{a-1}} N_a(t) + N_b(t)\right) L_{bM}(t) \end{array} \right] \\ &= (1+\alpha) \left[w_a(t) L_{aM}(t) + w_b(t) L_{bM}(t) \right] \\ &= (1+\alpha) L_M(t) \left[w_a(t) \left(\frac{L_{aM}(t)}{L_M(t)} \right) + w_b(t) \left(1 - \frac{L_{aM}(t)}{L_M(t)} \right) \right]. \end{aligned} \tag{F.8}$$

Rearranging (F.8) provides the expression for M-sector employment in (1.43) from Proposition 1. The expression for S-sector employment when the economy is in static equilibrium in (1.44) from Proposition 1 follows directly from the expression for $L_{kS}(t)$, $k = a, b$, from (1.26). Similarly, the expression for total employment in I-sector innovation when the economy is in static equilibrium in (1.45) from Proposition 1 follows directly from the market-clearing condition in the labor market, (1.31).

Substituting the expression for the expression for S-sector employment in region k at time t from (1.26) into the expression for S-sector production in region k at time t from (1.23) yields the expression for total production of S-sector when the economy is in static equilibrium in (1.47) from Proposition 1. Finally, the expression for output in I-sector innovation when the economy is in static equilibrium in (1.48) from Proposition 1 is simply the restatement of (1.29) from Chapter 1.4.2. ■

F.2 Proof of Theorem 1

Proof. Theorem 1 is proven in three parts. First, it is shown that $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{1, 1\}$ is a stable spatial equilibrium and that all $\{\lambda_M(t), N_a(t)/N(t)\} \in \{(0.5, 1), (0.5, 1)\}$

are not spatial equilibria. This result is demonstrated in two steps. First, using (1.42),

$$\begin{aligned}
\lambda_M(t) &\in (0.5,1] \Rightarrow \pi_{ai}(t) = (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}}\left[\lambda_M(t) + \Gamma^{\frac{\alpha}{a-1}}(1-\lambda_M(t))\right] L_M(t) \\
&> (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}}\left[\Gamma^{\frac{\alpha}{a-1}}\lambda_M(t) + (1-\lambda_M(t))\right] L_M(t) = \pi_{bI}\ (t) \Rightarrow \frac{N_a(t)}{N(t)} = 1.
\end{aligned}$$

The first step establishes that I-sector firms are more profitable in region a than region b when region a has a larger concentration of M-sector labor. These higher profits will lead to I-sector firms concentrating themselves in the region a. Second, using (1.37),

$$\begin{aligned}
\frac{N_a(t)}{N(t)} &\in (0.5,1] \Rightarrow w_a(t) = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}\left[\frac{N_a(t)}{N(t)} + \Gamma^{\frac{\alpha}{a-1}}\left(1-\frac{N_a(t)}{N(t)}\right)\right] N(t) \\
&> (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}\left[\Gamma^{\frac{\alpha}{a-1}}\frac{N_a(t)}{N(t)} + \left(1-\frac{N_a(t)}{N(t)}\right)\right] N(t) = w_b(t) \Rightarrow \lambda_M(t) = 1.
\end{aligned}$$

The second step establishes that the wage paid to M-sector labor is higher region a than region b when region a has a larger concentration of I-sector firms. This higher wage will lead M-sector labor to concentrate itself in the region a.

These two steps establish that $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{1,1\}$ is a spatial equilibrium in that no I-sector firm increase its profits and no M-sector worker can increase their wage by moving between regions when $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{1,1\}$. $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{1,1\}$ is a stable spatial equilibrium because there exists a bounded, convex set containing $\{1,1\}$, given by $\{(1-0.5, 1+0.5), (1-0.5, 1+0.5)\} \cap \{[0,1], [0,1]\} = \{(0.5,1], (0.5,1]\}$, such that from any point $\{\lambda_M(t), N_a(t)/N(t)\} \in \{(0.5,1], (0.5,1]\}$, M-sector workers can increase their wage and I-sector firms can increase their profits by moving to region a until $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{1,1\}$ is restored as the spatial distributions of M-sector labor and I-sector firms. An immediate corollary of this result is that no $\{\lambda_M(t), N_a(t)/N(t)\} \in \{(0.5,1), (0.5,1)\}$ is a spatial equilibrium.

Second, identical arguments establish that $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{0,0\}$ is a stable spatial equilibrium; that there exists a bounded, convex set containing

$\{0, 0\}$, given by $\{[0, 0.5), [0, 0.5)\}$, such that from any point $\{\lambda_M(t), N_a(t)/N(t)\} \in \{[0, 0.5), [0, 0.5)\}$, M-sector workers can increase their wage and I-sector firms can increase their profits by moving to region b until $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{0, 0\}$ is restored as the spatial distributions of M-sector labor and I-sector firms; and that all $\{\lambda_M(t), N_a(t)/N(t)\} \in \{(0, 0.5), (0, 0.5)\}$ are not spatial equilibria.

Third, it is shown that $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{0.5, 0.5\}$ is a spatial equilibrium. This results is demonstrated in two steps. First,

$$\lambda_M(t) = 0.5 \Rightarrow \pi_{ai}(t) = \pi_{bi}(t) = (1-\alpha)\,\alpha^{\frac{1+\alpha}{1-\alpha}}\frac{1}{2}\left(1 + \Gamma^{\frac{\alpha}{a-1}}\right)L_M(t),$$

which means that when $\lambda_M(t) = 0.5$, no I-sector firm can increase its profits by moving between regions because the demand for its variety of intermediate input by M-sector firms, which determines its profits, is the same in both regions. Second,

$$\frac{N_a(t)}{N(t)} = 0.5 \Rightarrow w_a(t) = w_b(t) = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}\frac{1}{2}\left(1 + \Gamma^{\frac{\alpha}{a-1}}\right)N(t),$$

which means that when $N_a(t)/N(t) = 0.5$, no M-sector worker can increase their wage by moving between regions. Taken together, these two steps imply that $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{0.5, 0.5\}$ is a spatial equilibrium. The first two parts of this proof establish that $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{0.5, 0.5\}$ is not a stable spatial equilibrium. This is because any bounded, convex set containing $\{0.5, 0.5\}$ defined by $\{(0.5-\rho, 0.5+\rho), (0.5-\rho, 0.5+\rho)\} \cap \{[0,1], [0,1]\}$, $\rho > 0$, will contain elements of the sets $\{[0, 0.5), [0, 0.5)\}$ and $\{(0.5, 1], (0.5, 1]\}$. As such, there does not exist a bounded, convex set containing $\{0.5, 0.5\}$ defined by $\{(0.5-\rho, 0.5+\rho), (0.5-\rho, 0.5+\rho)\} \cap \{[0,1], [0,1]\}$, $\rho > 0$, such that from any point in the set, M-sector workers and I-sector firms will change regions to restore $\{\lambda_M^*(t), N_a^*(t)/N(t)\} = \{0.5, 0.5\}$ as the spatial distribution of M-sector labor and I-sector firms. ∎

F.3 Proof of Theorem 2

Proof. Theorem 2 is proven in three steps. First, recall that the free-entry condition in I-sector innovation means that entrepreneurs will develop new varieties of intermediate inputs until the value of an I-sector firm, $V(t)$, is equal to the cost of developing a new variety at time t in the region with the lowest cost of development. From (1.40), when the M-sector labor and I-sector firms are in spatial equilibrium so that $w(t) = w_a(t) = w_b(t)$,

$$V(t) = \min\left(\frac{w(t)}{\Omega_a(t)}, \frac{w(t)}{\Omega_b(t)}\right), \tag{F.9}$$

where $\Omega_a(t) = N(t)[\lambda_I(t) + \eta(1-\lambda_I(t))]$ and $\Omega_b(t) = N(t)[\eta\lambda_I(t) + (1-\lambda_I(t))]$ are the marginal products of labor in I-sector innovation in regions a and b. That $\lambda_I^*(t) = 1$ is a stable spatial equilibrium and that all $\lambda_I(t) \in (0.5, 1)$ are not spatial equilibria follows from

$$\lambda_I(t) \in (0.5, 1] \Rightarrow V(t) = \frac{w(t)}{\Omega_a(t)} < \frac{w(t)}{\Omega_b(t)} \Rightarrow \lambda_I(t) = 1.$$

$\lambda_I^*(t) = 1$ is a spatial equilibrium because no firm in I-sector innovation can increase their profits by moving between regions when $\lambda_I(t) = 1$. $\lambda_I^*(t) = 1$ is a stable spatial equilibrium because there exists a bounded, convex set containing $\lambda_I^*(t) = 1$, given by $(1-0.5, 1+0.5) \cap [0,1] = (0.5, 1]$, such that from any point in $\lambda_I(t) \in (0.5, 1]$, firms in I-sector innovation can increase their profits by moving into region a until $\lambda_I^*(t) = 1$ is restored as the spatial distribution of labor in I-sector innovation.

Second, identical arguments establish that $\lambda_I^*(t) = 0$ is a stable spatial equilibrium; that there exists a bounded, convex set containing $\lambda_I^*(t) = 0$, given by $[0, 0.5)$, such that from any point $\lambda_I(t) \in [0, 0.5)$, firms in I-sector innovation can increase their profits by moving into region b until $\lambda_I^*(t) = 0$ is restored as the spatial distribution of labor in I-sector innovation; and that all $\lambda_I(t) \in (0, 0.5)$ are not spatial equilibrium.

Third, that $\lambda_I^* (t) = 0.5$ is a spatial equilibrium follows from the fact that

$$\lambda_I (t) = 0.5 \Rightarrow V (t) = \frac{w (t)}{\Omega_a (t)} = \frac{w (t)}{\Omega_b (t)},$$

which implies that no firm in I-sector innovation can increase its profits by moving between regions when $\lambda_I (t) = 0.5$. The first two parts of this proof establish that $\lambda_I^* (t) = 0.5$ is not stable spatial equilibrium. This is because any bounded, convex set containing $\lambda_I^* (t) = 0.5$ defined by $(0.5 - \rho, 0.5 + \rho) \cap [0, 1]$, $\rho > 0$, will contain elements of the sets $[0, 0.5)$ and $(0.5, 1]$. As such, there does not exist a bounded, convex set containing $\lambda_I^* (t) = 0.5$ defined by $(0.5 - \rho, 0.5 + \rho) \cap [0, 1]$, $\rho > 0$, such that from any point in the set, I-sector firms will move between regions so that $\lambda_I^* (t) = 0.5$ is restored as the spatial distribution of labor in I-sector innovation. ■

Appendix G

Comparative Static Results

G.1 Proof of Proposition 2

Proof. To begin, the price of S-sector output, $p(t)$, is defined implicitly in terms of $E(t)$ and $N(t)$ by the S-sector market-clearing condition, (1.39). To derive an expression for the elasticity of $p(t)$ with respect to $E(t)$, $N(t)$, and $L(t)$, rewrite (1.39) as

$$\psi\left(E(t), N(t), p^{SC}(t)\right) = \frac{E(t) w^{SC}(t)^{\frac{\beta}{1-\beta}}}{\beta^{\frac{\beta}{1-\beta}}(F_a + F_b)} - \left[p^{SC}(t)^{\frac{1}{1-\beta}} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon + \frac{\beta}{1-\beta}}\right] = 0, \tag{G.1}$$

$SC = FA, FD$. Applying the implicit function theorem to (G.1) and using (1.39) yields part 1 of Proposition 2:

$$\begin{aligned}
\frac{d\ln p^{SC}(t)}{d\ln N(t)} &= \frac{E(t) w^{SC}(t)^{\frac{\beta}{1-\beta}}}{\beta^{\frac{\beta}{1-\beta}}(F_a + F_b)} \frac{\beta}{1-\beta} \frac{1}{\left[p^{SC}(t)^{\frac{1}{1-\beta}} \frac{1}{1-\beta} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon + \frac{\beta}{1-\beta}} \left(\varepsilon + \frac{\beta}{1-\beta}\right)\right]} \\
&= \frac{\beta}{1-\beta} \frac{p^{SC}(t)^{\frac{1}{1-\beta}} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon + \frac{1}{1-\beta} - 1}}{p^{SC}(t)^{\frac{1}{1-\beta}} \frac{1}{1-\beta} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon + \frac{1}{1-\beta} - 1} \left(\varepsilon + \frac{\beta}{1-\beta}\right)} \\
&= \beta \frac{1 + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon - 1}}{1 + \left[(1-\beta)\varepsilon + \beta\right] \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon - 1}} > 0.
\end{aligned} \tag{G.2}$$

Next, the expression for the elasticity of $p^{SC}(t)$ with respect to $E(t)$ in part 2 of Proposition 2 is derived by applying the implicit function theorem to (G.1) and using

(1.39):

$$\frac{d\ln p(t)}{d\ln E(t)} = \frac{p^{SC}(t)^{\frac{1}{1-\beta}} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon+\frac{\beta}{1-\beta}}}{p^{SC}(t)^{\frac{1}{1-\beta}}\frac{1}{1-\beta} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon+\frac{\beta}{1-\beta}}\left(\varepsilon+\frac{\beta}{1-\beta}\right)} = \frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} > 0. \tag{G.3}$$

Finally, from (1.39), when the economy is in static equilibrium, $p^{SC}(t)$ is determined by $N(t)$ and $E(t)$ and not $L(t)$. As such, part 3 of Proposition 2 must hold for all t. ■

G.2 Proof of Proposition 3

Proof. Using the expression for M-sector employment when the economy is in a static equilibrium from (1.43), the elasticity of M-sector employment with respect to technology, $N(t)$, is

$$\frac{d\ln L_M^{SC}(t)}{d\ln N(t)} = -\left[\frac{d\ln w^{SC}(t)}{d\ln N(t)} + \frac{\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}}(1-\varepsilon)\frac{d\ln p^{SC}(t)}{d\ln N(t)}\right] = -\left[1+\frac{(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\frac{d\ln p^{SC}(t)}{d\ln N(t)}\right], \tag{G.4}$$

$SC = FA, FD$. Substituting the expression for $\frac{d\ln p(t)}{d\ln N(t)}$ from (G.2) into (G.4) yields the first expression in part 1 of Proposition 3:

$$\frac{d\ln L_M^{SC}(t)}{d\ln N(t)} = -\left\{1+\frac{(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\beta\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\} = -\left\{\frac{(1-\beta)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\} < 0. \tag{G.5}$$

Using the expression for S-sector employment when the economy is in a static equilibrium from (1.44), the elasticity of S-sector employment, $L_S^{SC}(t)$, with respect to technology, $N(t)$, is

$$\frac{d\ln L_S^{SC}(t)}{d\ln N(t)} = \frac{d\ln L_{kS}^{SC}(t)}{d\ln N(t)} = \frac{1}{1-\beta}\left(\frac{d\ln p^{SC}(t)}{d\ln N(t)} - \frac{d\ln w^{SC}(t)}{d\ln N(t)}\right) = \frac{1}{1-\beta}\left(\frac{d\ln p^{SC}(t)}{d\ln N(t)} - 1\right). \tag{G.6}$$

$k = a, b$. Substituting the the expression for $\frac{d\ln p(t)}{d\ln N(t)}$ from (G.2) into (G.6) yields the second expression in part 1 of Proposition 3:

$$\begin{aligned} \frac{d\ln L_S^{SC}(t)}{d\ln N(t)} &= \frac{1}{1-\beta}\left\{\beta\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} - 1\right\} \\ &= \frac{1}{1-\beta}\left\{\frac{\beta\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]}{\beta\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]+(1-\beta)\left[1+\varepsilon\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]} - 1\right\} \\ &= -\left\{\frac{1+\varepsilon\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\} < 0. \end{aligned} \tag{G.7}$$

The third expression in part 1 of Proposition 3 follows from (G.5) and (G.7) and the market-clearing condition in the labor market, (1.31).

Using (1.43), the elasticity of M-sector employment with respect to societal expenditure, $E(t)$, is given by

$$\frac{d\ln L_M^{SC}(t)}{d\ln E(t)} = 1 - \frac{\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}}(1-\varepsilon)\frac{d\ln p^{SC}(t)}{d\ln E(t)} = 1 - \frac{(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\frac{d\ln p^{SC}(t)}{d\ln E(t)}. \tag{G.8}$$

Substituting the expression for $\frac{d\ln p^{SC}(t)}{d\ln E(t)}$ from (G.3) into (G.8) yields the first expression in part 2 of Proposition 3:

$$\begin{aligned} \frac{d\ln L_M^{SC}(t)}{d\ln E(t)} &= 1 - \frac{(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{\left\{1+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right\}} \\ &= 1 - \frac{(1-\varepsilon)}{1+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} \\ &= \frac{\varepsilon+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} > 0. \end{aligned} \tag{G.9}$$

Using (1.44), the elasticity of S-sector employment with respect to $E(t)$ is given by

$$\frac{d\ln L_S^{SC}(t)}{d\ln E(t)} = \frac{d\ln L_{kS}^{SC}(t)}{d\ln E(t)} = \frac{1}{1-\beta}\left(\frac{d\ln p^{SC}(t)}{d\ln E(t)} - \frac{d\ln w^{SC}(t)}{d\ln E(t)}\right) = \frac{1}{1-\beta}\frac{d\ln p^{SC}(t)}{d\ln E(t)}. \tag{G.10}$$

Using the result that $\frac{d\ln p^{SC}(t)}{d\ln E(t)} > 0$ from Proposition 2, it follows that (G.10) is greater than zero, which establishes the second expression in part 2 of Proposition 3. The third expression in part 1 of Proposition 3 follows from (G.9), (G.10), and the market-clearing condition in the labor market, (1.31).

Finally, from (1.43) and (1.44), when the economy is in static equilibrium, $L_M^{SC}(t)$ and $L_S^{SC}(t)$ are determined by $N(t)$ and $E(t)$ and not $L(t)$. As such, part 5 of Proposition 3 must hold for all t. ■

G.3 Proof of Proposition 4

Proof. To derive part 1 of Proposition 4, the expression for $\frac{d\ln L_M^{SC}(t)}{d\ln N(t)}$ from (G.5) and the expression for $\frac{d\ln L_S^{SC}(t)}{d\ln N(t)}$ from (G.7) imply

$$\begin{aligned}\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln N(t)} &= -\left[1+\frac{(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}\frac{d\ln p^{SC}(t)}{d\ln N(t)}\right] \\ &\quad -\frac{1}{1-\beta}\left(\frac{d\ln p^{SC}(t)}{d\ln N(t)}-1\right) \\ &= \frac{\beta}{1-\beta}-\left\{\frac{(1-\varepsilon)(1-\beta)+\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\frac{d\ln p^{SC}(t)}{d\ln N(t)},\end{aligned} \tag{G.11}$$

$SC = FA, FD$. Substituting the expression for $\frac{d\ln p^{SC}(t)}{d\ln N(t)}$ from (G.2) into (G.11), and the fact that $\frac{d\ln p^{SC}(t)}{d\ln N(t)} > 0$ from Proposition 2, yields part 1 of Proposition 2:

$$\begin{aligned}
\frac{d\ln\left(L_M^{SC}(t)/L_S^{SC}(t)\right)}{d\ln N(t)} &= \left\{\begin{array}{c} \frac{\beta}{1-\beta}\left(\frac{d\ln p^{SC}(t)}{d\ln N(t)}\right)^{-1} \\ -\frac{(1-\varepsilon)(1-\beta)+\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\end{array}\right\}\frac{d\ln p^{SC}(t)}{d\ln N(t)} \qquad \text{(G.12)}\\
&= \left\{\begin{array}{c} \frac{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]} \\ -\frac{(1-\varepsilon)(1-\beta)+\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\end{array}\right\}\frac{d\ln p^{SC}(t)}{d\ln N(t)}\\
&= \left\{\frac{\begin{array}{c}1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\\ -(1-\varepsilon)(1-\beta)-\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]\end{array}}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\frac{d\ln p^{SC}(t)}{d\ln N(t)}\\
&= \left\{\frac{\begin{array}{c}[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]\\ -\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]\end{array}}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\frac{d\ln p^{SC}(t)}{d\ln N(t)}\\
&= \left\{\frac{-(1-\varepsilon)(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\frac{d\ln p^{SC}(t)}{d\ln N(t)}\\
&= (\varepsilon-1)\frac{d\ln p^{SC}(t)}{d\ln N(t)} > 0 \Leftrightarrow \varepsilon > 1.
\end{aligned}$$

Next, to derive part 2 of Proposition 4, the expression for $\frac{d\ln L_M^{SC}(t)}{d\ln E(t)}$ from (G.9) and the expression for $\frac{d\ln L_S^{SC}(t)}{d\ln E(t)}$ from (G.10) imply,

$$\begin{aligned}
\frac{d\ln\left(L_M^{SC}(t)/L_S(t)\right)}{d\ln E(t)} &= \frac{d\ln L_M^{SC}(t)}{d\ln E(t)} - \frac{d\ln L_S^{SC}(t)}{d\ln E(t)} \qquad \text{(G.13)}\\
&= 1-\frac{(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}\frac{d\ln p^{SC}(t)}{d\ln E(t)} - \frac{1}{1-\beta}\frac{d\ln p^{SC}(t)}{d\ln E(t)}\\
&= 1-\left\{\frac{(1-\varepsilon)(1-\beta)+\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\frac{d\ln p^{SC}(t)}{d\ln E(t)}.
\end{aligned}$$

Substituting the expression for $\frac{d\ln p^{SC}(t)}{d\ln E(t)}$ from (G.3) into (G.13), and the fact that $\frac{d\ln p^{SC}(t)}{d\ln E(t)} > 0$ from Proposition 2, yield part 2 of Proposition 2:

$$\begin{aligned}
\frac{d\ln\left(L_M^{SC}(t)/L_S(t)\right)}{d\ln E(t)} &= 1-\left\{\begin{array}{c}\frac{(1-\varepsilon)(1-\beta)+\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\\ \times\frac{\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}\end{array}\right\} \qquad \text{(G.14)}\\
&= 1-\frac{(1-\varepsilon)(1-\beta)+\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\beta)\left\{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right\}}\\
&= 1-\frac{\left\{\begin{array}{c}(1-\varepsilon)(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]\\ +\left\{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right\}\end{array}\right\}}{(1-\beta)\left\{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right\}}\\
&= -\frac{\left\{\begin{array}{c}\beta\left\{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right\}\\ +(1-\varepsilon)(1-\beta)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]\end{array}\right\}}{(1-\beta)\left\{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right\}}\\
&= -\left\{\frac{\beta}{(1-\beta)}+(1-\varepsilon)\frac{\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}\right\}\\
&= -\left[\frac{\beta}{(1-\beta)}+(1-\varepsilon)\frac{d\ln p^{SC}(t)}{d\ln E(t)}\right],
\end{aligned}$$

where

$$-\left[\frac{\beta}{(1-\beta)}+(1-\varepsilon)\frac{d\ln p^{SC}(t)}{d\ln E(t)}\right] > 0 \Leftrightarrow \varepsilon > \left[1+\frac{\beta}{(1-\beta)}\left(\frac{d\ln p^{SC}(t)}{d\ln E(t)}\right)^{-1}\right] > 1.$$

■

G.4 Proof of Proposition 5

Proof. Under full agglomeration, region a's total employment is given by

$$L_a^{FA}(t) = L_M^{FA}(t) + L_{aS}^{FA}(t) + L_I^{FA}(t).$$

Using the market-clearing condition for labor, (1.31), $L_a^{FA}(t)$ can be re-expressed as

$$L_a^{FA}(t) = L_M^{FA}(t) + L_{aS}^{FA}(t) + \left(L(t) - L_M^{FA}(t) - L_{aS}^{FA}(t) - L_{bS}^{FA}(t)\right) = L(t) - L_{bS}^{FA}(t) \tag{G.15}$$

The elasticity of $L_a^{FA}(t)$ with respect to technology, $N(t)$, is

$$\frac{d\ln L_a^{FA}(t)}{d\ln N(t)} = -\frac{L_{bS}^{FA}(t)}{L(t) - L_{bS}^{FA}(t)}\frac{d\ln L_{bS}^{FA}(t)}{d\ln N(t)} = -\frac{L_{bS}^{FA}(t)}{L(t) - L_{bS}^{FA}(t)}\frac{d\ln L_S^{SC}(t)}{d\ln N(t)}. \tag{G.16}$$

Substituting the expression for $\frac{d\ln L_S^{SC}(t)}{d\ln N(t)}$ from (G.7) into (G.16) yields the first term in part 1 of Proposition 5:

$$\frac{d\ln L_a^{FA}(t)}{d\ln N(t)} = \frac{L_{bS}^{FA}(t)}{L(t) - L_{bS}^{FA}(t)}\left\{\frac{1+\varepsilon\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+\left[(1-\beta)\varepsilon+\beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\} > 0. \tag{G.17}$$

Similarly, the elasticity of $L_a^{FA}(t)$ with respect to societal expenditure, $E(t)$, is

$$\frac{d\ln L_a^{FA}(t)}{d\ln E(t)} = -\frac{L_{bS}^{FA}(t)}{L(t) - L_{bS}^{FA}(t)}\frac{d\ln L_{bS}^{FA}(t)}{d\ln E(t)} = -\frac{L_{bS}^{FA}(t)}{L(t) - L_{bS}^{FA}(t)}\frac{d\ln L_S^{FA}(t)}{d\ln E(t)} < 0. \tag{G.18}$$

(G.18) uses the result that $\frac{d\ln L_S^{FA}(t)}{d\ln E(t)} < 0$ from Proposition 3 (G.18) establishes the second term from part 1 of Proposition 5:

$$\frac{d\ln L_a^{FA}(t)}{d\ln E(t)} = -\frac{L_{bS}^{FA}(t)}{L(t) - L_{bS}^{FA}(t)}\frac{1}{1-\beta}\frac{d\ln p^{FA}(t)}{d\ln E(t)} < 0. \tag{G.19}$$

The third term from part 1 of Proposition 5 follows directly from (G.15).

Under full dispersion, the region a's share of total employment is given by

$$L_a^{FD}(t) = \frac{1}{2}L_M^{FD}(t) + L_{aS}^{FD}(t) + \frac{1}{2}L_I^{FD}(t).$$

Using the market-clearing condition for labor, (1.31), and the expression for $L_{kS}^{FD}(t)$, $k = a, b$, from (1.26), $L_a^{FD}(t)$ can be re-expressed as

$$\begin{aligned} L_a^{FA}(t) &= \frac{1}{2}L_M^{FD}(t) + L_{aS}^{FD}(t) + \frac{1}{2}\left(L(t) - L_M^{FD}(t) - L_{aS}^{FD}(t) - L_{bS}^{FD}(t)\right) \\ &= \frac{1}{2}L(t) + \frac{1}{2}\left(L_{aS}^{FD}(t) - L_{bS}^{FD}(t)\right) \\ &= \frac{1}{2}L(t) + \frac{1}{2}L_S^{FD}(t)\left(\frac{F_a - F_b}{F_a + F_b}\right) \end{aligned} \tag{G.20}$$

Using this expression for $L_a^{FD}(t)$ and the fact that $\frac{d\ln L_S^{FD}(t)}{d\ln N(t)} < 0$ from Proposition 3, the elasticity of $L_a^{FD}(t)$ with respect to $N(t)$ is

$$\frac{d\ln L_a^{FA}(t)}{d\ln N(t)} = \frac{1}{2}\frac{L_S^{FD}(t)}{L_a^{FA}(t)}\left(\frac{F_a - F_b}{F_a + F_b}\right)\frac{d\ln L_S^{FD}(t)}{d\ln N(t)} > 0 \Longleftrightarrow F_a - F_b < 0. \tag{G.21}$$

(G.21) establishes the first term in part 2 of Proposition 5.

Similarly, the elasticity $L_a^{FD}(t)$ with respect to $E(t)$ is

$$\frac{d\ln L_a^{FA}(t)}{d\ln E(t)} = \frac{1}{2}\frac{L_S^{FD}(t)}{L_a^{FA}(t)}\left(\frac{F_a - F_b}{F_a + F_b}\right)\frac{d\ln L_S^{FD}(t)}{d\ln E(t)} > 0 \Longleftrightarrow F_a - F_b > 0. \tag{G.22}$$

(G.22) uses the result that $\frac{d\ln L_S^{FD}(t)}{d\ln E(t)} > 0$ from Proposition 3. (G.22) establishes the second term in part 2 of Proposition 5. The third term from part 2 of Proposition 5 follows directly from (G.20). ∎

Appendix H

Dynamic Equilibrium Results

H.1 Proof of Proposition 6

Proof. In Section 1.4.1, the optimal expenditure path for household j is characterized by the Euler equation, (1.13), the budget-flow constraint, (1.6), the initial conditions $l_j(0)$, $a_j(0)$, f_{ja}, and f_{jb}, and the transversality condition, (1.12). These first-order conditions are shown to be both necessary and sufficient to identify a unique maximum for household j's intertemporal program. This proof for Proposition 6 will demonstrate that having these first-order conditions hold for each household in the economy is a sufficient condition for (1.59) - (1.61) to hold. In doing so, this proof establishes that (1.59) - (1.61) are necessary conditions for the first-order conditions for household j's intertemporal program to hold for every household in the economy.

The first differential equation in Proposition 6, (1.58), governs the growth of technology. To understand how (1.58) is a necessary condition for the first-order conditions from household j's intertemporal program to hold for every household in the economy, first note that when the economy is in static equilibrium, the equation governing how household j's assets evolve over time, (1.6), can be expressed as

$$\dot{a}_j(t) = w^{SC}(t)l_j(t) + w^{F,SC}(t)\left(f_{ja} + f_{jb}\right) + r^{SC}(t)\,a_j(t) - E_j(t), \tag{H.1}$$

$SC = FA, FD$. (H.1) uses the fact when the economy is in static equilibrium

$w_j(t) = w^{SC}(t)$ for all $j = 1, ..., H$, where $w^{SC}(t)$ is defined in (1.37), and $w_a^{F,SC}(t) = w_b^{F,SC}(t) = w^{F,SC}(t)$, where $w_j^{F,SC}(t)$, $j = a, b$, is defined in (1.38). That (H.1) holds for every household in the economy implies

$$\begin{aligned}
\dot{a}^{SC}(t) &= \sum_{j=1}^{H} \dot{a}_j(t) \qquad \text{(H.2)} \\
&= \sum_{j=1}^{H} \left[w^{SC}(t) l_j(t) + w^{F,SC}(t)(f_{ja} + f_{jb}) + r^{SC}(t) a_j(t) - E_j(t) \right] \\
&= w^{SC}(t) \sum_{j=1}^{H} l_j(t) + w^{F,SC}(t) \sum_{j=1}^{H} (f_{ja} + f_{jb}) + r^{SC}(t) \sum_{j=1}^{H} a_j(t) - \sum_{j=1}^{H} E_j(t) \\
&= w^{SC}(t) L(t) + w^{F,SC}(t)(F_a + F_b) + r^{SC}(t) a^{SC}(t) - E(t),
\end{aligned}$$

where $a^{SC}(t)$ is the total value of assets in the economy at time t.

(H.2) is an economy-wide resource constraint. To understand why this is the case, first note that because households invest in I-sector firms and the economy is closed, $a^{SC}(t)$ is equal to the total value of I-sector firms at time t, i.e., $a^{SC}(t) = V^{SC} N(t)$ (recall from (1.40) that value of an I-sector firm is constant). This implies that $\dot{a}^{SC}(t) = V^{SC} \dot{N}^{SC}(t)$, so that the growth in the total value of household assets is a linear function of the growth in the number of I-sector firms (i.e., of the growth in the number of varieties of intermediate inputs). Second, aggregate income in the economy has three parts: labor income, $w^{SC}(t) L(t)$, rents to the fixed factor, $w_a^{F,SC}(t)(F_a + F_b)$, and total profits of I-sector firms, $\pi_I^{SC}(t) N(t)$, which, using (1.42) and (1.57), equal $r^{SC}(t) a^{SC}(t)$. As such, the first three terms on the right-hand side of (H.2) are aggregate income in the economy at time t. Given these facts, (H.2) states that, at every point in time, aggregate income in the economy must be allocated either to expenditure, $E(t)$, which is directed towards household consumption of M- and S-sector output, or to investment in the creation of new varieties of intermediate inputs.

Next, it will be demonstrated that (H.2) is equivalent to (1.58) when the expressions for $\dot{a}^{SC}(t)$, $a^{SC}(t)$, $w^{SC}(t)$, $w^{F,SC}(t)$, and $r^{SC}(t)$ when the economy is in static equilibrium are inserted into (H.2). Using the expression for $w^{F,SC}(t)$ from (1.38) and the expression for $L_S^{SC}(t)$ from (1.44), we obtain

$$\begin{aligned}
w^{F,SC}(t)\left(F_a+F_b\right) &= (1-\beta)\,p^{SC}(t)\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{\frac{\beta}{1-\beta}}\left(F_a+F_b\right) \qquad\qquad \text{(H.3)}\\
&= (1-\beta)\,p^{SC}(t)\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{-1}\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{\frac{1}{1-\beta}}\left(F_a+F_b\right)\\
&= (1-\beta)\,\beta^{-1}w^{SC}(t)\,L_S^{SC}(t)\,.
\end{aligned}$$

Next, using the expressions $r^{FA}(t)=\alpha L_M^{FA}(t)$ and $r^{FD}(t)=\alpha\frac{1}{2}(1+\eta)\,L_M^{FD}(t)$ from (1.57), the expressions $V^{FA}=(1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}$ and $V^{FD}=(1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}\left(1+\Gamma^{\frac{\alpha}{a-1}}\right)(1+\eta)^{-1}$ from (1.40), and the fact that $a^{SC}(t)=V^{SC}N(t)$, yields

$$r^{SC}(t)\,a^{SC}(t)=\alpha w^{SC}(t)\,L_M^{SC}. \qquad \text{(H.4)}$$

Substituting (H.3) and (H.4) into (H.2), and using the fact that $\dot{a}^{SC}(t)=V^{SC}\dot{N}^{SC}(t)$,

$$\begin{aligned}
V^{SC}\dot{N}^{SC}(t) &= w^{SC}(t)L(t)+(1-\beta)\,\beta^{-1}w^{SC}(t)\,L_S^{SC}(t)+\alpha w^{SC}(t)\,L_M^{SC}-E(t)\\
&= w^{SC}(t)\left(L(t)-L_M^{SC}(t)-L_S^{SC}(t)\right) \qquad\qquad \text{(H.5)}\\
&\quad +w^{SC}(t)\left[\beta^{-1}w^{SC}(t)\,L_S^{SC}(t)+(1+\alpha)\,w^{SC}(t)\,L_M^{SC}(t)-E(t)\right]\\
&= w^{SC}(t)\left(L(t)-L_M^{SC}(t)-L_S^{SC}(t)\right).
\end{aligned}$$

Substituting the expressions for $L_M^{SC}(t)$ and $L_S^{SC}(t)$ from (1.43) and (1.44) into (H.5) yields (1.58). This establishes that (H.2) is equivalent to (1.58), and that having (1.6) hold for each household in the economy is a sufficient condition for (1.58) to hold. In addition, this proof demonstrates that expressions for $\hat{N}^{FA}(t)$ and $\hat{N}^{FD}(t)$ in (1.58) are consistent with the aggregate resource constraint for the economy holding for all t.[1]

[1] In addition, notice that (1.58) is the expression for the production function in the I-sector innovation, (1.29), using the market-clearing condition for labor (1.31), and substituting in the expressions for $L_M^{SC}(t)$ and $L_S^{SC}(t)$ from (1.43) and (1.44). This equivalence is reassuring, and also not surprising given that the production function for I-sector innovation is embedded in the economy-wide resource constraint implied by (H.2).

Notice that (H.5) holds because

$$\beta^{-1}w^{SC}(t)\,L_S^{SC}(t)+(1+\alpha)\,w^{SC}(t)\,L_M^{SC}-E(t)=0 \tag{H.6}$$

for all values of $E(t)$, $N(t)$, and $L(t)$. To establish this result, the first term in (H.6) can be re-expressed as, using the expression for $L_S^{SC}(t)$ from (1.44),

$$\begin{aligned}\beta^{-1}w^{SC}(t)\,L_S^{SC}(t) &= \beta^{-1}w^{SC}(t)\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{\frac{1}{1-\beta}}(F_a+F_b) \qquad \text{(H.7)}\\ &= \beta^{-1}w^{SC}(t)\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{\frac{\beta}{1-\beta}}(F_a+F_b)\\ &= p^{SC}(t)\left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{\frac{\beta}{1-\beta}}(F_a+F_b).\end{aligned}$$

Further, using the S-sector market-clearing condition from (1.39), (H.7) becomes

$$\beta^{-1}w^{SC}(t)\,L_S^{SC}(t)=p^{SC}(t)\frac{E(t)}{p^{SC}(t)+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon}}=\frac{E(t)}{1+\left[\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{SC}(t)^{1-\varepsilon}\right]^{-1}} \tag{H.8}$$

Similarly, using the M-sector market-clearing condition from (1.32), the second term in (H.6) can be expressed as,

$$(1+\alpha)\,w^{SC}(t)\,L_M^{SC}(t)=\frac{E(t)}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{SC}(t)^{1-\varepsilon}}. \tag{H.9}$$

Substituting (H.8) and (H.9) into (H.6),

$$\left\{\frac{1}{1+\left[\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{SC}(t)^{1-\varepsilon}\right]^{-1}}+\frac{1}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon}p^{SC}(t)^{1-\varepsilon}}\right\}E(t)-E(t)=E(t)-E(t)=0.$$

Thus, (H.6) holds for all values of $E(t)$, $N(t)$, and $L(t)$.[2]

[2] To understant why the above expressoin holds, note that

$$\frac{1}{1+a^{-1}}+\frac{1}{1+a}=\frac{1+a+1+a^{-1}}{(1+a^{-1})(1+a)}=\frac{1+a^{-1}+1+a}{1+a^{-1}+1+a}=1.$$

The second differential equation in Proposition 6, (1.59), governs the evolution of total societal expenditure, $E(t)$. As the Euler equation from household j's program, (1.13), must hold for every household in the economy, the following ordinary differential equation must hold for total societal expenditure:

$$\hat{E}^{SC}(t) = \frac{\sum_{j=1}^{H} \dot{E}_j(t)}{\sum_{j=1}^{H} E_j(t)} = \frac{\left(\sum_{j=1}^{H} E_j(t)\right)\left[r^{SC}(t) - (\delta - n)\right]}{\sum_{j=1}^{H} E_j(t)} = r^{SC}(t) - (\delta - n) \quad \text{(H.10)}$$

$SC = FA, FD$. Using the expression for $r^{FA}(t) = \alpha L_M^{FA}(t)$ and $r^{FD}(t) = \frac{1}{2}[1+\eta]\,\alpha L_M^{FD}(t)$ from (1.57) and the expression for $L_M^{SC}(t)$ from (1.43) yields (1.59). This implies that a sufficient condition for (1.59) is for the Euler equation, (1.13), to hold for every household in the economy.

The third differential equation in Proposition 6, (1.60), governs the population growth. As (1.1), which governs growth in household size, must hold for every household in the economy, population growth in the economy is $\hat{L}(t) = n$.

Finally, it is left to demonstrate that having each household's transversality condition, (1.15), hold is a sufficient condition for (1.61) to hold. To begin, that the transversality condition in (1.15) must hold for every household in the economy implies

$$\begin{aligned}\sum_{j=1}^{H} \lim_{t\to\infty} a_j(t) \exp\left(-\int_0^t r(\tau)d\tau\right) &= \lim_{t\to\infty}\left(\sum_{j=1}^{H} a_j(t)\right)\exp\left(-\int_0^t r(\tau)d\tau\right) \quad \text{(H.11)}\\ &= \lim_{t\to\infty} a^{SC}(t)\exp\left(-\int_0^t r(\tau)d\tau\right) = 0.\end{aligned}$$

As is explained above, $a^{SC}(t) = V^{SC}N(t)$ at all time t, which implies $\hat{a}(t) = \hat{N}^{SC}(t)$.

Integrating $\hat{a}(t) = \hat{N}^{SC}(t)$ with respect to time,

$$\begin{aligned}
\dot{a}(s) - a(s)\hat{N}^{SC}(s) &= 0 \qquad \text{(H.12)} \\
\frac{d}{ds}\left[a(s)\exp\left(-\int_0^s \hat{N}^{SC}(\tau)d\tau\right) + k_0\right] &= 0 \\
\int_0^t \frac{d}{ds}\left[a(s)\exp\left(-\int_0^s \hat{N}^{SC}(\tau)d\tau\right) + k_0\right] ds &= 0 \\
\left[a(s)\exp\left(-\int_0^s \hat{N}^{SC}(\tau)d\tau\right) + k_0\right]_0^t &= 0 \\
a(t)\exp\left(-\int_0^t \hat{N}^{SC}(\tau)d\tau\right) - a(0) &= 0.
\end{aligned}$$

(H.12) can be re-expressed as

$$a(t) = a(0)\exp\left(\int_0^t \hat{N}^{SC}(\tau)\,d\tau\right). \qquad \text{(H.13)}$$

Substituting (H.13) into (H.11) yields (1.61). This establishes that having (1.15) hold for every household in the economy is a sufficient condition for (1.61) to hold. This establishes that (1.61) is a necessary condition for each household's transversality condition to hold. ■

H.2 Proof of Theorem 3

Proof. The EGP is defined as a dynamic equilibrium for the economy that features a constant real interest rate and constant growth in expenditure. To begin, from (1.57), a constant real interest rate, $\bar{r}^{SC}$, implies that M-sector employment is also constant along the EGP. Using the expression for $\bar{r}^{SC}$, $SC = FA, FD$, from (1.57) and the expressions for $\hat{E}^{SC}(t)$ from (1.59),

$$g_E^{FA} = \alpha\bar{L}_M^{SC} - (\delta - n) \text{ and } g_E^{FD} = \frac{1}{2}[1+\eta]\,\alpha\bar{L}_M^{FD} - (\delta - n). \qquad \text{(H.14)}$$

Next, the asymptotic growth rates for technology, $N(t)$, and the relative price of S-sector output, $p(t)$, along the EGP are established using the S-sector market clearing condition, (1.39), and the fact that M-sector employment is constant along the EGP. To begin, the expression for $L_M^{SC}(t)$ from (1.43) and the fact that M-sector employment is constant along the EGP imply that

$$\hat{E}^{SC}(t) - \hat{N}(t) - (1-\varepsilon)\frac{1}{\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]}\hat{p}^{SC}(t) = 0 \tag{H.15}$$

must hold along the EGP. Similarly, totally differentiating the S-sector market-clearing condition (1.39) implies that

$$\begin{aligned}(1-\beta)\hat{E}^{SC}(t) + \beta\hat{N}(t) - \left\{\frac{p^{SC}(t)^{\frac{1}{1-\beta}} + [(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon+\frac{1}{1-\beta}-1}}{p^{SC}(t)^{\frac{1}{1-\beta}} + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon+\frac{1}{1-\beta}-1}}\right\}\hat{p}^{SC}(t) &= 0 \\ (1-\beta)\hat{E}^{SC}(t) + \beta\hat{N}(t) - \left\{\frac{1 + [(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1 + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\}\hat{p}^{SC}(t) &= 0\end{aligned}$$

must hold along the EGP. Substituting (H.15) into this expression yields

$$\begin{array}{c}(1-\beta)\hat{E}^{SC}(t) + \beta\left[\hat{E}^{SC}(t) - (1-\varepsilon)\frac{1}{\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]}\hat{p}^{SC}(t)\right] \\ -\left(\frac{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right)\hat{p}^{SC}(t)\end{array} = 0.$$

This implies that the following relationship between $\hat{E}^{SC}(t)$ and $\hat{p}^{SC}(t)$ must hold along the EGP:

$$\begin{aligned}\hat{E}^{SC}(t) &= \left\{\frac{\beta(1-\varepsilon)}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} + \frac{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\}\hat{p}^{SC}(t) \\ &= \left\{\frac{(1-\varepsilon) + [(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right]}{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}\right\}\hat{p}^{SC}(t). \end{aligned} \tag{H.16}$$

Similarly, substituting (H.16) into (H.15) implies that the following relationship between $\hat{E}^{SC}(t)$ and $\hat{N}^{SC}(t)$ must hold along the EGP:

$$\begin{aligned}
\hat{N}^{SC}(t) &= \hat{E}^{SC}(t) - (1-\varepsilon)\frac{1}{\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\hat{p}^{SC}(t) \qquad \text{(H.17)}\\
&= \hat{E}^{SC}(t) - \left\{\begin{array}{c}(1-\varepsilon)\frac{1}{\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\\ \times\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\end{array}\right\}\hat{E}^{SC}(t)\\
&= \hat{E}^{SC}(t) - \left\{\frac{(1-\varepsilon)}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\hat{E}^{SC}(t)\\
&= \left\{\frac{[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\hat{E}^{SC}(t).
\end{aligned}$$

The fact that $\hat{E}^{SC}(t) = g_E^{SC}$ along the EGP and L'Hôpital's rule are used to get expressions for $\hat{p}^{SC}(t)$ and $\hat{N}^{SC}(t)$ along the EGP. For example, applying L'Hôpital's Rule and using (H.16),

$$\begin{aligned}
g_p^{SC} &= \lim_{t\to\infty}\left\{\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}g_E^{SC} \qquad \text{(H.18)}\\
&= \lim_{t\to\infty}\left\{\frac{(\varepsilon-1)\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-2}\dot{p}^{SC}(t)}{(\varepsilon-1)[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-2}\dot{p}^{SC}(t)}\right\}g_E^{SC}\\
&= [(1-\beta)\varepsilon+\beta]^{-1}g_E^{SC}.
\end{aligned}$$

Similarly, applying L'Hôpital's Rule and using (H.17),

$$\begin{aligned}
g_N^{SC} &= \lim_{t\to\infty}\left\{\frac{[\varepsilon(1-\beta)+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\varepsilon)+[\varepsilon(1-\beta)+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}g_E^{SC} \qquad \text{(H.19)}\\
&= \lim_{t\to\infty}\left\{\frac{(\varepsilon-1)[\varepsilon(1-\beta)+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-2}\dot{p}^{SC}(t)}{(\varepsilon-1)[\varepsilon(1-\beta)+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-2}\dot{p}^{SC}(t)}\right\}g_E^{SC}\\
&= g_E^{SC}.
\end{aligned}$$

Using the expression for M-sector production from (1.46),

$$\hat{Y}_M^{SC}(t) = \hat{L}_M^{SC}(t) + \hat{N}^{SC}(t). \qquad \text{(H.20)}$$

Together, (H.20) and the fact that M-sector employment is constant along the EGP imply

$$g_M^{SC} = \lim_{t\to\infty} \hat{Y}_M^{SC}(t) = \lim_{t\to\infty} \hat{N}^{SC}(t) = g_N^{SC} = g_E^{SC}. \tag{H.21}$$

Similarly, using the expression for S-sector production from (1.47) and the expression for S-sector employment from (1.44),

$$Y_{kS}^{SC}(t) = \left(\frac{\beta p^{SC}(t)}{w^{SC}(t)}\right)^{\frac{\beta}{1-\beta}} F_k, \tag{H.22}$$

$k = a, b$. (H.22) implies

$$\hat{Y}_{kS}^{SC}(t) = \hat{Y}_S^{SC}(t) = \frac{\beta}{1-\beta}\left(\hat{p}^{SC}(t) - \hat{N}^{SC}(t)\right), \tag{H.23}$$

Using the fact that (H.18) and (H.17) imply that $\hat{N}(t) = [(1-\beta)\varepsilon + \beta]\,\hat{p}^{SC}(t)$, (H.23) becomes

$$\hat{Y}_{kS}^{SC}(t) = \hat{Y}_S^{SC}(t) = \frac{\beta}{1-\beta}\left(\frac{1}{(1-\beta)\varepsilon+\beta} - 1\right)\hat{N}^{SC}(t) = \left(\frac{(1-\varepsilon)\beta}{(1-\beta)\varepsilon+\beta}\right)\hat{N}^{SC}(t). \tag{H.24}$$

(H.24) implies that

$$g_S^{SC} = \lim_{t\to\infty} \hat{Y}_S^{SC}(t) = \left(\frac{(1-\varepsilon)\beta}{(1-\beta)\varepsilon+\beta}\right) g_N^{SC} = \left(\frac{(1-\varepsilon)\beta}{(1-\varepsilon)\beta+\varepsilon}\right) g_E^{SC}. \tag{H.25}$$

(H.25) implies that S-sector output is growing along the EGP ($g_S^{SC} > 0$) if and only if $\varepsilon < 1$. In addition, (H.25) implies that S-sector output grows at a lower rate (g_S^{SC}) than M-sector output ($g_M^{SC} = g_E^{SC}$) along the EGP.

Next, the asymptotic growth rates for S-sector employment, $L_S^{SC}(t)$, is derived by using the condition that marginal revenue product of labor in the S-sector must equal the wage for all t, which, using (1.47), is given by

$$\beta p^{SC}(t)\, L_S^{SC}(t)^{\beta-1} F_k^{1-\beta} = w^{SC}(t) \tag{H.26}$$

$k = a, b$, $SC = FA, FD$, where, from (1.37), $w^{FA}(t) = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}} N(t)$ and $w^{FD}(t) = (1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}} \frac{1}{2}\left[1 + \Gamma^{\frac{\alpha}{a-1}}\right] N(t)$, and $p^{SC}(t)$ is defined implicitly by the

S-sector market-clearing condition (1.39). (H.26) implies that

$$\hat{L}_S^{SC}(t) = (1-\beta)^{-1}\left(\hat{p}^{SC}(t) - \hat{N}^{SC}(t)\right). \tag{H.27}$$

Using (H.16) and (H.17), (H.27) becomes

$$\begin{aligned}\hat{L}_S^{SC}(t) &= (1-\beta)^{-1}\left\{\begin{array}{c}\left\{\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\} \\ -\left\{\frac{[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\end{array}\right\}\hat{E}^{SC}(t) \\ &= \left\{\frac{(1-\varepsilon)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}\hat{E}^{SC}(t).\end{aligned} \tag{H.28}$$

Applying L'Hôpital's Rule and using (H.28),

$$\begin{aligned} n_S^{SC} &= \lim_{t\to\infty}\left\{\frac{(1-\varepsilon)\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}{(1-\varepsilon)+[(1-\beta)\varepsilon+\beta]\left[1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-1}\right]}\right\}g_E^{SC} \\ &= \lim_{t\to\infty}\left\{\frac{(1-\varepsilon)\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-2}\dot{p}^{SC}(t)}{[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon}p^{SC}(t)^{\varepsilon-2}\dot{p}^{SC}(t)}\right\}g_E^{SC} \\ &= \left[\frac{(1-\varepsilon)}{(1-\beta)\varepsilon+\beta}\right]g_E^{SC} = \left[\frac{(1-\varepsilon)}{(1-\varepsilon)\beta+\varepsilon}\right]g_E^{SC}.\end{aligned} \tag{H.29}$$

(H.29) is greater than zero if and only if $\varepsilon < 1$.

Next, from (1.29), $\hat{N}^{FA}(t) = L_I^{FA}(t)$ and $\hat{N}^{FD}(t) = \frac{1}{2}(1+\eta)L_I^{FD}(t)$. These expression imply, using (H.19), that I-sector employment is constant along the EGP,

$$\bar{L}_I^{FA} \equiv \lim_{t\to\infty} L_I^{FA}(t) = \lim_{t\to\infty}\hat{N}^{SC}(t) = g_E^{FA} \tag{H.30}$$

and

$$\bar{L}_I^{FD} \equiv \lim_{t\to\infty} L_I^{FD}(t) = \lim_{t\to\infty}\left[\frac{1}{2}(1+\eta)\right]^{-1}\hat{N}^{FD}(t) = \left[\frac{1}{2}(1+\eta)\right]^{-1}g_E^{FD}. \tag{H.31}$$

That both M- and I-sector employment are constant along the EGP implies that for the EGP to be consistent with a growing population (i.e., $n > 0$), S-sector employment must be growing along the EGP. (H.29) establishes that $n_S^{SC} > 0$ if and only if $\varepsilon < 1$. Hence, an EGP with a constant real interest rate and constant growth in expenditure is only consistent with the conditions for a dynamic equilibrium for the economy if $\varepsilon < 1$. This is the basis for the claim in Theorem 3 that there does not exist an EGP when $\varepsilon \geq 1$.

That both M- and I-sector employment are constant along the EGP implies that the asymptotic rate of growth in S-sector employment equals the rate of population growth. (i.e., $n_S^{SC} = n$). From (H.29), this result implies that

$$n_S^{SC} = n = \left(\frac{(1-\varepsilon)}{(1-\varepsilon)\beta+\varepsilon}\right) g_E^{SC} \rightarrow g_E^{SC} = \left(\frac{(1-\varepsilon)\beta+\varepsilon}{(1-\varepsilon)}\right) n \tag{H.32}$$

(1.62)-(1.66) in Theorem 3 follow directly from (H.32). In particular, (H.32) along with the result that $g_E^{SC} = g_N^{SC} = g_M^{SC}$ establishes (1.62). Next, substituting the expression for g_E^{SC} from (H.32) into the expression for S-sector production along the EGP, (H.25), establishes the second term in (1.63).

Concerning prices, substituting the expression for g_E^{SC} from (H.32) into the expression for the growth of $p^{SC}(t)$ from (H.18) yields the first term in (1.64). The second term in (1.64) follows from (1.37).The third term in (1.64) follows from (1.38), the expression for g_p^{SC}, and the fact that, from (H.18) and (H.17), $\hat{N}(t) = [(1-\beta)\varepsilon+\beta]\,\hat{p}^{SC}(t)$. The fourth and final term in (1.64) follows from (H.32) and (1.36).

Substituting the expression for g_E^{SC} from (H.32) into (H.14) and rearranging terms establishes the expression for the constant level of M-sector employment along the EGP in (1.65). Similarly, substituting the expression for g_E^{SC} from (H.32) into (H.30) and (H.31) establishes the expressions for the constant level of I-sector employment along the EGP in (1.66).

Concerning regional population growth and regional population shares along the EGP, the populations of regions a and b are, under full agglomeration,

$$L_a^{FA}(t) = L_M^{FA}(t) + L_I^{FA}(t) + L_{aS}^{FA}(t) \text{ and } L_b^{FA}(t) = L_{bS}^{FA}(t), \tag{H.33}$$

and, under full dispersion,

$$L_a^{FD}(t) = \frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right) + L_{aS}^{FD}(t) \text{ and } L_b^{FD}(t) = \frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right) + L_{bS}^{FD}(t). \tag{H.34}$$

It follows from (H.33) and (H.34) that

$$n_a^{FA} = \lim_{t\to\infty} \hat{L}_a^{FA}(t) = \lim_{t\to\infty} \left(\frac{L_{aS}^{FA}(t)}{L_a^{FA}(t)}\right) \hat{L}_{aS}^{FA}(t) = \lim_{t\to\infty} \hat{L}_S^{FA}(t) = n. \tag{H.35}$$

(H.35) uses the facts that $\hat{L}_{kS}^{FA}(t) = \hat{L}_S^{FA}(t)$, $k = a, b$, which follows from the expression for $L_{kS}^{FA}(t)$ from (1.26), and that both M- and I-sector employment are constant along the EGP, which implies that $\lim_{t\to\infty}\left(L_{aS}^{FA}(t)/L_a^{FA}(t)\right) = 1$. Identical arguments establish that the remaining terms in (1.67) hold (i.e., that $n_b^{FA} = n_a^{FD} = n_b^{FD} = n$). Concerning regional population shares, (H.33) and (H.34) imply that, under full agglomeration,

$$\overline{\frac{L_a^{FA}(t)}{L_b^{FA}(t)}} = \lim_{t\to\infty} \frac{L_a^{FA}(t)}{L_b^{FA}(t)} = \lim_{t\to\infty}\left[\frac{L_M^{FA}(t) + L_I^{FA}(t)}{L_{bS}^{FA}(t)} + \frac{L_{aS}^{FA}(t)}{L_{aS}^{FA}(t)}\right] = 0 + \frac{F_a}{F_b} = \frac{F_a}{F_b},$$

and, under full dispersion,

$$\begin{aligned}
\overline{\frac{L_a^{FD}(t)}{L_b^{FD}(t)}} &= \lim_{t\to\infty} \frac{L_a^{FD}(t)}{L_b^{FD}(t)} \\
&= \lim_{t\to\infty}\left[\frac{\frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right)}{\frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right) + L_{bS}^{FA}(t)} + \frac{L_{aS}^{FA}(t)}{\frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right) + L_{bS}^{FA}(t)}\right] \\
&= 0 + \frac{F_a}{F_b} = \frac{F_a}{F_b}.
\end{aligned}$$

These result establish (1.68).

To demonstrate that this EGP allocation is consistent with the necessary conditions for a dynamic equilibrium from Proposition 6, it is left to establish that it satisfies the transversality condition given in (1.61) (note that this proof has already used (1.58), (1.59), and (1.60) from Proposition 6). A sufficient condition for the transversality condition in (1.61) to hold is for the rate of growth of technology, $\hat{N}(t)$, to be less than the real interest rate, $r(t)$, as $t \to \infty$. This condition holds along the EGP: $\bar{r}^{SC} = \alpha \bar{L}_M^{SC} = g_E + (\delta - n) > g_E = g_N$, $SC = FA, FD$. Hence, the transversality condition in (1.61) from Proposition 6 holds along the EGP.

The EGP is unique by construction in that it is the only dynamic path for the economy with a constant real interest rate and constant growth rate of expenditure that satisfies the necessary conditions for a dynamic equilibrium from Proposition 6.

■

H.3 Proof of Proposition 7

Proof. To establish Proposition 7, it is sufficient to demonstrate that $L_M^{SC}(t)$ tends to a constant as $t \to \infty$ when (i) the economy is on the dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ and (ii) $\hat{L}_S^{FA}(t)$ approaches n from above as $t \to \infty$. Recall that the EGP is defined as a dynamic equilibrium for the economy that features a constant real interest rate (as well as constant expenditure growth). From (1.57), a constant $L_M^{SC}(t)$ implies a constant $r^{SC}(t)$.

To obtain an expression for $L_M^{SC}(t)$ when the economy is on a dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$, using the expression for $\hat{N}^{SC}(t)$ from (1.58) and the expression for $\hat{E}^{SC}(t)$ from (1.58),

$$\hat{N}^{FA}(t) = L(t) - L_M^{FA}(t) - L_S^{FA}(t) = \alpha L_M^{FA}(t) - (\delta - n) = \hat{E}^{FA}(t)$$

and

$$\hat{N}^{FD}(t) = \frac{1}{2}(1+\eta)\left(L(t) - L_M^{FD}(t) - L_S^{FD}(t)\right) = \frac{1}{2}(1+\eta)\alpha L_M^{FD}(t) - (\delta - n) = \hat{E}^{FD}(t).$$

These two expressions imply that when $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$,

$$L_M^{FA}(t) = (1+\alpha)^{-1}\left[L(t) - L_S^{FA}(t) + (\delta - n)\right] \tag{H.36}$$

and

$$L_M^{FD}(t) = (1+\alpha)^{-1}\left[L(t) - L_S^{FD}(t) + \frac{2}{(1+\eta)}(\delta - n)\right]. \tag{H.37}$$

From (H.36) and (H.37), $L_M^{FA}(t)$ and $L_M^{FD}(t)$ will tend to a constant as $t \to \infty$ if $\left(L(t) - L_S^{SC}(t)\right)$ tends to constant as $t \to \infty$. When $\hat{L}_S^{FA}(t)$ approaches n from above as $t \to \infty$ then $\lim_{t\to\infty} \left(L(t) - L_S^{SC}(t)\right) = \lim_{t\to\infty} L(t)\left(1 - \frac{L_S^{SC}(t)}{L(t)}\right) = 0$ because $\lim_{t\to\infty} L_S^{SC}(t)/L(t) = 1$ (i.e., because the S-sector dominates the asymptotic distribution of employment. Hence, that $\hat{L}_S^{FA}(t)$ approaches n from above as $t \to \infty$ is a sufficient condition for (H.36) and (H.37) to tend to constants as $t \to \infty$, and for the dynamic trajectory defined by $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ to approach the EGP. ■

H.4 Proof of Proposition 8

Proof. To begin, totally differentiating the expression for M-sector employment from (1.43) with respect to time yields

$$\hat{L}_M^{SC}(t) = \hat{E}^{SC}(t) - \hat{N}^{SC}(t) - \frac{(1-\varepsilon)}{1 + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} \hat{p}^{SC}(t), \tag{H.38}$$

$SC = FA, FD$. The expression for $\hat{p}^{SC}(t)$ is obtained by totally differentiating the S-sector market clearing condition (1.39) with respect to time:

$$\hat{p}^{SC}(t) = \frac{1 + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1 + \left[(1-\beta)\varepsilon + \beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} \left[(1-\beta)\left(\hat{E}^{SC}(t) - \hat{N}^{SC}(t)\right) + \hat{N}^{SC}(t)\right]. \tag{H.39}$$

When $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, (H.38) and (H.39) yield part 1 of Proposition 8:

$$\hat{L}_M^{SC}(t) = \frac{(\varepsilon - 1)}{1 + \left[(1-\beta)\varepsilon + \beta\right]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} \hat{N}^{SC}(t) > 0 \Leftrightarrow \varepsilon > 1. \tag{H.40}$$

Next, using the expression for S-sector employment from (1.44) and the S-sector market clearing condition from (1.39),

$$L_S^{SC}(t) = \frac{\beta E(t)}{w^{SC}(t)} \frac{1}{\left(1 + \left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}\right)} \tag{H.41}$$

Totally differentiating (H.41) with respect to time yields

$$\hat{L}_S^{SC}(t) = \hat{E}^{SC}(t) - \hat{N}^{SC}(t) + \frac{(1-\varepsilon)}{1+\left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}} \hat{p}^{SC}(t). \tag{H.42}$$

When $\hat{N}^{SC}(t) = \hat{E}^{SC}(t)$, substituting the expression for $\hat{p}(t)^{SC}$ from (H.39) into (H.42) yields part 2 of Proposition 8:

$$\hat{L}_S^{SC}(t) = \frac{(1-\varepsilon)}{[(1-\beta)\varepsilon+\beta] + \left(\frac{1-\mu}{\mu}\right)^{\varepsilon} p^{SC}(t)^{1-\varepsilon}} \hat{N}^{SC}(t) > 0 \Leftrightarrow \varepsilon < 1. \tag{H.43}$$

Part 3 of Proposition 8 can be derived using the expression for M-sector production from (1.46), and using the expression for $\hat{L}_M^{SC}(t)$ when $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ from (H.40)

$$\begin{aligned}
\hat{Y}_M^{SC}(t) &= \hat{L}_M^{SC}(t) + \hat{N}^{SC}(t) \\
&= \left(\frac{(\varepsilon-1)}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} + 1 \right) \hat{N}^{SC}(t) \\
&= \left(\frac{\varepsilon+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} \right) \hat{N}^{SC}(t) > 0.
\end{aligned}$$

Part 4 of Proposition 8 can be derived using the expression for S-sector production from (1.47), the expression for S-sector employment from (1.44), the fact that $\hat{Y}_{kS}^{SC}(t) = \hat{Y}_S^{SC}(t)$, $k = a, b$, and the expression for $\hat{p}^{SC}(t)$ when $\hat{E}^{SC}(t) = \hat{N}^{SC}(t)$ from (H.39),

$$\begin{aligned}
\hat{Y}_S^{SC}(t) &= \frac{\beta}{1-\beta}\left(\hat{p}^{SC}(t) - \hat{N}^{SC}(t)\right) \\
&= \frac{\beta}{1-\beta}\left(\frac{1+\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} - 1 \right) \hat{N}^{SC}(t) \\
&= \frac{\beta}{1-\beta}\left(\frac{(1-\beta)(1-\varepsilon)\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}}{1+[(1-\beta)\varepsilon+\beta]\left(\frac{\mu}{1-\mu}\right)^{\varepsilon} p^{SC}(t)^{\varepsilon-1}} \right) \hat{N}^{SC}(t) > 0 \Leftrightarrow \varepsilon < 1.
\end{aligned}$$

Finally, part 5 of Proposition 8 follows from totally differentiating the market clearing condition in the labor market with respect to time,

$$\hat{L}_I^{SC}(t) = \left[nL(t) - \left(\dot{L}_M^{SC}(t) + \dot{L}_S^{SC}(t) \right) \right] L_I^{SC}(t)^{-1}. \tag{H.44}$$

and from the the fact that $\hat{Y}_I^{SC}(t) = \hat{L}_I^{SC}(t)$. ■

H.5 Proof of Proposition 9

Proof. To being, under full agglomeration, employment in regions a and b are

$$\begin{aligned} L_a^{FA}(t) &= L_M^{FA}(t) + L_I^{FA}(t) + L_{aS}^{FA}(t) = L(t) - L_{bS}^{FA}(t) = L(t) - L_S^{FA}(t)\left(\frac{F_b}{F_a+F_b}\right) \text{ and} \\ L_b^{FA}(t) &= L_{bS}^{FA}(t) = L_S^{FA}(t)\left(\frac{F_b}{F_a+F_b}\right). \end{aligned}$$

Totally differentiating these expressions for $L_a^{FA}(t)$ and $L_a^{FA}(t)$ with respect to time and rearranging terms yields part 1 of Proposition 9:

$$\hat{L}_a^{FA}(t) - \hat{L}_b^{FA}(t) = \frac{L(t)}{L_S^{FA}(t)}\frac{F_a+F_b}{F_b}\left(n - \hat{L}_S^{FA}(t)\right) > 0 \Leftrightarrow n > \hat{L}_S^{FA}(t).$$

Similarly, under full dispersion, employment in regions a and b are

$$\begin{aligned} L_a^{FD}(t) &= \frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right) + L_{aS}^{FD}(t) = \frac{1}{2}L(t) + \frac{1}{2}L_S^{FD}(t)\left(\frac{F_a-F_b}{F_a+F_b}\right) \text{ and} \\ L_b^{FD}(t) &= \frac{1}{2}\left(L_M^{FD}(t) + L_I^{FD}(t)\right) + L_{bS}^{FD}(t) = \frac{1}{2}L(t) - \frac{1}{2}L_S^{FD}(t)\left(\frac{F_a-F_b}{F_a+F_b}\right). \end{aligned}$$

Totally differentiating these expressions for $L_a^{FD}(t)$ and $L_a^{FD}(t)$ with respect to time yields

$$\begin{aligned} \hat{L}_a^{FD}(t) &= \frac{1}{2}\frac{1}{L_a^{FA}(t)}\left[nL(t) + \hat{L}_S^{FD}(t)\,L_S^{FD}(t)\left(\frac{F_a-F_b}{F_a+F_b}\right)\right] \text{ and} \\ \hat{L}_b^{FD}(t) &= \frac{1}{2}\frac{1}{L_b^{FD}(t)}\left[nL(t) - \hat{L}_S^{FD}(t)\,L_S^{FD}(t)\left(\frac{F_a-F_b}{F_a+F_b}\right)\right]. \end{aligned}$$

Rearranging terms yields part 2 of Proposition 9:

$$
\begin{aligned}
\hat{L}_a^{FD}(t) - \hat{L}_b^{FD}(t) &= \frac{1}{2}\frac{1}{L_a^{FA}(t)\,L_b^{FD}(t)}\frac{1}{F_a + F_b}\begin{bmatrix} nL(t)(F_a + F_b)\,L_b^{FD}(t) \\ +\hat{L}_S^{FD}(t)\,L_S^{FD}(t)(F_a - F_b)\,L_b^{FD}(t) \\ -nL(t)(F_a + F_b)\,L_a^{FD}(t) \\ +\hat{L}_S^{FD}(t)\,L_S^{FD}(t)(F_a - F_b)\,L_a^{FD}(t) \end{bmatrix} \\
&= \frac{1}{2}\frac{1}{L_a^{FA}(t)\,L_b^{FD}(t)}\frac{1}{F_a + F_b}\begin{bmatrix} nL(t)(F_a + F_b)\left(L_b^{FD}(t) - L_a^{FD}(t)\right) \\ +\hat{L}_S^{FD}(t)\,L_S^{FD}(t)(F_a - F_b)\,L(t) \end{bmatrix} \\
&= \frac{1}{2}\frac{1}{L_a^{FA}(t)\,L_b^{FD}(t)}\frac{1}{F_a + F_b}\begin{bmatrix} -nL(t)\,L_S^{FD}(t)(F_a - F_b) \\ +\hat{L}_S^{FD}(t)\,L_S^{FD}(t)(F_a - F_b)\,L(t) \end{bmatrix} \\
&= \frac{1}{2}\frac{L(t)\,L_S^{FD}(t)}{L_a^{FA}(t)\,L_b^{FD}(t)}\frac{F_a - F_b}{F_a + F_b}\left(\hat{L}_S^{FD}(t) - n\right),
\end{aligned}
$$

where

$$
\frac{1}{2}\frac{L(t)\,L_S^{FD}(t)}{L_a^{FA}(t)\,L_b^{FD}(t)}\frac{F_a - F_b}{F_a + F_b}\left(\hat{L}_S^{FD}(t) - n\right) > 0 \Leftrightarrow (F_a - F_b)\left(\hat{L}_S^{FD}(t) - n\right) > 0.
$$

■

Appendix I

Appendix for Chapter 3

We prove that

$$U^m > U^p \text{ for some } B \text{ if and only if } U^m > U^p \text{ for all } B. \tag{I.1}$$

To prove (I.1), we show that $U^m > U^p$ for all B if and only if

$$\frac{\mu_{11}^2}{4\gamma_1} > \left\{ \sum_{i=1}^{2} \frac{\left[E\left(P_{e_i}\right)\right]^2 + Var\left(P_{e_i}\right)}{4\gamma_i} \right\} \left(\beta^p\right)^2, \tag{I.2}$$

which we demonstrate in three steps. First, it follows directly from the expressions for U^m and U^p that $U^m > U^p$ for all $B \geq \max\{B^m, B^p\}$ if and only if (I.2) holds. Second, from (3.27), when $B < \min\{B^m, B^p\}$,

$$U^m = U^{nr} + \left[\pi\left(e_1^{nr}, e_2^{nr}\right) + B\right] \left[\sqrt{\frac{\mu_{11}^2}{4\gamma_1}} - \sqrt{\pi\left(e_1^{nr}, e_2^{nr}\right) + B} \right]. \tag{I.3}$$

It follows from (3.40) that when $B < \min\{B^m, B^p\}$,

$$\begin{aligned} U^p = {} & U^{nr} + \left[\pi\left(e_1^{nr}, e_2^{nr}\right) + B\right] \\ & \times \left[\sqrt{\left\{ \sum_{i=1}^{2} \frac{\left[E\left(P_{e_i}\right)\right]^2 + Var\left(P_{e_i}\right)}{4\gamma_i} \right\} \left(\beta^p\right)^2} - \sqrt{\pi\left(e_1^{nr}, e_2^{nr}\right) + B} \right]. \end{aligned} \tag{I.4}$$

Using (I.3) and (I.4), we obtain that $U^m > U^p$ for all $B < \min\{B^m, B^p\}$ if and only if (I.2) holds. Third, it is left to demonstrate that $U^m > U^p$ for all $B \in [\min\{B^m, B^p\}, \max\{B^m, B^p\})$ if and only if (I.2) holds. It follows from the expressions for B^m and B^p that $B^m > B^p$ if and only if (I.2) holds. By continuity of U^m

and U^p as functions of B and equivalence between (I.2) and $B^m > B^p$, we have that $U^m > U^p$ for all $B \in [\min\{B^m, B^p\}, \max\{B^m, B^p\})$ if and only if $B^m > B^p$. Hence, $U^m > U^p$ for all $B \in [\min\{B^m, B^p\}, \max\{B^m, B^p\})$ if and only if (I.2) holds.

Combining the results proved in the above three steps, we obtain that $U^m > U^p$ for all B if and only if (I.2) holds. The inequality (I.2) is independent of B. Hence, $U^m > U^p$ for some B if and only if $U^m > U^p$ for all B.

BIBLIOGRAPHY

[1] Abdel-Rahman, H. and M. Fujita. "Product Variety, Marshallian Externalities, and City Sizes." *Journal of Regional Science*, 1990, 30(2), pp. 165-83.

[2] Acemoglu, D. and V. Guerrieri. "Capital Deepening and Non-Balanced Economic Growth." National Bureau of Economic Research, Inc, NBER Working Papers: 12475, 2006.

[3] ____. "Capital Deepening and Nonbalanced Economic Growth." *Journal of Political Economy*, 2008, 116(3), pp. 467-98.

[4] Anthon, S., P. Bogetoft, and B. Jellesmark Thorsen. "Socially Optimal Procurement with Tight Budgets and Rationing." *Journal of Public Economics*, 2007, 91(7-8), pp. 1625-42.

[5] Anthon, S., S. Garcia, and A. Stenger. "Incentive Contracts for Natura 2000 Implementation in Forest Areas." *Environmental and Resource Economics,* 2010, 46(3), pp. 281-302.

[6] Arrow, K.J. and G. Debreu. "Existence of an Equilibrium for a Competitive Economy." *Econometrica*, 1954, 22(3), pp. 265-90.

[7] Attfield, C.L.F. and J.R.W. Temple. 2006. Balanced growth and the great ratios: new evidence for the US and UK. Centre for Growth and Business Cycle Research discussion paper no. 75. http://www.socialsciences.manchester.ac.uk/cgbcr/dpcgbcr/dpcgbcr75.pdf

[8] Baker, G.P. "Incentive Contracts and Performance Measurement." *Journal of Political Economy*, 1992, 100(3), pp. 598-614.

[9] Baldwin, R.E. and R. Forslid. "The Core-Periphery Model and Endogenous Growth: Stabilizing and Destabilizing Integration." *Economica*, 2000, 67(267), pp. 307-24.

[10] Barro, R.J. and X. Sala-i-Martin. *Economic Growth*. Second edition. Cambridge and London: MIT Press, 2004.

[11] Baumol, W.J. "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis." *American Economic Review*, 1967, 57, pp. 415-26.

[12] Baumol, W.J., S.A.B. Blackman, and E.N. Wolff. "Unbalanced Growth Revisited: Asymptotic Stagnancy and New Evidence." *American Economic Review*, 1985, 75(4), pp. 806-17.

[13] Beaudry, P. and E. van Wincoop. "The Intertemporal Elasticity of Substitution: An Exploration Using a Us Panel of State Data." *Economica*, 1996, 63(251), pp. 495-512.

[14] Becker, C.M., E.S. Mills, and J.G. Williamson. "Modeling Indian Migration and City Growth, 1960-2000." *Economic Development and Cultural Change*, 1986, 35(1), pp. 1-33.

[15] Bergstrand, J.H. "Structural Determinants of Real Exchange Rates and National Price Levels: Some Empirical Evidence." *American Economic Review*, 1991, 81(1), pp. 325-34.

[16] Berliant, M., P. Wang, R. Capello, and P. Nijkamp. "Dynamic Urban Models: Agglomeration and Growth," *Urban Dynamics and Growth: Advances in Urban*

Economics. Contributions to Economic Analysis, vol. 266. Amsterdam, San Diego and Oxford: Elsevier, 2004, 533-81.

[17] ____. "Urban Evolution in the USA." *Journal of Economic Geography*, 2003, 3(4), pp. 343-72.

[18] Billings, W.D. "Bromus Tectorum, A Biotic Cause of Ecosystem Impoverishment in the Great Basin." In *The Earth in Transition: Patterns and Processes of Biotic Impoverishment*. G.M. Woodwell ed. New York: Cambridge University Press, 1990, pp. 301-22.

[19] Blanchard, O. J. and L.F. Katz. "Regional Evolutions." *Brookings Papers on Economic Activity*, 1992, 0(1), pp. 1-61.

[20] Bontems, P., and J. Bourgeon. "Optimal Environmental Taxation and Enforcement Policy." *European Economic Review*, 2005, 49(2), pp. 409-435.

[21] Bontems, P., G. Rotillon, and N. Turpin. "Self-Selecting Agri-environmental Policies with an Application to the Don Watershed." *Environmental and Resource Economics*, 2005, 31(3), pp. 275-301.

[22] Bontems, P., and A. Thomas. "Regulating Nitrogen Pollution with Risk Averse Farmers under Hidden Information and Moral Hazard." *American Journal of Agricultural Economics*, 2006, 88:1, pp. 57-72.

[23] Bosworth, B.P., and J.E. Triplett. "The early 21st century U.S. productivity expansion is still in services." *International Productivity Monitor*, 2007, 14 (Spring), pp. 3–19.

[24] Brown, J., J. Angerer, J.W. Stuth, and R. Blaisdell. "Soil Carbon Dynamics in the Southwest: Effects of Changes in Land Use and Management. Phase I Final Report of the Southwest Regional Partnership on Carbon Sequestration."

Accessed on June 18th, 2010, from the Southwest Carbon Partnership Web Site: http://southwestcarbonpartnership.org.

[25] Buera, F.J. and J.P. Kaboski. "Can Traditional Theories of Structural Change Fit the Data?" *Journal of the European Economic Association*. 2009, 7(2-3), pp. 469-477.

[26] Caputo, M.R. *Foundations of Dynamic Economic Analysis: Optimal Control Theory and Applications*. Cambridge, United Kingdom: Cambridge University Press, 2005.

[27] Cartigny, P. and P. Michel. "On a sufficient transversality condition for infinite horizon optimal control problems." *Automatica*, 2003, 39(6), pp. 1007-10.

[28] Caselli, F. and W.J. Coleman II. "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy*, 2001, 109(3), pp. 584-616.

[29] Chambers, R.G. *Applied Production Analysis: A Dual Approach*. Cambridge, New York and Melbourne: Cambridge University Press, 1988.

[30] Chenery, H.B. "Patterns of Industrial Growth." *American Economic Review*, 1960, 50, pp. 624-54.

[31] Cheshire, P.C. and S. Magrini. "Population Growth in European Cities: Weather Matters–but Only Nationally." *Regional Studies*, 2006, 40(1), pp. 23-37.

[32] Christenson, L.R., D. Cummings, and D.W. Jorgenson. 1980. "Ecomomic Growth, 1947-1973: An International Comparison." In J.W. Kendrick and B. Vaccara (Eds.), *New Developments in Productivity Measurement and Analysis*, NBER Conference Report. Chicago: University of Chicago Press.

[33] Clark, C. *The Conditions of Economic Progress.* London: Macmillan, 1940.

[34] Clemente, J., A. Montanes, and M. Ponz. "Are the Consumption/Output and Investment/Output Ratios Stationary? An International Analysis." *Applied Economic Letters*, 1999, 6(1), pp. 687-691.

[35] Combes, P-P. "Economic Structure and Local Growth: France, 1984-1993." *Journal of Urban Economics*, 2000, 47(3), pp. 329-55.

[36] Cory, D.C. and W.E. Martin. "Valuing Wildlife for Efficient Multiple Use: Elk Versus Cattle." *Western Journal of Agricultural Economics*, 1985, 10(2), pp. 282-93.

[37] Davis, J.C. and Henderson, J.V. "Evidence on the Political Economy of the Urbanization Process." *Journal of Urban Economics*, 2003, 53(1), pp. 98-125.

[38] Dennis, B.N., and T.B. Iscan, T.B. "Engel versus Baumol: Accounting for structural change using two centuries of US data." *Explorations in Economic History*, 2009, 46(2), pp. 186-202.

[39] Dixit, A.K. and J.E. Stiglitz. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review*, 1977, 67(3), pp. 297-308.

[40] Duranton, G. "Labor Specialization, Transport Costs, and City Size." *Journal of Regional Science*, 1998, 38(4), pp. 553-73.

[41] ____. "Urban Evolutions: The Fast, the Slow, and the Still." *American Economic Review*, 2007, 97(1), pp. 197-221.

[42] Duranton, G. and D. Puga. " Micro-Foundations of Urban Agglomeration Economies," *Handbook of Regional and Urban Economics.* Elsevier, 2004, 2063-117.

[43] Ellison, G. and E.L. Glaeser. "Geographic Concentration in U S Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*, 1997, 105(5), pp. 889-927.

[44] Eltis, W. "How Low Profitability and Weak Innovativeness Undermined Uk Industrial Growth." *The Economic Journal*, 1996, 106(434), pp. 184-95.

[45] Engle, S., S. Pagiola, and S. Wunder. "Designing Payments for Environmental Services in Theory and Practice: An Overview of the Issues." *Ecological Economics*, 2008, 65(4), pp. 663-74.

[46] Feng, H. 2007. "Green Payments and Dual Policy Goals." *Journal of Environmental Economics and Management*, 54:3, pp. 323-335.

[47] Flavey, R.E. and N. Gemmell. "Are Service Income-Elastic? Some New Evidence." *Review of Income and Wealth*, 1996, 42(3), pp. 257-69.

[48] ____. "Explaining Service-Price Differences in International Comparisons." *American Economic Review*, 1991, 81(5), pp. 1295-309.

[49] Foellmi, R. and J. Zweimuller. "Structural change, Engel's consumption cycles and Kaldor's facts of economic growth." *Journal of Monetary Economics*, 2008, 55(7), pp.1371-1328.

[50] Follett, R.F., J.M. Kimble, and R. Lal eds. *The potential for U.S. grazing lands to sequester carbon and mitigate the greenhouse effect*. Boca Raton, FL, USA: CRC Press, 2001.

[51] Fowler, J.M., L.A. Torell, and G. Gallacher. "Competitive Pricing for the McGregor Range: Implications for Federal Grazing Fees." *Journal of Range Management*, 1994, 47(2), pp. 155-58.

[52] Fujita, M. and J-F. Thisse. "Does Geographical Agglomeration Foster Economic Growth? And Who Gains and Loses from It?" *Japanese Economic Review*, 2003, 54(2), pp. 121-45.

[53] ____. *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth.* Cambridge: Cambridge University Press, 2002.

[54] Gabaix, X. "Zipf's Law and the Growth of Cities." *American Economic Review*, 1999, 89(2), pp. 129-32.

[55] Gautier, A. "Regulation Under Financial Constraints." *Annals of Public and Cooperative Economics*, 2004, 75(4), pp. 645-56.

[56] Glaeser, Ed.L. and M.E. Kahn. "Decentralized Employment and the Transformation of the American City." *Brookings-Wharton Papers on Urban Affairs*, 2001, 0(0), pp. 1-47.

[57] Glaeser, E.L., J.A. Scheinkman, and A. Shleifer. "Economic Growth in a Cross-Section of Cities." *Journal of Monetary Economics*, 1995, 36(1), pp. 117-43.

[58] Glaeser, E.L. and J.M. Shapiro. "Urban Growth in the 1990s: Is City Living Back?" *Journal of Regional Science*, 2003, 43(1), pp. 139-65.

[59] Glaeser, E.L. and K. Tobio. "The Rise of the Sunbelt," National Bureau of Economic Research, Inc, NBER Working Papers: 13071, 2007.

[60] Green, H.A.J. *Aggregation in Economic Analysis: An Introductory Survey.* Princeton, New Jersey: Princeton University Press, 1964.

[61] Gollin, D., S. Parente, and R. Rogerson. "The Role of Agriculture in Development." *American Economic Review*, 2002, 92(2), pp. 160-64.

[62] Gomme, P., and P. Rupert. "Theory, measurement and calibration of macroeconomic models." Journal of Monetary Economics, 2007, 54(2), pp. 460–497.

[63] Gorman, W. M. "Community Preference Fields." *Econometrica*, 1953, 21(1), pp. 63-80.

[64] Guillo, M.D., C. Papageorgiou, and F. Perez-Sebastian. "A Unified Theory of Structural Change." Journal of Economic Dynamics and Control. 2011, 35(9), pp. 1393-1404.

[65] Harris, C.D. "The Market as a Factor in the Localization of Production " *Annals of the Association of American Geographers*, 1954, 44, pp. 315-48.

[66] Havranek, T., R. Hovath, Z. Irsova, and M. Rusnak. "Cross-Country Heterogeneity in Intertemporal Substitution." William Davidson Institute Working Paper Number 1056. August 2013.

[67] Havstad, K.M., D.P.C. Peters, R. Skaggs, J. Brown, B. Bestelmeyer, E. Fredrickson, J. Herrick, and J. Wright. "Ecological services to and from rangelands of the United States." *Ecological Economics*, 2007, 64:2, pp. 261-68.

[68] Helsley, Robert W. and W.C. Strange. "Matching and Agglomeration Economies in a System of Cities." *Regional Science and Urban Economics*, 1990, 20(2), pp. 189-212.

[69] Henderson, J.V. "Urbanization and Growth," P. Aghion and S. N. Durlauf, *Handbook of Economic Growth.* Amsterdam: North-Holland, 2006, 1543-91.

[70] Hess, K. and J.L. Holechek. "Policy Roots of Land Degradation in the Arid Region of the United States: An Overview." *Environmental Monitoring and Assessment*, 1995, 37:1, pp. 123-41.

[71] Holechek, J.L. 1988. "An Approach for Setting the Stocking Rate." *Rangelands*, 10:1, pp. 10-14.

[72] Holmstrom, B. and P. Milgrom. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization*, 1991, 7, pp. 24-52.

[73] Hodrick, R.J., and E.C. Prescott. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit, and Banking*, 1997, 29(1), pp. 1-16.

[74] Houthakker, H.S. "An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel's Law " *Econometrica*, 1957, 25(4), pp. 532-51.

[75] Houthakker, H.S. and L. Taylor. *Consumer Demand in the United States: Analysis and Projections*. Cambridge: Harvard University Press, 1970.

[76] Hueth, B. and T. Melkonyan. "Standards and the Regulation of Environmental Risk." *Journal of Regulatory Economics*, 2009, 36(3), pp. 219-46.

[77] Huffaker, R.G., J.E. Wilen, and B.D. Gardner. "Multiple Use Benefits on Public Rangelands: An Incentive-Based Fee System." *American Journal of Agricultural Economics*, 1989, 71(3), pp. 670-78.

[78] Iscan, T.B. "How Much Can Engel's Law and Baumol's Disease Explain the Rise of Service Employment in the United States?" *B.E. Journal of Macroeconomics*, 2010, 10(1).

[79] Innes, R.D. "Environmental Standards and Liability with Limited Assets and Private Information." In *Designing Institutions for Environmental Management*, ed. E.T. Loehman and D.M. Kilgour, Cheltenham, U.K. - Northampton, MA, USA: Edward Elgar. 1998, pp.209-230.

[80] Jaffe, A.B., M. Trajtenberg, and R. Henderson. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics*, 1993, 108(3), pp. 577-98.

[81] Johnson, R.N. and M.J. Watts. "Contractual Stipulations, Resource Use, and Interest Groups: Implications from Federal Grazing Contracts." *Journal of Environmental Economics and Management*, 1989, 16(1), pp. 87-96.

[82] Jorgenson, D.W. and K.J. Stiroh. "U.S. Economic Growth at the Industry Level." *American Economic Review*, 2000, 90(2), pp. 161-67.

[83] Jorgenson, D. and E. Yip. 2001. "Whatever Happened to Productivity Growth?" In E.R. Dean, M.J. Harper, and C. Hulten (Eds.), *New Developments in Productivity Analysis*, p. 205-246. Chicago: University of Chicago Press.

[84] Judd, K. *Numerical Methods in Economics*. Cambridge: MIT Press, 1998.

[85] Junius, K. "Primacy and Economic Development: Bell Shaped or Parallel Growth of Cities?" *Journal of Economic Development*, 1999, 24(1), pp. 1-22.

[86] Kamihigashi, T. "Necessity of Transversality Conditions for Infinite Horizon Problems." *Econometrica*, 2001, 69(4), pp. 995-1012.

[87] Kapur, B.K. "Progressive services, asymptotically stagnant services, and manufacturing: Growth and structural change." *Journal of Economic Dynamics and Contro*l, 2012, 36(9), pp. 1322-1339.

[88] Keith, J.E. and K.S. Lyon. "Valuing Wildlife Management: A Utah Deer Herd." *Western Journal of Agricultural Economics*, 1985, 10(2), pp. 216-22.

[89] King, R.G., C.I. Plosser, J.H. Stock, and M.W. Watson. "Stochastic Trends and Economic Fluctuations." *American Economic Review*. 1991, 81(4), pp. 819-840.

[90] Knapp, P.A. "Cheatgrass (Bromus tectorum L) Dominance in the Great Basin Desert: History, Persistence, and Influences to Human Activities." *Global Environmental Change*, 1996, 6(1), pp. 37-52.

[91] Kolko, J. "Can I Get Some Service Here? Information Technology, Services Industries, and the Future of Cities." Mimeo, Harvard University, 1999.

[92] Kolstad, C.D., T.S. Ulen, and G.V. Johnson. "Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?" *The American Economic Review*, 1990, 80(4), pp. 888-901.

[93] Kongsamut, P, S. Rebelo, and D. Xie. "Beyond Balanced Growth." National Bureau of Economic Research, Inc, NBER Working Papers: 6159, 1997.

[94] ____. "Beyond Balanced Growth." *Review of Economic Studies*, 2001, 68(4), pp. 869-82.

[95] Kravis, I.B., A.W. Heston, and R. Summers. "The Share of Services in Economic Growth," In *Global Econometrics: Essays in Honor of Lawrence R. Klein*, ed. F. G. Adams and B. G. Hickman, 188-218. Cambridge, MA, and London: MIT Press, 1983.

[96] Krugman, P. "*First Nature, Second Nature, and Metropolitan Location.*" Journal of Regional Science, 1993, 33(2), pp. 129-44.

[97] ____. "*Increasing Returns and Economic Geography.*" Journal of Political Economy, 1991, 99(3), pp. 483-99.

[98] Kuznets, S. "Quantitative Aspects of the Economic Growth of Nations: II." *Economic Development and Cultural Change*, 1957, Supplement to Vol. 5, pp. 3-111.

[99] Laffont, J., and J. Tirole. *A Theory of Incentives in Procurement and Regulation*. Cambridge: MIT University Press, 1993.

[100] LaFrance, J.T. and M.J. Watts. "Public Grazing in the West and "Rangeland Reform '94"." *American Journal of Agricultural Economics*, 1995, 77(3), pp. 447-61.

[101] Laitner, J. "Structural Change and Economic Growth." *Review of Economic Studies*, 2000, 67(3), pp. 545-61.

[102] Lange, A. and M.F. Quaas. "Economic Geography and the Effect of Environmental Pollution on Agglomeration." *The B.E. Journal of Economic Analysis & Policy*, 2007, 7(1).

[103] Lee, Dl and Wolpin, K.I. "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica*, 2006, 74(1), pp. 1-46.

[104] Lei, C.C., R.G. Zhang, and B.C. Wu. "Labor reallocation in China: 1978-2011." *Economic Modelling*. 2013, 35, pp. 668-673.

[105] Li, H. and V. Daly. "Testing the Balanced Growth Hypothesis: Evidence from China." *Empirical Economics*. 2009, 37(1), pp. 185-200.

[106] Loomis, J.B. *Integrated Public Lands Management: Principles and Applications to National Forests, Parks, Wildlife Refuges, and BLM Lands*. New York: Columbia University Press, 2002.

[107] Maddison, A. "Economic Growth and Structural Change in the Advanced Countries," In *Western Economies in Transition: Structural Change and Adjustment Policies in Industrial Countries*, ed. I. Leveson and J. W. Wheeler, 41-65. London: Croom Helm, 1980.

[108] Maher, A. “The Economic Impacts of Sagebrush Steppe Wildfires on an Eastern Oregon Ranch.” Agricultural and Resource Economics, Vol. Master of Science: 172. Oregon State University: Corvallis, 2007.

[109] Mas-Colell, A., M.D. Whinston, and J.R. Green. *Microeconomic Theory.* New York, Oxford: Oxford University Press, 1995.

[110] Mao, R., and Y. Yao. "Structural Change in a Small Open Economy: An Application to South Korea." *Pacific Economic Review.* 2012, 17(1), pp. 29-56.

[111] Matsuyama, K. "Productivity-Based Theory of Manufacturing Employment Decline: A Global Perspective," 2007. Unpublished.

[112] McCarl, B.A. and R.F. Brokken. “An Economic Analysis of Alternative Grazing Fee Systems.” *American Journal of Agricultural Economics*, 1985, 67(4), pp. 769-78.

[113] Mills, T.C. "Great ratios and common cycles: do they exist for the UK?" *Bulletin of Economic Research,* 2001, 53(1), pp. 35–51.

[114] Muellbauer, J. "Community Preferences and the Representative Consumer." *Econometrica*, 1976, 44(5), pp. 979-99.

[115] Murphy, K.M., A. Shleifer, and R.W. Vishny. "Industrialization and the Big Push." *Journal of Political Economy*, 1989, 97(5), pp. 1003-26.

[116] Ngai, L.R., and C.A. Pissarides. "Structural Change in a Multisector Model of Growth." *American Economic Review*, 2007, 97(1), pp. 429-43.

[117] Nordhaus, W.D. "Alternative Methods for Measuring Productivity Growth." National Bureau of Economic Research, Inc, NBER Working Paper No. 8095, 2001.

[118] Obstfeld, M., and K. Rogoff. *Foundations of International Macroeconomics.* Cambridge, Mass. and London: MIT Press, 1996.

[119] Palivos, T., and P. Wang. "Spatial Agglomeration and Endogenous Growth." *Regional Science and Urban Economics*, 1996, 26(6), pp. 645-69.

[120] Papanastasis, V.P. "Restoration of Degraded Grazing Lands through Grazing Management: Can It Work?" *Restoration Ecology*, 2009, 17(4), pp. 441-45.

[121] Pitelis, C., and N. Antonakis. "Manufacturing and Competitiveness: The Case of Greece." *Journal of Economic Studies*, 2003, 30(5-6), pp. 535-47.

[122] Poterba, J.M. "Rate of return to corporate capital and factor shares: new estimates using revised national income accounts and capital stock data." *Carnegie–Rochester Conference Series on Public Policy*, 1998, 48, pp. 211–246.

[123] Prendergast, C. "The Tenuous Trade-off between Risk and Incentives." *Journal of Political Economy*, 2002, 110(5), pp. 1071-102.

[124] Roach, B., J. Loomis, and R. Motroni. "Economic Analysis of Deer Management Alternatives on Public Lands in Northern California." *Human Dimensions of Wildlife*, 1996, 1(2), pp. 14-23.

[125] Roback, J. "Wages, Rents, and the Quality of Life." *Journal of Political Economy*, 1982, 90(6), pp. 1257-78.

[126] Romer, P.M. "Endogenous Technological Change." *Journal of Political Economy*, Part 2, 1990, 98(5), pp. S71-102.

[127] Rosenthal, S.S. and W.C. Strange. "Evidence on the Nature and Sources of Agglomeration Economies," *Handbook of Regional and Urban Economics.* Elsevier, 2004, 2119-71.

[128] Rowthorn, R. and R. Ramaswamy. "Growth, Trade, and Deindustrialization." *IMF Staff Papers*, 1999, 46(1), pp. 18-41.

[129] Sachs, J.D. and H.J. Shatz. "Trade and Jobs in U S Manufacturing." *Brookings Papers on Economic Activity*, 1994, 0(1), pp. 1-69.

[130] Saxenian, A. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge: Havard University Press, 1996.

[131] Seierstad, A. and K. Sydsaeter. "Sufficient Conditions in Optimal Control Theory." *International Economic Review*, 1977, 18(2), pp. 367-91.

[132] Shavell, S. 1984. "A Model of the Optimal Use of Liability and Safety Regulation." *The RAND Journal of Economics*, 15:2, pp. 271-80.

[133] Sheriff, G. "Optimal Environmental Regulation of Politically Influential Sectors with Asymmetric Information." *Journal of Environmental Economics and Management*, 2008, 55(1), pp. 72-89.

[134] Shonkwiler, J.S. and J. Englin. "Welfare Losses Due to Livestock Grazing on Public Lands: A Count Data Systemwide Treatment." *American Journal of Agricultural Economics*, 2005, 87(2), pp. 302-13.

[135] Siegel, J.J. "The real rate of interest from 1800–1990: a study of the U.S. and the U.K." *Journal of Monetary Economics,* 1992, 29, 227–252.

[136] Simon, C.J. "Industrial Reallocation across U.S. Cities, 1977-1997." *Journal of Urban Economics*, 2004, 56(1), pp. 119-43.

[137] Solow, R.M. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, 1956, 70, pp. 65-94.

[138] Sundaram, R.K. *A First Course in Optimization Theory.* Cambridge, United Kingdom: Cambridge University Press, 1996.

[139] Swanson, S. Personal Interview. November 17^{th}, 2008.

[140] Swanson, S., ed. *Nevada Rangeland Monitoring Handbook, Second Edition.* Cooperative Extension Educational Bulletin 06-03, University of Nevada, 2006.

[141] Taylor, M.H., K. Rollins, M. Kobayashi, and R.J. Tausch. "The Economics of Fuel Management: Wildfire, Invasive Plants, and the Evolution of Sagebrush Rangelands in the Western United States". University of Nevada, Reno, Joint Economics Working Paper 11-002, 2011.

[142] Torell, L.A. and J.P. Doll. "Public Land Policy and the Value of Grazing Permits." *Western Journal of Agricultural Economics*, 1991, 16(1), pp. 174-84.

[143] Torell, L.A. K.S. Lyon, and E.B. Godfrey. "Long-Run versus Short-Run Planning Horizons and the Rangeland Stocking Rate Decision." *American Journal of Agricultural Economics*, 1991, 73(3), pp. 795-807.

[144] Torell, L.A., N.R. Rimbey, O.A. Ramirez, and D.W. McCollum. "Income Earning Potential versus Consumptive Amenities in Determining Ranchland Values." *Journal of Agricultural and Resource Economics*, 2005, 30(3), pp. 537-60.

[145] Triplett, J.E. and B.P. Bosworth. "Productivity Measurement Issues in Services Industries: 'Baumol's Disease' Has Been Cured." *Federal Reserve Bank of New York Economic Policy Review*, 2003, 9(3), pp. 23-33.

[146] U.S. Department of Agriculture, Economic Research Service, 2009. International Food Consumption Patterns Database. http://www.ers.usda.gov/data/InternationalFoodDemand/

[147] U.S. Department of the Interior, Bureau of Land Management. "Investing in Range Improvement on Public Lands: Factsheet." 2008. Retrieved June 16th, 2010, from the Bureau of Land Management Web Site:

http://www.blm.gov/pgdata/etc/medialib/blm/wo/Information_Resources_Management /policy/im_attachments/2008.Par.37258.File.dat/IM2008-170_att2.pdf.

[148] U.S. Department of the Interior, Bureau of Land Management. “RAC Standards and Guideline Rangeland Health.” 2007. Retrieve June 16th, 2010, from Bureau of Land Management Web Site: http://www.blm.gov/nv/st/en/res/resource_advisory/sierra_front-northwestern/ standards_and_guideline.html

[149] U.S. Department of the Interior, Bureau of Land Management. “Resource Advisory Councils in Nevada.” 2010. Accessed on June 16th, 2010, from the Bureau of Land Management Web Site: http://www.blm.gov/nv/st/en/res/resource_advisory.3.html.

[150] U. S. Government Accountability Office (GAO). *Livestock Grazing: Federal Expenditures and Receipts Vary, Depending on the Agency and the Purpose of the Fee Charged*. Washington, DC, 2005.

[151] U.S. General Accounting Office (GAO). *Rangeland Management: Interior’s Monitoring Has Fallen Short of Agency Requirements*. Washington, DC, 1992.

[152] Valentinyi, A. and B. Herrendorf. "Measuring Factor Income Shares at the Sectoral Level." *Review of Economic Dynamics*, 2008, 11(4), pp. 820-835.

[153] Venables, A.J. "Equilibrium Locations of Vertically Linked Industries." *International Economic Review*, 1996, 37(2), pp. 341-59.

[154] Verburg, P.S.J., J.A. Arnone III, D. Obrist, D.E. Schorran, R.D. Evans, D. Leroux-Swarthout, D.W. Johnson, Y. Luo, and J.S. Coleman. “Net Ecosystem Carbon Exchange in Two Experimental Grassland Ecosystems.” *Global Change Biology*, 2004, 10, pp. 498-508.

[155] Wallsten, S.J. "An Empirical Test of Geographic Knowledge Spillovers Using Geographic Information Systems and Firm-Level Data." *Regional Science and Urban Economics*, 2001, 31(5), pp. 571-99.

[156] Walz, U. "Transport Costs, Intermediate Goods, and Localized Growth." *Regional Science and Urban Economics*, 1996, 26(6), pp. 671-95.

[157] Watts, M.J., J. Shimshack, and J.T. LaFrance. “Grazing Fees versus Stewardship on Federal Lands.” Working Paper, Department of Agricultural & Resource Economics, University of California, Berkeley, 2006.

[158] Weitzman, M.L. “Prices vs. Quantities.” *The Review of Economic Studies*, 1974, 41(4), pp. 477-91.

[159] Wood, A. "How Trade Hurt Unskilled Workers." *Journal of Economic Perspectives*, 1995, 9(3), pp. 57-80.

[160] Xu, F., R.C. Mittelhammer, and L.Al. Torell. “Modeling Nonnegativity via Truncated Logistic and Normal Distributions: An Application to Ranch Land Price Analysis.” *Journal of Agricultural and Resource Economics*, 1994, 19(1), pp. 102-14.

[161] Zhao, R.R. “All-or-Nothing Monitoring.” *American Economic Review*, 2008, 98(4), pp. 1619-28.