博士論文

論文題目　Functional Specialization of Cities

（都市の機能的特化）

氏　　名　長町　康平

Contents

Introduction

My dissertation focuses on functional specialization of cities, a modern, increasingly important form of a system of cities, in which new technologies and varieties of goods and services are developed in larger cities with the help of greater urban diversity and positive externalities, while in smaller cities production based on technologies developed in the former cities are conducted, often characterized as less skill-intensive, more routine economic activity (Duranton and Puga, 2005).

Except for the concluding chapter, the dissertation consists of three chapters, each corresponding to an independently readable paper. In Chapter 1, a model of a system of cities is developed in order to formalize the mechanism behind functional specialization of cities and then consider welfare implications of income redistribution across cities. Then, in Chapter 2, an implication of the specialization for income inequality (skill premium) is investigated using a two-region model, which can be viewed as a simplified and modified version of the model in the previous chapter more suitable for analysis of skill premium. In Chapter 3, functional specialization of cities is interpreted as team production, production by organizations with high-skilled managers supervising low-skilled workers at the top, and policy implications are investigated, especially focusing on the allocation of creativity in light of the increasing importance of knowledge and creativity in overall economic activity.

The followings are slightly more detailed summaries of the chapters.

Chapter 1: A Simple Model of Functional Specialization of Cities

This chapter develops a static equilibrium model of a system of cities in which ex ante identical locations specialize in stages of production different in the degree of dependence on routine and nonroutine local services sectors, the latter of which is tied to an agglomeration force due to (Dixit-Stiglitz type) monopolistic competition. The model is simple in that the system is summarized by a second-order difference/differential equation, which has a unique non-degenerate city size distribution with the comovement of income, population, factor prices, and urban diversity as observed for the U.S. cities.

The model is an extension of Matsuyama's (2013) to an urban context, and circular causation is the key mechanism that induces a market equilibrium to exhibit functional specialization of cities. If stages of production which demand for nonroutine services more than the other stages concentrate in a particular location, then a large number of differentiated firms enter the local nonroutine services sector there. Since the larger number of varieties reduces the price index of non-routine services in that location, the concentration of stages with higher skill intensities strengthens further.

Due to this circular causation, a market outcome results in an inefficient allocation (Matsuyama, 1995). In this sense, the modeling approach is in contrast to standard urban economics models of a system of cities à la Henderson (1974) and Black and Henderson (1999) including Duranton and Puga's (2005) model of functional specialization of cities, where a market allocation is efficient thanks to the existence of an institution, e.g., competitive city developers, that fixes sources of inefficiency.

Therefore, this chapter considers a simple income redistribution policy as a tool that corrects inefficiency to some extent. Although introducing income redistribution makes makes the model analytically intractable, a market equilibrium is still characterized by a difference/differential equation easily solved with a numerical method and thus useful for further analyses.

Chapter 2: Comparative Advantage and Skill Premium of Regions

This chapter provides one explanation for why a positive correlation between the size and skill premium of a region emerges by providing a comparative advantage model with a continuum of mobile heterogeneous individuals as well as a continuum of final goods sectors that are different in terms of their skill intensities of intermediate goods. All individuals choose their occupations depending on their productivity, and any occupation can freely migrate across regions. This location-occupation choice then interacts with the regional comparative advantage in final goods sectors which depends on the regional offer prices of two different types of intermediate goods, one of which features (Dixit-Stiglitz type) monopolistic competition. Although regions are ex-ante identical, interactions between individuals' location-occupation choices and regional comparative advantage result in a self-organized positive correlation between the skill premium and income of regions. The theory can also accommodate the interpretation that the regional difference in skill premium is caused by specialization in task trade within firms, not industries.

The crucial difference of the model in this chapter from that in the previous chapter is that there is individual heterogeneity: individuals are heterogeneous in productivity of producing a differentiated good. Given that the specification follows Melitz (2003), the model can be view as an extension of Matsuyama (2013) to a class of heterogeneous agents. In the model, circular causation is again the key mechanism that induces a market equilibrium to exhibit a positive correlation between the size and the skill premium of a region. In this sense, the current approach, i.e., functional specialization of cities, differs from Davis and Dingel (2012) who resort to knowledge spillover as a mechanism of the concentration of economic activity.

Chapter 3: Team Production and the Allocation of Creativity across Global and Local Sectors

This chapter develops a two-sector Ricardian comparative advantage model with team production in order to obtain implications for policies encouraging team production in a sector with only low-skilled agents to attract high-skilled agents to that sector. It is shown that team production changes the nature of comparative advantage, providing a policy device for attracting creativity. It is also shown that policy targets, reducing cost of communication within teams, reducing cost of learning within team, and increasing productivity gain from such learning, should be carefully selected because likelihoods of success in attracting creativity are different across the targets, and they exhibit some non-monotonicity in the dynamics in the allocation of creativity.

The model is based on team production models developed by Garicano and Rossi-Hansberg (2006a,b) who investigate how knowledge is used in the economy especially focusing on the emergence of organizations. Similar to models of functional specialization of cities including Duranton and Puga (2001) and Duranton and Puga (2005), there is an equilibrium where some location specializes in skill-intensive economic activity, while the other specializes in less skill-intensive one. Unlike these studies, the primal focus of this chapter is on comparative statics of the allocation of skilled agents with respect to policy-related parameters.

Acknowledgment

I would like to thank Takatoshi Tabuchi, Kiminori Matsuyama, Takaaki Takahashi, Federico Trionfetti, Marcus Berliant, Hajime Takatsuka, Akihiko Matsui, Tomoya Mori, Rafael Gonzáles-Val, Shota Fujishima, Daisuke Oyama, Kenichi Amaya, Tadashi Morita, Daichi Shirai, and Dao-Zhi Zheng for their helpful and insightful comments and suggestions. I would also like to thank comments from participants of various seminars, workshops, and conferences. This research is financially supported by the JSPS Grant-in-Aid for Research Activity Start-up No.24830025 and the Project "Spatial Economic Analysis on Regional Growth" undertaken at Research Institute of Economy, Trade and Industry (RIETI).

Chapter 1

A Simple Model of Functional Specialization of Cities

1.1 Introduction

In the urban economics literature, it is documented that the form of urban specialization has been changing from traditional, sectoral to functional specialization (Duranton and Puga, 2005). Although there is no clear-cut relationship between the population size and the number of executives and managers of a city decades ago, we now observe the contrary. Meanwhile, the degree of sectoral specialization of cities, which is measured by a Gini specialization index, tends to decrease. As in the trade literature such as Grossman and Rossi-Hansberg (2008), Duranton and Puga (2005) argue that a decrease in the costs of remote management including improvements in information and communication technologies allows firms efficiently to fragment stages of production across cities, resulting in functional specialization of cities.

Since larger cities are likely to be associated with higher living costs and a larger number of varieties of goods and services, it is natural to guess that the modern form of urban specialization should imply that skill intensity of cities is positively correlated with these variables. Figure 1.1 actually confirms these positive correlations. Panel (a)-(e) in the figure plot skill intensity, the market size, the wage rate, the land value (as a proxy for the land rent), and urban diversity, respectively, against the total employment size for Metropolitan Statistical Areas (hereafter, MSAs).[1] Each of the

[1] The data on the variables except for the last three are based on the *May 2011 Occupational Employment Statistics* (hereafter, OES). Skill intensity of an MSA is defined as the share of detailed occupations listed in the *Standard Occupational Classification System* (hereafter, SOC), more than 20% of employments of which have a master's degree or a higher one, in the labor compensation of all detailed occupations reported by OES. The market size of an MSA is approximated by the total labor compensation. The wage rate of an MSA is the median of annual wage rates of detailed occupations in the MSA. The land value (per acre) is taken from Table A3 in Albouy and Ehrlich (2012). As for urban diversity and the establishment size, we calculate these as the total number of establishments and the number of employments per establishment, respectively, of "Professional, Scientific, and Technical Services" (in the *North American Industry Classification System*), using the *2011 Statistics U.S. Businesses* (hereafter, SUSB).

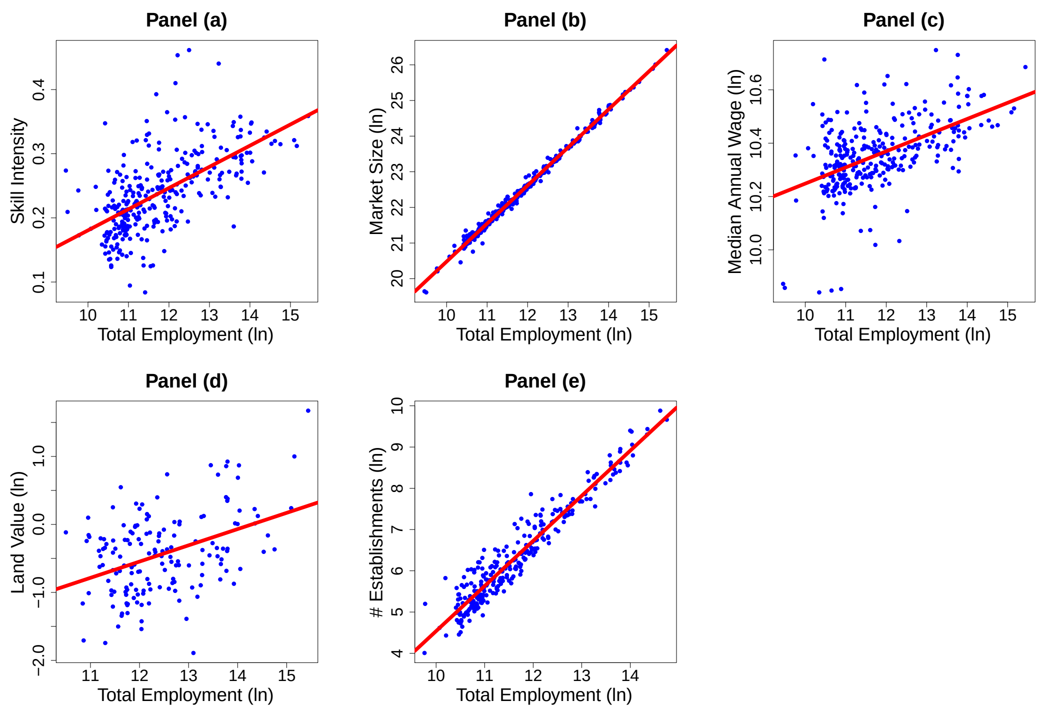

Figure 1.1: Correlations

panels also depicts a red line corresponding to the linear regression. Spearman's rank correlation coefficients of the panels are 0.650, 0.994, 0.503, 0.318, 0.937, and 0.471, respectively.

Are these positive correlations just an ad-hoc connection of the facts mentioned in the literature? Answering to this question is the purpose of this chapter, and the answer is YES: it is shown that these positive correlations are a general result of a market economy.

To this end, we extend Matsuyama's (2013) international trade model to a spatial equilibrium model of a system of cities with a continuum of stages of production in which ex ante identical locations specialize ex post in different sets of stages of production. Each stage of production differs in the degree of dependence on routine and nonroutine local services sectors, characterized by perfect competition and monopolistic competition á la Dixit and Stiglitz (1977), respectively, and those with higher skill intensity, equivalently higher cost share of the nonroutine local services, benefit more from varieties of nonroutine local services or urban diversity facilitating complex problem-solving activities. Therefore, urban diversity acts as an agglomeration force, that ensures that concentration into a particular location is sustainable despite being associated with a higher rental price of land

Here, we focus on the MSAs for which OES report non-military detailed occupations that cover more than 75% of the total employments in order to reduce a sample selection bias in the correlation between the total employment size and skill intensity. This results in the sample size of 303 in the cases which use OES only, i.e., Panel (a)-(c). The sample size of Panel (d) is reduced to 160, because the definitions of MSAs in Albouy and Ehrlich (2012) are not exactly the same as in our study, which focuses on MSAs listed in OES. The sample sizes for the case of Panel (e) and (f), which use both OES and SUSB, are in the middle of the above two cases, 279 and 278, respectively.

which constitutes production inputs together with labor and is also consumed by households. Specialization of cities then, by generating differences in the average skill intensity among cities, results in a non-degenerate city size distribution. As in Matsuyama (2013), an equilibrium is characterized as a single, second-order difference or differential equation which we call the fundamental equation.

Using the model, we show that an equilibrium with functional specialization of cities exists and is unique in the sense that there exists a unique, non-degenerate size distribution of cities. Furthermore, this equilibrium is shown to be characterized by the comovement of income, population, the wage rate, the land rent, the average establishment size (in the monopolistic competition sector), and the number of varieties of local nonroutine services, being actually observed for the U.S. metropolitan areas (Figure 1.1). The analytical tractability of the model also allows us to derive the necessary and sufficient condition for the size distribution of cities to obey a power law including Zipf's law as a special case.

It should be noted that an important feature of the model is the non-existence of large, competitive city developers or local governments which are often introduced in standard urban economics models such as Black and Henderson (1999) and Rossi-Hansberg and Wright (2007) among others. In those models, developers or local governments act as a device internalizing technological externalities and coordination failure associated with agents' location choices, resulting in an efficient market outcome.

Therefore, there are two implications of the current modeling approach. The first is that the positive correlations need not be related with existence of active developers or local governments. As a result, it is expected that functional specialization of cities tends to be characterized as the above mentioned positive correlations in various countries with highly developed information and communication technologies irrelevant to whether countries have active developers or local governments. The second is that without such developers or local governments, functional specialization of cities is in general associated with an inefficient market outcome. In this case, there is room for policy intervention, and relevant questions are "How can we improve welfare?", "To what extent can we improve welfare?", and etc.

As for the second, this chapter focuses on a simple income redistribution policy scheme: the central government levies a proportional income tax and redistributes the collected income across cities as a lump-sum transfer. It is found that extending the model in such a way still results in a set of equilibrium conditions which are characterized as a single, second-order difference/differential equation. Although the model is no longer analytically solvable unlike the laissez-faire case, an equilibrium is easily solved with a numerical method.

Given the form of income redistribution, we then ask what is the optimal tax rate, and to what extent welfare is improved with the optimal tax rate. It is suggested that the optimal tax rate decreases as the economy is more characterized as monopolistic competition. Consistent with this result, it is also found that the optimal tax rate is decreasing as the (constant) elasticity of substitution between varieties decreases if a sufficiently large part of the economy exhibits monopolistic competition. When the importance of monopolistic competition is not sufficiently large, the optimal tax rate has an inverted-U-shape relationship between the elasticity of substitution. The associated welfare gain from income redistribution is limited, e.g., less than about 2.3% increase in life-time utility in the most of an empirically relevant range of parameters, consistent with Desmet and Rossi-Hansberg (2013).

This chapter is related to three strands of literature. The first strand is the literature on the international trade theory of Ricardian comparative advantage. We extend the Matsuyama's (2013) model to a regional context and also go a step further by conducting a welfare analysis of income redistribution. He introduces monopolistic competition à la Dixit and Stiglitz (1977) into Dornbusch et al. (1977), resulting in a model with symmetry break through endogenous comparative advantage. He also develops a new method allowing us to characterize an equilibrium of the world with multiple and arbitrarily large number of countries as a second-order difference or differential equation, the latter of which is analytically solvable. Our model has the same degree of analytical tractability as in Matsuyama's (2013) international case, even if we include migration across cities and an immobile factor of land, that is used not only for production but also for consumption as in Pflüger and Tabuchi (2010). The current urban application of Ricardo's theory with endogenous comparative advantage brings us a clear relationship between comparative advantage and agglomeration and dispersion forces of economic activity.

The second strand of literature is the research on a system of cities. To our knowledge, this is the first paper that, using a system of cities model with functional specialization, conducts a welfare analysis of an income redistribution policy. Economic theories of the size distribution of cities range from stochastic growth models to static deterministic ones. Some models of the former type such as Eeckhout (2004) do not focus on the specialization of cities, whereas models such as Rossi-Hansberg and Wright (2007) and Duranton (2007) (of the former type) and Hsu (2012) (of the latter type) focus on cross-city variation in industries. Recent studies by Behrens et al. (2014a) and Behrens et al. (2014b) advance our understanding on the effects of heterogeneous agents and urban amenities on the size distribution of cities, respectively. However, the main focus of the above studies is on the positive side. While providing a model of a system of cities with functional specialization, Duranton and Puga (2005) preclude the possibility of an income redistribution policy by introducing competitive developers á la Henderson (1974). In addition, unlike their analysis, we do not focus on the transformation of urban specialization and thus assume a single sector. This can be interpreted as a situation where information and communication technologies are highly developed, resulting in a costless pooling of stages of production across sectors, e.g., head quarters' services tend to agglomerate in the largest city in a country. From a different perspective, Desmet and Rossi-Hansberg (2013) advances our understanding on welfare effects of efficiency, amenities, and frictions by applying the Business Cycle Accounting method in the macroeconomics literature to the urban context, but an optimal redistribution is not their scope of analysis.

The third is trade models which specify the production of goods as a continuum of intermediate goods, stages of production or tasks. The current model does not include trade costs of stages of production for analytical tractability and thus heterogeneous trade costs, introduced by Grossman and Rossi-Hansberg (2008). In this sense, the model specification is somewhat similar to previous international trade models.[2] Instead of heterogeneous trade costs, heterogeneity in skill intensity allows the fraction of stages of production is determined endogenously, reflecting their skill intensities and the comparative advantages of cities, the latter of which are in turn determined endogenously, i.e., a circular causation process.

The remainder of this chapter is organized as follows. We first provide the model in Section 1.2. In Section 1.3, we discuss the equilibrium properties including the welfare gain from functional specialization. The model is then extended by introducing an income redistribution policy using an income tax and lump-sum transfers across cities to analyze the optimal income redistribution policy in Section 1.4. Finally, we conclude this chapter in Section 1.5.

1.2 The Model

In this section, we provide a simple spatial equilibrium model with a continuum of stages of production with heterogeneity in skill intensity that are fragmented and outputs of which are traded across cities and within firms. The economic agents consist of mobile workers, final good firms, and city-specific local services firms, the last of which include monopolistically competitive and perfectly competitive firms. Here, local services firms provide outsourcing services that are necessary for conducting stages of production, and the variety effect introduced via monopolistic competition á la Dixit and Stiglitz (1977) captures urban diversity. In the latter sense, we call the monopolistically competitive sector the nonroutine sector and the other the routine sector. This stylized specification makes clear the distinction between the two sectors in a way that the nonroutine sector conducts nonroutine tasks that often require complex problem-solving activities which could be facilitated with the help of local network. The model is essentially an application of Matsuyama's (2013) framework to the urban context.

In the following, we explain the optimization problems of the agents in order. For the sake of convenience, we first assume that there are $J \in \mathbb{N}$ ex ante identical locations in the economy,[3] each of which is endowed with one unit of land. We subsequently modify this assumption by making J diverge to infinity but making the measure of each location converges to zero in such a way that the

[2] The current model differs from these studies in more respects. Unlike Grossman and Rossi-Hansberg (2008), each task is not related to a particular production factor. Rather, all tasks use the same set of production inputs; the skill intensity of each task is different from each other; and there is a continuous distribution of such skill intensity. In this sense, except for the use of labor and land instead of labor and capital as production inputs, the specification of our model is close to those of Dixit and Grossman (1982), Yi (2003), and Kohler (2004). However, the current model also shares the same assumption with Grossman and Rossi-Hansberg (2008) as well as Feenstra and Hanson (1996) in that there is no vertical linkage between different tasks or between intermediate inputs. Importantly, the current model differs from all these studies in that it deals with an arbitrarily large number of locations rather than just two countries or a single small open economy.

[3] In this chapter, we do not distinguish cities, regions, and locations and use these words interchangeably.

total measure of locations is equal to unity, which allows us to consider the size distribution of cities.[4]

1.2.1 Workers

There is a unit mass of identical workers in the economy. Each worker is freely mobile across locations and thus decides her location as well as consumption of goods and services.

Suppose that a worker had already chosen her residence $j \in \{1, 2, \cdots, J\}$, which is also her workplace. Then, she solves the following utility maximization problem:

$$U_j = \max_{c_j, h_j \geq 0} c_j^{1-\alpha} h_j^{\alpha} \quad s.t. \quad Pc_j + R_j h_j = W_j + \bar{R}_j, \qquad 0 < \alpha < 1.$$

Her income consists of labor income W_j and subsidy $\bar{R}_j$.[5] Assuming that the land rents in a city are received by the residents of that city as in Behrens et al. (2014b) and Michaels et al. (2013),[6] $\bar{R}_j$ is given by

$$\bar{R}_j = \frac{R_j}{N_j} \quad \forall j. \tag{1.1}$$

She uses these sources of income to consume homogeneous tradeable good c_j and housing h_j, the prices of which are P and R_j, respectively.

The associated indirect utility function, together with free migration, then implies

$$\frac{W_j + \bar{R}_j}{W_{j'} + \bar{R}_{j'}} = \left(\frac{R_j}{R_{j'}}\right)^{\alpha} \qquad \forall j \neq j', \tag{1.2}$$

which imposes a restriction on the relationship between income and land rent differentials.

1.2.2 Final Good Sector

The tradeable homogeneous final good is produced using a constant-returns-to-scale (CRS) Cobb-Douglas production technology. More specifically, the production of one unit of the final good re-

[4] Whether the distribution corresponding to a discrete J converges to that with $J = \infty$ as J increases itself is an important research topic. However, answering such a question is beyond the scope of this chapter. In the following, we simply assume that the case of $J = \infty$ can approximate the one with a large J, or we focus only on the limiting case.

[5] We assume that each worker is endowed with one unit of time and supplies it inelastically. As discussed later, there are two different sectors, the nonroutine and routine sectors, that mainly conduct nonroutine and routine tasks, respectively. For simplicity, we assume that each worker has an equal ability to perform each type of task. Thus, in an equilibrium where both sectors locate at almost every locations, wages rates within each location are equalized.

[6] An equivalent specification is to introduce immobile, competitive land owners in each location. Although we can consider an environment where land in each location is equally owned by all workers, this makes analytical characterization infeasible, since the individual income then includes the economy-wide average land rent, a term independent of location index j.

quires to process a continuum of stages of production (hereafter, stages) $\{t : 0 \leq t \leq 1\}$:[7]

$$Y = \exp\left[\int_0^1 \ln(y(t))dt\right], \tag{1.3}$$

where Y and $y(t)$ denote outputs of the final good and stage t, respectively. The equal weights and unit mass of the stages imply that the total sales, which are equal to the economy-wide income $E = \sum_{j=1}^{J}(W_j N_j + R_j)$ times the expenditure share $(1-\alpha)$ of the final good, are distributed to each stage t.

Firms decide where each of these stages is performed. For each fixed stage $t \in [0,1]$, a typical firm decides the quantity $y_j(t)$ of production of stage t at a location $j \in \{1,2,\cdots,J\}$. Once the stage has been processed at each location, outputs $\{y_j(t)\}_{j=1}^{J}$ are aggregated and transported without cost to use them as a production input of the final good:[8]

$$y(t) = \sum_{j=1}^{J} y_j(t) \quad \forall t. \tag{1.4}$$

Furthermore, stage t is processed at location j with the help of the outsourcing services by local nonroutine services $X_{n,j}(t)$ and local routine services $X_{r,j}(t)$. Its output $y_j(t)$ is given by[9]

$$y_j(t) = X_{n,j}(t)^{\gamma(t)} X_{r,j}(t)^{1-\gamma(t)} \quad \forall j,t, \tag{1.5}$$

where $\gamma(t) \in [0,1]$ represents the skill intensity of stage t. In the following, we assume that $\gamma(t)$ is strictly monotonically increasing. For analytical convenience, we also assume that $\gamma(t)$ is continuously differentiable, i.e., $\gamma'(t) > 0$ for all t. We also assume that $\gamma(0) = 0$ and $\gamma(1) = 1$.

Therefore, letting $P_{n,j}$ and $P_{r,j}$ denote the prices of the nonroutine and routine services, respectively, we can write the profit maximization of the final good firm as follows:

$$\max PY - \sum_{j=1}^{J} \int_{\mathbb{T}_j} \left[P_{n,j} X_{n,j}(t) + P_{r,j} X_{r,j}(t)\right] dt \quad s.t. \quad (1.3)-(1.5), \text{ and } \mathbb{T}_j \subseteq [0,1],$$

where the control variables consist of the measurable set $\mathbb{T}_j$ of stages processed at location j as well

[7] In Matsuyama (2013), what we consider to be a continuum of stages of production here is interpreted as a continuum of sectors, and the technology specified by (1.3) directly enters the utility function. However, in the regional context, this interpretation is not favorable given the transition from sectoral specialization of cities to functional one argued by Duranton and Puga (2005).

[8] Transportation costs can be introduced by dividing goods and services into tradeable and non-tradeable parts with CRS Cobb-Douglas composite technology, where the share parameter corresponding to the tradeable is interpreted as the freeness of trade.

[9] The implicit assumption here is that either type of local services alone cannot be traded across locations. This exactly corresponds to what we do when communicating with people in different locations. For example, sending an interesting idea to your friend via e-mail breaks down to two stages: developing or formulating the idea, a nonroutine task, and writing or sending an e-mail, a routine task. Although each task is not tradeable itself, the output, i.e., the combination of those tasks, is now tradeable.

as quantities $(Y,\{y(t)\}_t,\{y_j(t),X_{n,j}(t),X_{r,j}(t)\}_{j,t})$. Defining $|\mathbb{T}_j|$ as the Lebesgue measure of $\mathbb{T}_j$, we can see that if $\mathbb{T}_j$'s are all mutually exclusive, which is indeed the case, $|\mathbb{T}_j|$ fraction of the total sales $(1-\alpha)E$ is distributed to location j. In addition, $\gamma(t)$ and $1-\gamma(t)$ fractions of the distribution $(1-\alpha)E$ to stage t are distributed to the nonroutine and routine sectors, respectively. Thus, the nonroutine sector at location j receives $\int_{\mathbb{T}_j}\gamma(t)dt(1-\alpha)E \equiv \Gamma_j|\mathbb{T}_j|(1-\alpha)E$, where $\Gamma_j \equiv |\mathbb{T}_j|^{-1}\int_{\mathbb{T}_j}\gamma(t)dt$ is the average skill-intensity of location j.

In addition, as we discuss in the next two subsections, we assume that the market structures of the nonroutine and routine sectors are monopolistically and perfectly competitive, respectively.[10] More specifically, $X_{n,j}(t)$ denotes the composite of a continuum of varieties $\{x_{n,j}(v,t)\}_{v\in[0,D_j]}$. The quantity of each variety is also a control variable, with the following technology:

$$X_{n,j}(t) = \left[\int_0^{D_j} x_{n,j}(v,t)^{\frac{\sigma-1}{\sigma}}dv\right]^{\frac{\sigma}{\sigma-1}} \quad \forall j,t\in\mathbb{T}_j,$$

where D_j is the number of varieties produced at location j, which we call *urban diversity*, and $\sigma > 1$ is the elasticity of substitution between any two different varieties.

The profit maximization then implies the following demand for variety v from stage t processed at location j:

$$x_{n,j}(v,t) = \left[\frac{p_{n,j}(v)}{P_{n,j}}\right]^{-\sigma} X_{n,j}(t) \quad \forall v,j,t\in\mathbb{T}_j,$$

where $P_{n,j}$ is the price index of the nonroutine services at location j given by

$$P_{n,j} = \left[\int_0^{D_j} p_{n,j}(v)^{-\frac{1}{\theta}}dv\right]^{-\theta} \quad \forall j. \tag{1.6}$$

Here, $\theta \equiv 1/(\sigma-1)$.

1.2.3 Nonroutine Local Service Sector

As mentioned in the previous subsection, the nonroutine local service sector is characterized by monopolistic competition á la Dixit and Stiglitz (1977), and each firm produces one variety. In addition, as in Pflüger and Tabuchi (2010), we assume that production inputs consist of both labor and land. More specifically, the fixed and marginal costs of production are both measured in terms of their Cobb-Douglas composite with a cost share parameter $\beta\in(0,1)$ for land.

Formally, variety-v firm at location j solves

$$\pi_j(v) = \max_{p_{n,j}(v),q_j(v)} \left[p_{n,j}(v) - mR_j^{\beta}W_j^{1-\beta}\right]q_j(v) - R_j^{\beta}W_j^{1-\beta}f \quad s.t. \quad q_j(v) = \int_{\mathbb{T}_j} x_{n,j}(v,t)dt,$$

[10] This stylized specification reflects our view of the nature of nonroutine services such as management, research and development, and legal services as well as routine services such as line production based on previously developed blueprints.

where $q_j(v)$ is the output of variety-v firm at location j, and m and f denote the shift parameters of marginal and fixed costs, respectively.

The profit maximization then implies the optimal pricing rule specified by $p_{n,j}(v) = (1+\theta)mR_j^{\beta}W_j^{1-\beta}$, and substituting this into (1.6) and taking the ratio across two different locations, j and j', we obtain

$$\frac{P_{n,j}}{P_{n,j'}} = \left(\frac{R_j}{R_{j'}}\right)^{\beta}\left(\frac{W_j}{W_{j'}}\right)^{1-\beta}\left(\frac{D_j}{D_{j'}}\right)^{-\theta} \quad \forall j \neq j'. \tag{1.7}$$

1.2.4 Routine Local Service Sector

Unlike the nonroutine local service sector, the routine sector is characterized by perfect competition with a CRS Cobb-Douglas production technology:

$$\max_{H_{r,j},\ L_{r,j}} P_{r,j}H_{r,j}^{\beta}L_{r,j}^{1-\beta} - R_jH_{r,j} - W_jL_{r,j}.$$

Note that the production of routine services also requires both labor and land. For simplicity, we assume the same cost share parameter β as in the nonroutine sector.[11]

The profit maximization then implies

$$\frac{P_{r,j}}{P_{r,j'}} = \left(\frac{R_j}{R_{j'}}\right)^{\beta}\left(\frac{W_j}{W_{j'}}\right)^{1-\beta} \quad \forall j \neq j'. \tag{1.8}$$

1.2.5 Equilibrium

We now define a market equilibrium. Since the locations are ex ante identical by assumption, the symmetric configuration always exists. However, this configuration does not seem robust to exogenous shocks to the economy, and thus, we focus on another type of equilibrium, which is the only equilibrium other than the symmetric one and which is unique at least in the limiting case $J \to \infty$, which is of general interest.

Specifically, we define an equilibrium with rankings. Without loss of generality, assume that $0 < |\mathbb{T}_1| < |\mathbb{T}_2| < \cdots < |\mathbb{T}_{j-1}| < |\mathbb{T}_j| < \cdots < |\mathbb{T}_J|$, i.e., as j increases, the associated market share increases. It is then immediately demonstrated that there must exist an increasing sequence $\{T_j\}_{j=0}^{J}$ of thresholds such that $\mathbb{T}_j = (T_{j-1}, T_j]$ for all $j \in \{1, 2, \cdots, J\}$; $T_0 = 0$; and $T_J = 1$ under free migration and free entry into the nonroutine sector. That is, if cities are different, we observe a perfect sorting of stages of production,[12] and the higher the location index j is, the higher the average skill intensity Γ_j is. Note that this configuration of the equilibrium is consistent with the argument advanced by

[11] More precisely, there are two reasons why we introduce this assumption. The first is for simplicity. Without this assumption, we cannot obtain the analytical solution discussed later. The second is to make clear the distinction between the two local services sectors. With this assumption, the nonroutine sector differs from the routine one in one way, market structure.

[12] Section .1 explains why the economy exhibits a perfect sorting in equilibrium.

Duranton and Puga (2001) that cities sort themselves into specialized cities, some of which host the research and development sectors testing prototype products, while others host the production sectors producing goods in a stylized manner. It should, however, be noted that in our model, there is no perfect specialization with respect to the service sectors in the sense that every city hosts both nonroutine and routine service sectors. Stated differently, the important characteristic that distinguishes one city from another is its average skill-intensity.

Therefore, we call the equilibrium on which we focus a *sorting equilibrium* and define it as follows:

Definition 1.1. *A sorting equilibrium is a set of prices* $(P, \{P_{n,j}, P_{r,j}, R_j, W_j\}_j, \{p_{n,j}(v)\}_{v,j})$*, quantities* $(Y, \{c_j, h_j, H_{r,j}, L_{r,j}\}_j, \{y(t)\}_t \{y_j(t), X_{n,j}(t), X_{r,j}(t)\}_{j,t}, \{x_{n,j}(v,t)\}_{v,t,j}, \{q_j(v)\}_v)$*, transfers* $\{\bar{R}_j\}_j$*, populations* $\{N_j\}_j$*, diversities* $\{D_j\}_j$*, and a sequence* $\{T_j\}_{j=0}^J$ *of thresholds such that*

1. *workers maximize their utility by choosing quantities and locations;*

2. *firms maximize their profits;*

3. *markets clear:*

$$\text{(Land)} \qquad R_j = (1-\alpha)\beta|\mathbb{T}_j|E + \alpha(W_j N_j + \bar{R}_j N_j), \tag{1.9}$$

$$\text{(Labor)} \qquad W_j N_j = (1-\alpha)(1-\beta)|\mathbb{T}_j|E; \tag{1.10}$$

4. *there is free entry into the nonroutine local service sectors; and*

5. *thresholds* $\{T_j\}_{j=0}^J$ *are consistent with comparative advantage:*

$$\left(\frac{P_{n,j+1}}{P_{n,j}}\right)^{\gamma(T_j)} \left(\frac{P_{r,j+1}}{P_{r,j}}\right)^{1-\gamma(T_j)} = 1, \tag{1.11}$$

where $P_{n,j+1}/P_{n,j} < 1$ *and* $P_{r,j+1}/P_{r,j} > 1$ *for all* j*.*

Here, the market clearing conditions, i.e., (1.9) and (1.10), are evident from the specifications of the utility and production technologies presented in the previous subsections.[13] The conditions that $P_{n,j+1}/P_{n,j} < 1$ and $P_{r,j+1}/P_{r,j} > 1$ imply that location $j+1$, compared to location j, has a comparative advantage in the nonroutine service sector and thus has a comparative advantage in stages with higher skill intensity. The threshold T_j here is the stage for which locations j and $j+1$ have the same comparative advantage, i.e., both locations j and $j+1$ have the same unit cost of producing stage T_j output.

[13]The market clearing condition for the final good is $PY = E$, where P denotes the price index of the final good. We omit this condition from the definition because we do not use it in deriving analytical results presented in the next section.

1.3 Equilibrium Properties

In this section, we first consolidate the equilibrium conditions presented in the previous section to obtain an equation governing the equilibrium of the economy in Subsection 1.3.1. We then prove the existence and uniqueness of a sorting equilibrium in Subsection 1.3.2, which has clear and empirically valid predictions about the relationships between the city size and some variables. The relation of our model to the size distribution of cities is also derived in Subsection 1.3.3.

1.3.1 Fundamental Equation

For a given J, we first show that the equilibrium system of the economy reduces to the following *fundamental equation*:

$$\left(\frac{T_{j+1}-T_j}{T_j-T_{j-1}}\right)^{\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma(T_j)} = \left(\frac{\Gamma_{j+1}}{\Gamma_j}\right)^{\theta\gamma(T_j)} \qquad \forall j \in \{1,2,\cdots,J-1\},$$

with $T_0 = 0$ and $T_J = 1$. Given the definition of Γ_j, i.e., $\Gamma_j = |T_j - T_{j-1}|^{-1}\int_{T_{j-1}}^{T_j}\gamma(t)dt$, the fundamental equation is a second-order difference equation with two boundary conditions.

For this purpose, we start with the result that consolidating market clearing conditions (1.9) and (1.10) along with the distribution of land rents, (1.1), yields

$$\frac{R_{j+1}}{R_j} = \frac{W_{j+1}N_{j+1}+R_{j+1}}{W_jN_j+R_j} = \frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|} \quad \forall j. \tag{1.12}$$

That is, given the ordering of $|\mathbb{T}_j|$, the higher the market share $|\mathbb{T}_j|$ is, the higher the land rent R_j and regional income $W_jN_j + R_j \equiv E_j$ are. In addition, it is also implied that differentials of land rent R_j and market share $|\mathbb{T}_j|$ are equal. That $R_j < R_{j+1}$ is simply because both locations have the same area of land while income concentrates in location $j+1$.

This result is then combined with the free-migration condition (1.2) to obtain

$$\frac{N_{j+1}}{N_j} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{1-\alpha} \quad \forall j, \tag{1.13}$$

which states that the ordering of population N_j is the same as the market share $|\mathbb{T}_j|$, and that the differential of population N_j is less than proportional to that of the market share $|\mathbb{T}_j|$ or, using (1.12), the land rent R_j or local income E_j. The latter is interpreted as the spatial equilibrium imposing some upper bound on the population differential. As a result, the ordering of the wage rate W_j is also the same as $|\mathbb{T}_j|$ because the labor market clearing condition (1.10) implies that the differential of labor compensation W_jN_j is equal to that of the market share $|\mathbb{T}_j|$:

$$\frac{W_{j+1}}{W_j} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{\alpha} \quad \forall j. \tag{1.14}$$

That $W_j < W_{j+1}$ is interpreted as a differential compensating the disparity in the cost of living, i.e., $R_j < R_{j+1}$. The spatial equilibrium (1.2) limits the degree of the wage differential, allowing the comovement between the wage rate and population.

These results immediately imply that a higher market share $|\mathbb{T}_j|$ is associated with more severe market competition, which is a *dispersion force*, in the sense that the unit production cost and thus the price $P_{r,j}$ of the routine local services is higher in that location. More specifically, substituting (1.12) and (1.14) into (1.8), we obtain

$$\frac{P_{r,j+1}}{P_{r,j}} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{\alpha(1-\beta)+\beta} \qquad \forall j. \tag{1.15}$$

Another important implication of this result is that a location with a higher market share $|\mathbb{T}_j|$ exhibits comparative advantage in the production of nonroutine services. This can be seen from the determination of urban diversity, a factor that generates comparative advantage in nonroutine services. Substituting (1.7), (1.12), (1.14) and (1.15) into (1.11), we obtain

$$\left(\frac{R_{j+1}}{R_j}\right)^{\beta}\left(\frac{W_{j+1}}{W_j}\right)^{1-\beta} = \left(\frac{D_{j+1}}{D_j}\right)^{\theta\gamma(T_j)}, \text{ or } \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{\alpha(1-\beta)+\beta} = \left(\frac{D_{j+1}}{D_j}\right)^{\theta\gamma(T_j)}$$

for all j. That is, if a sorting equilibrium exists, more severe market competition, i.e., higher wage rage and land rent, in a location with a higher market share $|\mathbb{T}_j|$ are associated with greater urban diversity D_j, leading the location to exhibit comparative advantage in the production of nonroutine services. We here observe an *agglomeration force* represented by urban diversity D_j. Therefore, we can see that in this model, there is a close relationship between these agglomeration and dispersion forces through regional comparative advantage, which is a result of the fragmentation of production stages across locations. In a sorting equilibrium, if it were to exist, regional disparities would emerge as a balance between the agglomeration and dispersion forces,[14] and this balance would be affected by functional specialization of locations reflecting the function $\gamma(t)$ and, thus, the distribution of the skill intensities in the economy.

Finally, by utilizing the free-entry condition for the nonroutine sector, we can derive the fundamental equation. First, note that the free entry, $\pi_j(v) = 0$ for all v and j, together with the optimal pricing rule, implies that the output $q_j(v)$ of each variety v at location j is constant, i.e., $q_j(v) = f/(\theta m) \equiv q$ for all v and j. Then, the market clearing condition for the nonroutine sector yields $D_j p_{n,j} q = (1-\alpha)\Gamma_j|\mathbb{T}_j|E$, or taking the ratio of this equation, we obtain

$$\frac{D_{j+1}}{D_j} = \frac{\Gamma_{j+1}}{\Gamma_j}\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\left(\frac{p_{n,j+1}}{p_{n,j}}\right)^{-1} = \frac{\Gamma_{j+1}}{\Gamma_j}\left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{(1-\alpha)(1-\beta)} \qquad \forall j.$$

[14] Note that the equality between agglomeration and dispersion forces holds at a threshold T_j. Since $R_{j+1}/R_j, W_{j+1}/W_j, D_{j+1}/D_j > 1$ and $\gamma(t)$ is increasing in t, compared with location j, location $j+1$ has comparative advantage in stage $t > T_j$ due to dominance of agglomeration force over dispersion force implying lower unit costs of conducting more skill-intensive tasks.

To avoid this arbitrariness of the finiteness of J, we investigate the distribution of cities in a sorting equilibrium. We accomplish this by making J diverge to infinity while limiting the total measure of cities to unity.[15] Then, applying the Matsuyama's (2013) method to the fundamental equation,[16] we obtain the following boundary value problem for a second-order ordinary differential equation (ODE):

$$\frac{\Phi''}{\Phi'} = \frac{\theta\gamma'(\Phi)\Phi'}{G(\Phi)} \quad \text{with } \Phi(0) = 0 \text{ and } \Phi(1) = 1, \tag{1.16}$$

where

$$G(\Phi) \equiv \alpha(1-\beta) + \beta - (1-\alpha)(1-\beta)\theta\gamma(\Phi).$$

Each city is characterized by the stage t that it hosts because as J diverges to infinity, $|\mathbb{T}_j|$ converges to zero or, stated differently, $\mathbb{T}_j$ converges to a point that characterizes one of the cities. Here, $\Phi(t)$ is the Lorenz curve of the market share that corresponds to $\sum_{k=1}^{j} |\mathbb{T}_k|$ for some j. Thus, given the uniformity of stage t, $\Phi'(t)$ corresponds to $|\mathbb{T}_j|$, the market share of a location. In the following, given the one-to-one correspondence between a city and a stage, we call the city that hosts stage t city t. We assume that $G(1) > 0$ in order to focus on a meaningful case.[17]

1.3.2 Existence and Uniqueness of Endogenous Rankings

In order to prove the existence of a sorting equilibrium in the limiting case, it suffices to show that there exists a solution to the fundamental equation (1.16). Importantly, we can obtain a unique solution to the fundamental equation analytically, which also implies the uniqueness of a sorting equilibrium. The economic interpretation of this result is as follows: although cities may differ, the associated variations in city characteristics are limited to a range that is consistent with the unique distribution. Since cities are ex ante identical, we cannot identify which stage of production each city specializes in ex post.[18]

More specifically, we obtain the inverse function of the Lorenz curve (let $H : z \to t$ denote the

---

[15] Note that the unit endowment of land is fixed for each location.

[16] Essentially, the method involves interpreting $1/J$ as a differential dt when J is sufficiently large and then applying the asymptotic expansion. Here, it is crucial that as J diverges to infinity, each city hosts only one stage of production in the limit and thus is characterized by t.

[17] Intuitively, this assumption implies that the magnitude of market competition, i.e., the power $\alpha(1-\beta)+\beta$ appearing in the fundamental equation, is larger than that of the agglomeration force, $(1-\alpha)(1-\beta)\theta\gamma(t)$, net of the effect of the average skill intensity Γ_j, thus resulting in "bounded" city sizes. It can be easily shown that as $G(1)$ converges to zero from above, the max-min ratio of population diverges to infinity.

[18] This resembles the implication of a stochastic process that is specified by a Markov chain with a non-degenerate unique invariant distribution. That is, the realization of a random variable varies randomly in a manner that is consistent with the invariant distribution. Importantly, this uniqueness of the sorting equilibrium allows us to conduct the numerical exercises discussed later. If we have a cross-sectional dataset of cities, we can calibrate the model.

function, i.e., $H(z) \equiv \Phi^{-1}(z)$):

$$H(z) = \frac{\int_0^z G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du}{\int_0^1 G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du} \quad \forall z \in [0,1]. \tag{1.17}$$

For a given $\gamma(t)$, this equation yields a unique inverse Lorenz curve $H(z)$. Given that $H'(z) > 0$ and $H''(z) < 0$ for all t, $\Phi(t)$ is unique and has a property: $\Phi'(t) > 0$ and $\Phi''(t) > 0$.

Then, using this result and normalizing the economy-wide income E to unity without loss of generality, we can establish the following proposition:

Proposition 1.1. *Suppose that $\gamma'(t) > 0$ and $G(1) > 0$. Then, a sorting equilibrium, characterized by a Lorenz curve of the market size $\Phi(t)$, exists and is unique. In addition, this equilibrium has the following properties: The market share $E(t)$ of city t is given by $\Phi'(t)$, and population $N(t)$, land rent $R(t)$, wage rate $W(t)$, diversity $D(t)$ and the average establishment size $\zeta(t)$ (in the nonroutine sector) in city t are given by*

$$\begin{aligned}
N(t) &= \frac{\Phi'(t)^{1-\alpha}}{\int_0^1 \Phi'(t)^{1-\alpha} dt}, \\
R(t) &= [\alpha(1-\beta)+\beta]\Phi'(t), \\
W(t) &= (1-\alpha)(1-\beta)\int_0^1 \Phi'(t)^{1-\alpha} dt \Phi'(t)^{\alpha}, \\
D(t) &= \frac{\theta}{(1+\theta)f}\frac{1-\alpha}{[\alpha(1-\beta)+\beta]^{\beta}}\left[(1-\alpha)(1-\beta)\int_0^1 \Phi'(t)^{1-\alpha} dt\right]^{-(1-\beta)} \gamma(\Phi(t))\Phi'(t)^{(1-\alpha)(1-\beta)}, \\
\zeta(t) &= \frac{(1+\theta)f}{\theta}(1-\beta)[\alpha(1-\beta)+\beta]^{\beta}\left[(1-\alpha)(1-\beta)\int_0^1 \Phi'(t)^{1-\alpha} dt\right]^{-\beta} \Phi'(t)^{(1-\alpha)\beta}
\end{aligned}$$

for all $t \in [0,1]$. Therefore, as t increases, i.e., as a location specializes in a more nonroutine stage of production, the values of all of these variables increase.

Proof. The market share $\Phi'(t)$ is simply a definition. Population function $N(t)$ follows if we apply the asymptotic expansion to (1.13):

$$\frac{N(t+\Delta t)}{N(t)} = \left[1 + \frac{\Phi''(t)}{\Phi'(t)}\Delta t + o(|\Delta t|)\right]^{1-\alpha} = 1 + (1-\alpha)\frac{\Phi''(t)}{\Phi'(t)}\Delta t + o(|\Delta t|).$$

Arranging this result and making Δt converge to zero, we obtain

$$\frac{N'(t)}{N(t)} = (1-\alpha)\frac{\Phi''(t)}{\Phi'(t)} \quad \forall t,$$

which, together with the normalization, i.e., $\int_0^1 N(t)dt = 1$, implies the desired result of $N(t)$. The land rent function $R(t)$ follows immediately from the land and labor market clearing conditions, where $|\mathbb{T}_j|$ is now replaced with $\Phi'(t)$. The wage function $W(t)$ also follows from the labor market clearing condition with $|\mathbb{T}_j|$ replaced with $\Phi'(t)$ and the population function $N(t)$. The diversity function $D(t)$ derives from the market clearing condition for nonroutine services, i.e., $D_j p_{n,j} q = (1-\alpha)\Gamma_j|\mathbb{T}_j| = (1-\alpha)\int_{T_{j-1}}^{T_j}\gamma(t)dt$, together with the results for the land rent and wage rate functions. Here, $\int_{T_{j-1}}^{T_j}\gamma(t)dt$ is replaced with $\gamma(\Phi(t))\Phi'(t)$. The establishment size function $\zeta(t)$ is given by the consistency, i.e., the labor compensation calculated by $W(t)D(t)\zeta(t)$ must be equal to the market size $(1-\alpha)\Phi'(t)$ times the skill-intensity $\gamma(\Phi(t))$ times the labor share $1-\beta$. The statement that all of these functions are increasing in t follows from the result that $\gamma'(t), \Phi'(t), \Phi''(t) > 0$. □

As argued in the previous subsection, the result that $R'(t) > 0$ derives from competition in the land rental markets. The comovement between $W(t)$ and $N(t)$ is an implication of spatial equilibrium, and comparative advantage leads to the result that $D'(t) > 0$. $\zeta'(t) > 0$ because the differential of the wage rate times the number of firms in the nonroutine sector is less than proportional to that of the market size of the sector.[19]

1.3.3 Size Distribution of Cities

As has been argued in the literature, the upper tail of the population size of cities in the United States is well approximated by a power law or, more specifically, a Pareto distribution with a coefficient of one, the so-called Zipf's law (Gabaix and Ioannides, 2004; Gabaix, 2009). Economic mechanisms resulting in Zipf's law have also been proposed, ranging from random growth models such as Rossi-Hansberg and Wright (2007) and Duranton (2007) to static models such as Hsu (2012) and Behrens et al. (2014a).

The purpose of this subsection is therefore to relate our model to the size distribution of cities. More specifically, we derive the necessary and sufficient condition under which the associated sorting equilibrium exhibits a power law including Zipf's law as a special case. The next proposition states that this is equivalent to identifying the functional form of $\gamma(t)$ guaranteeing that the size distribution of cities obeys a power law.[20]

[19] This is easily understood in the discrete version. Using the market clearing condition $W_j D_j \zeta_j = (1-\alpha)(1-\beta)\Gamma_j|\mathbb{T}_j|E$, we get

$$\frac{W_{j+1}}{W_j}\frac{D_{j+1}}{D_j}\frac{\zeta_{j+1}}{\zeta_j} = \frac{\Gamma_{j+1}}{\Gamma_j}\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}.$$

Then, using the differentials derived in the previous subsection, we obtain

$$\frac{\zeta_{j+1}}{\zeta_j} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{(1-\alpha)\beta} \quad \forall j.$$

[20] One might interpret this kind of approach as searching for a knife-edge condition and thus conclude that the result presented below is not robust compared with previous theories. Testing the model, however, is not a purpose of this chapter. Proposition is simply used for calibration of the model (Subsection 1.3.4).

Proposition 1.2. *The size distribution of cities in the sorting equilibrium is characterized by a truncated Pareto distribution if and only if $\gamma(t)$ is given by*

$$\gamma(t) = \begin{cases} a - \{a^{\eta} - [a^{\eta} - (a-1)^{\eta}]\,t\}^{\frac{1}{\eta}} & \text{if } \eta \in \left(-\frac{1}{(1-\alpha)(1-\beta)}, 0\right) \cup (0, +\infty) \\ a - \exp\{\ln a - [\ln a - \ln(a-1)]\,t\} & \text{if } \eta = 0, \end{cases}$$

where

$$a \equiv \frac{\alpha(1-\beta)+\beta}{(1-\alpha)(1-\beta)\theta} > 1.$$

Furthermore, if $\eta = -\alpha/[(1-\alpha)(1-\beta)]$, the size distribution of cities is consistent with Zipf's law.

Proof. Only the essence is discussed here. The first part of the proposition is proven in four steps: The first step is to notice, from Proposition 1.1, that $N(t)$ obeys a power law if and only if $\Phi'(t)$ obeys a power law.

The second step is to show that $\Phi'(t)$ follows a power law if and only if λ defined by $\lambda \equiv G(\Phi(t))^{-1}$ obeys a power law. In order to prove this statement, we begin by differentiating $t = H[\Phi(t)]$ with respect to t to obtain $1 = H'[\Phi(t)]\Phi'(t)$. Using (1.17), we then obtain

$$\Phi'(t) \propto G(\Phi(t))^{-\frac{1}{(1-\alpha)(1-\beta)}}.$$

The third step is to prove that λ obeys a power law if and only if

$$\gamma'\left[\gamma^{-1}(B(\lambda))\right] \propto \lambda^{\tilde{\eta}}, \quad B(\lambda) \equiv \frac{\alpha(1-\beta)+\beta-\lambda^{-1}}{(1-\alpha)(1-\beta)\theta}. \tag{1.18}$$

Here, $\tilde{\eta}$ is defined by $\tilde{\eta} = \eta - 1$. In order to validate this statement, we first note that the density function $f_Z(z)$ of $z = \Phi(t)$ is given by $f_Z(z) = f_T[H(z)]H'(z) = H'(z)$, where the first equality uses the relationship between t and z, i.e., $t = H(z)$, and the second uses the uniformity of stage t. Then, using the relationship between λ and z given by $\lambda = G(z)^{-1}$ and the density function $f_Z(z)$, we obtain the density function $f_\Lambda(\lambda)$ of λ as follows:

$$f_\Lambda(\lambda) = f_Z\left[\gamma^{-1}(B(\lambda))\right]\frac{\partial}{\partial\lambda}\gamma^{-1}(B(\lambda)) \propto \frac{\lambda^{-\left[\frac{1}{(1-\alpha)(1-\beta)}+2\right]}}{\gamma'\left[\gamma^{-1}(B(\lambda))\right]}.$$

The final step is to show that (1.18) holds if and only if $\gamma(t)$ is given by the one specified in the proposition.

The second part of the proposition is demonstrated using the results discussed above and the fact that a random variable X_1 given by $X_1 = X_2^{\omega}$ ($\omega > 0$), where X_2 follows a Pareto distribution with coefficient $\delta > 0$, obeys a Pareto distribution with coefficient δ/ω. □

α	β	σ	$\tilde{a}$	η
0.095	0.085	5.877	1.022	-0.561

Table 1.1: Calibrated Parameters

1.3.4 Welfare Gain from Functional Specialization

Benchmark

As a numerical exercise, we compute the sorting equilibrium and the welfare gain from functional specialization for an example of parameter values, where the welfare gain is defined as the difference in log-utility between the laissez-faire case described in the previous subsections and the case where there are continuum of locations all of which hosts the whole range of stages of production *and* the same number of workers.

For this purpose, we specify the functional form of $\gamma(t)$ as follows:

$$\gamma(t) = \begin{cases} \tilde{a} - \{\tilde{a}^{\eta} - [\tilde{a}^{\eta} - (\tilde{a}-1)^{\eta}]\, t\}^{\frac{1}{\eta}} & \text{if } \eta \neq 0, \\ \tilde{a} - \exp\{\ln \tilde{a} - [\ln \tilde{a} - \ln(\tilde{a}-1)]\, t\} & \text{if } \eta = 0, \end{cases}$$

where $\tilde{a} > 1$ is a parameter that is different from a in Proposition 20. Thus, instead of assuming an exact power law, here we consider a slightly general $\gamma(t)$ that includes an exact power law as a special case. Based on a guess that the upper tail of market sizes $E(t) = \Phi'(t)$ of MSAs is well approximated by a power law, we expect that the above specification would successfully match the data.[21] A benchmark set of parameter values we use is presented in Table 1.1, which is obtained by the calibration described in Appendix .7.

Figure 1.2 depicts the equilibrium Lorenz curve $\Phi(t)$ and the associated profiles of population $N(t)$, the wage rate $W(t)$, the land rent $R(t)$, urban diversity $D(t)$, and the average establishment $\zeta(t)$. As predicted by Proposition 1.1, all functions are monotonically increasing in t.

Letting $\bar{U}$ and $\bar{U}_s$ denote utilities in the laissez-faire and symmetric cases, respectively, the welfare gain is given by[22]

$$\ln \bar{U} - \ln \bar{U}_s = \left\{ \ln \left[\frac{E(t)/N(t)}{R(t)^{\alpha}} \right] - \ln \left(\frac{1}{R_s^{\alpha}} \right) \right\} - (1-\alpha)\left(\ln P - \ln P_s\right),$$

where the first and the second terms represent the effects of the differences in the land-rent-adjusted average income and the price index of the final good, respectively. Variables with subscript A cor-

[21] Note that since the population $N(t)$ of city t is proportional to the market size $E(t)$ of city t (Proposition 1.1), if the distribution of population is closed to a power law, then the market size is also close to the power law.

[22]Since both the total population and the measure of the continuum of locations are normalized to unity, the population size of each location in the symmetric case is equal to unity. In addition, given the normalization of the economy-wide income, i.e., $E = 1$, the per capita income of each location is also equal to one. Therefore, the symmetric outcome is equivalent to full agglomeration, i.e., workers and stages of production all concentrate in a single location.

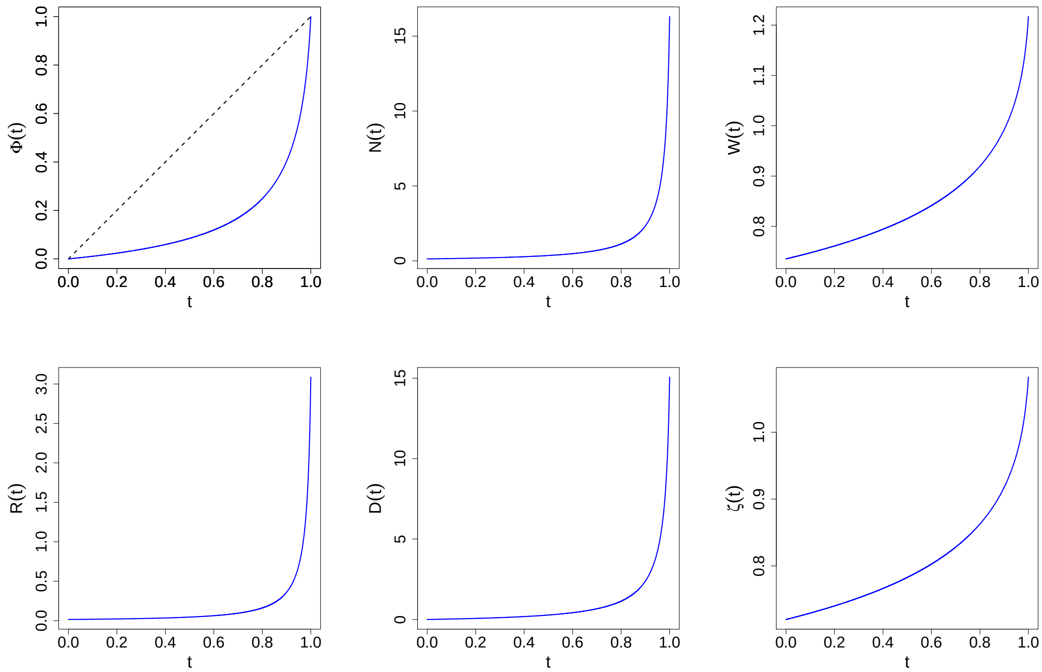

Figure 1.2: Lorenz Curve $\Phi(t)$ and Profiles, $N(t)$, $W(t)$, $R(t)$, $D(t)$, $\zeta(t)$

respond to those in the symmetric case. Note that spatial equilibrium implies that the rent-adjusted income $E(t)/[N(t)R(t)^{\alpha}]$ is independent of t. Here, the price index is further decomposed as follows:

$$\begin{aligned}\ln P-\ln P_s &= \int_0^1 \beta[\ln R(t)-\ln R_s]\Phi'(t)dt+\int_0^1 (1-\beta)[\ln W(t)-\ln W_s]\Phi'(t)dt \\ &\quad -\int_0^1 \theta\left[\gamma(\Phi(t))\ln D(t)-\Gamma_s\ln D_s\right]\Phi'(t)dt,\end{aligned}$$

where

$$\begin{aligned}&R_s=\alpha(1-\beta)+\beta, \quad W_s=(1-\alpha)(1-\beta),\\ &D_s=\frac{\theta(1-\alpha)\Gamma_s}{(1+\theta)f}R_s^{-\beta}W_s^{-(1-\beta)}, \quad \Gamma_s\equiv\int_0^1\gamma(z)dz.\end{aligned}$$

Table 1.2 reports that the welfare gain from functional specialization is about 1.6% of the utility level in the symmetric case, implying about 40% increase in life-time utility with log utility and the discount factor of 0.96.

The decomposition shows that the effect of the final good price slightly dominates that of the rent-adjusted income. In each of the two effects, increases in land rents due to agglomeration play important roles, i.e., reduce the rent-adjusted income and increase the price index of the final good. However, a much larger positive effect, 18.8% increase in utility, is generated by urban diversity, i.e.,

Item	Value (%)
Welfare Gain	1.6
Rent-adjusted Income	-8.7
Final Good Price	10.3
Land Rent	-7.8
Wage Rate	-0.7
Urban Diversity	18.8

Table 1.2: Welfare Gain from Functional Specialization: Benchmark

Note: The welfare effect of each item is measured in terms of % increase in utility from the symmetric case.

agglomeration increases the number of varieties of professional services.

Robustness

How does this result depend on the elasticity σ of substitution and the distribution of skill intensities of stages of production? In order to answer this question, fixing the share parameters, α and β, and $\tilde{a}$, a parameter less crucial than η in determining the distribution of skill intensity, we compute the welfare gain for each pair (σ,η) or (σ,Γ_s) given the one-to-one correspondence between η and Γ_s. Since $G(1) > 0$ in Proposition 1.1 must hold, we need the following restriction to σ:

$$\sigma > \max\left\{1, \frac{1}{\alpha(1-\beta)+\beta}\right\} \approx 5.817.$$

The upper bound for σ is set to 10 which corresponds to the upper bound of the typical range of the elasticity of substitution in the literatures (Anderson and van Wincoop, 2004). The lower and upper bounds, -0.685 and 9.210, for η are set in such a way that the average skill intensity Γ_s belongs to a range $[0.1, 0.9]$, which covers the benchmark case, 0.882. Figure 1.3 depicts the density of skill intensity for three values of the average skill intensity Γ_s, showing that as Γ_s increases, the composition of production processes shifts toward those with higher skill intensity.

Figure 1.4 depicts the welfare gain and its decomposition for each pair (σ,Γ_s) of the elasticity of substitution and the average skill intensity as well as the benchmark case represented as a point $(5.877, 0.882)$. Although the absolute magnitude of each decomposed term changes significantly depending on parameter values, its relative magnitude is the same as in the benchmark case, and the magnitude of the overall welfare gain is robustly high, e.g., in a large region of (σ,Γ_s), the total welfare gain is more than 1%, implying 25% increase in life-time utility.

Numerical comparative statics also shows that for a fixed average skill intensity Γ_s, an increase in the elasticity σ of substitution reduces the welfare gain from functional specialization; and that for a fixed elasticity σ of substitution, there is an inverted-U-shaped relationship between the average

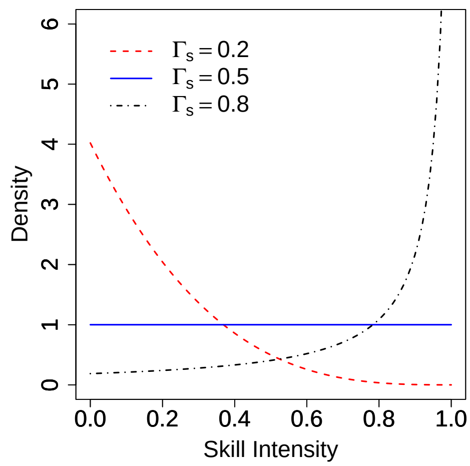

Figure 1.3: Density Function of Skill Intensity: $\Gamma_s = 0.2, 0.5, 0.8$

skill intensity Γ_s and the overall welfare gain. The former result is interpreted that as an increase in the elasticity σ of substitution weakens the urban diversity effect, resulting in a lower overall welfare gain. The other effects, i.e., rent-adjusted income, land rent, and wage rate, are all in the opposite direction reflecting weaker market competition. The latter result is simply because as the density of skill intensity concentrates on either the upper or lower corners, i.e., $\Gamma_s \to 1$ or $\Gamma_s \to 0$, heterogeneity across production processes with respect to skill intensity disappears, leading to less benefit from specialization or trade. This interpretation is supported by the result that each decomposed term has the same inverted-U-shaped relationship. However, we note that the absolute magnitude of each effect is strengthened when the density of skill intensity is skewed to the right.

1.4 Welfare Analysis of Income Redistribution Policy

Unlike standard urban economics models of a system of cities à la Henderson (1974) and Black and Henderson (1999), there is no competitive city developers or local governments that are often though of as an institution that internalizes technological externalities and coordinates agents' location choices. This implies that a market outcome is in general inefficient in the model. More specifically, given that there is no technological externalities, coordination failure among firms determining the location of stages of production and potential entrants into the monopolistic competition (non-routine

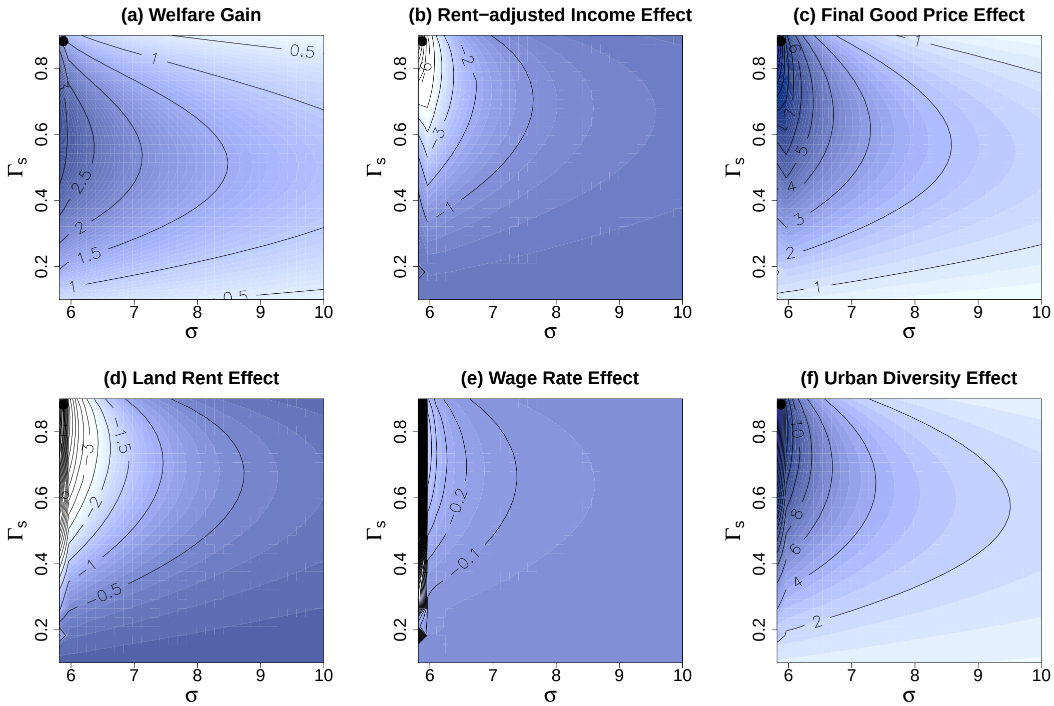

Figure 1.4: Welfare Gain from Functional Specialization

Note: Panels show the welfare effects in terms of % increase in utility from the symmetric case. A darker color within a fixed panel means a higher welfare gain. Colors in different panels are not comparable. The point $(5.877, 0.882)$ represents the benchmark case. For a given triplet $(\alpha, \beta, \tilde{a})$, there is a negative relationship between η and the average skill intensity Γ_s.

services) sector could be a source of inefficiency in the model (Matsuyama, 1995).[23] Simultaneity of these choices generates circular causation allowing inefficient concentration of economic activity:[24] If stages of production with higher skill intensities concentrate in a particular location, then the demand for non-routine services increases, leading to more entrants into the non-routine services sector there. Since the larger number of varieties reduces the price index of non-routine services in that location, the concentration of stages with higher skill intensities strengthens further.

Given the above inefficiency, this section considers a simple income redistribution policy, where the government levies a proportional income tax on workers and then redistributes the revenue across cities with a lump-sum transfer. Under this policy, a market equilibrium is determined such that workers and firms optimally choose locations of residence and production, and potential firms in the

[23] As discussed in Dixit and Stiglitz (1977) and more recently argued by Dhingra and Morrow (2012) in a class of models with heterogeneous firms à la Melitz (2003), monopolistic competition with a constant elasticity of substitution is itself not a source of inefficiency. The optimal allocation requires a "social markup" that finances fixed costs of production of a variety, and the social markup coincides with the "private" markup exactly with a constant elasticity of substitution.

[24] Although free migration is itself not a source of inefficiency, it could affect the degree of inefficiency due to coordination failure, since migration changes prices based on which firms make their decisions.

monopolistic competition sector enter the market if it is profitable to enter the market. The motivation of this focus is to evaluate effectiveness of a simple, feasible income redistribution in improving welfare instead of designing a skill-intensity-dependent tax or subsidy system that is complicated and seems infeasible in reality due to the government's imperfect knowledge of firms' technologies.

For this purpose, we describe the specific form of the policy is described in Subsection 1.4.1 and then show that the equilibrium is still characterized as a first-order difference/differential equation as in the laissez-faire case in Subsection 1.4.2. Finally, we conduct numerical comparative statics of the optimal tax rate and welfare in Subsection 1.4.3.

1.4.1 Income Redistribution Policy Rule

We begin with a finite number of locations J. Let $E_{b,j}$ and $E_{a,j}$ denote before- and after- tax regional income, respectively, i.e., $E_{b,j} = N_j(W_j + \bar{R}_j) = W_j N_j + R_j$. Then, let us introduce a government that implements the following income redistribution policy rule:[25][26]

$$E_{a,j} = (1-\tau)E_{b,j} + \frac{\tau E}{J} \quad \forall j,$$

where $\tau \in [0,1]$ is the proportional income tax rate; the case where $\tau = 0$ corresponds to the laissez-faire economy described and analyzed in the previous sections. That is, by levying a tax on individuals' incomes $E_{b,j}$ that are distributed by the market, the policy makes the equilibrium outcome more equalized than in the laissez-faire case. In addition, the lump-sum transfers among cities make this dispersion effect more effective because as individuals concentrate in a city, the *per capita lump-sum transfer within the city*, i.e., $\tau E/(JN_j)$, decreases.

In order to reflect this government policy rule, the land market clearing and free-migration conditions should be modified as

$$\begin{aligned} R_j &= (1-\alpha)\beta|\mathbb{T}_j|E + \alpha E_{a,j}, \\ \frac{E_{a,j+1}/N_{j+1}}{E_{a,j}/N_j} &= \left(\frac{R_{j+1}}{R_j}\right)^{\alpha}, \end{aligned}$$

[25] The government here cannot control the distribution of population across cities directly. Instead, the choice variable of the government is the income tax rate τ only. For a given tax rate, the population distribution of cities is implicitly determined by free migration expressed as the equalization of utility across cities. This is in contrast to the analysis in Pflüger and Tabuchi (2010) where the government controls the population distribution across (two) locations, and lump-sum transfers are automatically implied by the free-migration condition.

[26] An optimal allocation would be attained by the benevolent central government choosing taxes and subsidies in such a way that they depend on the distribution of economic activity including stages of production and workers. Therefore, although the location-dependent per-capita net transfer introduced by the central government is partially consistent with this view, an optimal allocation cannot be inferred from the policy. The purse here is to provide an example of how to analyze a model with income redistribution and to quantify the welfare gain from a simple income redistribution scheme. A more detailed analysis is a future work, which requires more complicated mathematics such as variational calculus.

where $E_{a,j}/N_j$ represents the after-tax per capita income at location j, implying

$$\begin{aligned}\frac{R_{j+1}}{R_j} &= \frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_{j+1}|+\alpha\tau J^{-1}}{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_j|+\alpha\tau J^{-1}},\\ \frac{N_{j+1}}{N_j} &= \frac{(1-\alpha)(1-\tau)|\mathbb{T}_{j+1}|+\tau J^{-1}}{(1-\alpha)(1-\tau)|\mathbb{T}_j|+\tau J^{-1}}\left\{\frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_{j+1}|+\alpha\tau J^{-1}}{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_j|+\alpha\tau J^{-1}}\right\}^{-\alpha}, \quad \forall j.\end{aligned}$$

1.4.2 Modified Fundamental Equation

Using the modified equilibrium conditions, we obtain the following modified fundamental equation:

$$\begin{aligned}&\left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{1-\beta[1+\theta\gamma(T_j)]}\left\{\frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_{j+1}|+\alpha\tau J^{-1}}{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_j|+\alpha\tau J^{-1}}\right\}^{[\alpha(1-\beta)+\beta][1+\theta\gamma(T_j)]}\\ &\times\left[\frac{(1-\alpha)(1-\tau)|\mathbb{T}_{j+1}|+\tau J^{-1}}{(1-\alpha)(1-\tau)|\mathbb{T}_j|+\tau J^{-1}}\right]^{-(1-\beta)[1+\theta\gamma(T_j)]} = \left(\frac{\Gamma_{j+1}}{\Gamma_j}\right)^{\theta\gamma(T_j)} \qquad \forall j\in\{1,2,\cdots,J-1\}.\end{aligned}$$

Replacing J^{-1} with Δt as J becomes sufficiently large and using the method of asymptotic expansion, we obtain the fundamental equation in the limiting case:

$$g(\Phi,\Phi')\Phi'' = \theta\gamma'(\Phi)\Phi',$$

where

$$\begin{aligned}g(\Phi,\Phi') &\equiv \frac{1-\beta[1+\theta\gamma(\Phi)]}{\Phi'}+\frac{[\alpha(1-\beta)+\beta][1+\theta\gamma(\Phi)]}{\Phi'+\tilde{\tau}_1}-\frac{(1-\beta)[1+\theta\gamma(\Phi)]}{\Phi'+\tilde{\tau}_2},\\ \tilde{\tau}_1 &\equiv \frac{\alpha\tau}{(1-\alpha)[\beta+\alpha(1-\beta)(1-\tau)]}, \qquad \tilde{\tau}_2\equiv\frac{\tau}{(1-\alpha)(1-\tau)}.\end{aligned}$$

With $\tau = 0$, this fundamental equation reduces to the one in the laissez-faire case.

This case is clearly complicated relative to the previous one, and thus we must resort to a numerical method such as the fourth-order Runge-Kutta method. It should also be noted that a solution to this ODE is not necessarily a sorting equilibrium because, unlike in the previous case, we do not have any analytical characterization stating that $\Phi(t)$ is a positive, strictly convex function. Instead, after solving the fundamental equation numerically, we need to check whether the function $\Phi(t)$ has this property.[27] However, an advantageous property of this ODE is that it is very easy to check the

[27] Using the land rent and wage rate functions stated in the next subsection, we have

$$P_L(t)\propto\left[(1-\alpha)(1-\tau)+\tau/\Phi'(t)\right]^{-(1-\beta)}\left\{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]\Phi'(t)+\alpha\tau\right\}^{\alpha(1-\beta)+\beta} \quad \text{for all } t\in[0,1],$$

which implies that the price $P_L(t)$ of the routine good and thus market competition are increasing in t if ($\Phi'(t)>0$ and) $\Phi''(t)>0$ for all $t\in[0,1]$. Then, the argument in Subsection 1.3.1 suggests that if a solution to the fundamental equation exists, we should have a $D(t)$ that is increasing in t, which is the implication of endogenous comparative advantage of cities with higher t in the nonroutine sector. That is, if a solution to the fundamental equation exists and if $\Phi'(t)>0$ and $\Phi''(t)>0$ for all $t\in[0,1]$, the solution is actually a sorting equilibrium.

uniqueness of a solution for a given set of parameters. This is because we know that $\Phi(t)$ is a Lorenz curve, and thus, $\Phi'(0)$ must be less than one. In addition, we have boundary conditions, $\Phi(0) = 0$ and $\Phi(1) = 1$. Therefore, by simply discretizing the interval $(0,1)$ and using each point as an initial guess for $\Phi'(0)$, we can obtain all the possible solutions to the ODE with the help of forward shooting.

1.4.3 Optimal Income Tax Rate

Benchmark

For the benchmark set of parameter values in Table 1.1, it was verified that for each $\tau \in [0,1]$, there is a unique solution to the fundamental equation that has the property $\Phi'(t), \Phi''(t) > 0$ for all t, implying that the solution is actually a sorting equilibrium.

Given the unique solution, the equilibrium utility is then calculated by

$$\ln \bar{U} = \ln \left[\frac{E_a(t)}{N(t)R(t)^{\alpha}} \right] - (1-\alpha)\ln P \qquad \forall t \in [0,1],$$

where

$$\begin{aligned}
E_a(t) &= [1-\alpha(1-\tau)]^{-1}[(1-\alpha)(1-\tau)\Phi'(t)+\tau], \\
\ln P &= \int_0^1 [\beta \ln R(t) + (1-\beta)\ln W(t) - \theta\gamma(\Phi(t))\ln D(t)]\,\Phi'(t)dt, \\
R(t) &= [1-\alpha(1-\tau)]^{-1}\left\{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]\Phi'(t)+\alpha\tau\right\}, \\
W(t) &= (1-\alpha)(1-\beta)\frac{\Phi'(t)}{N(t)}, \\
N(t) &= e^{c_0}\left[(1-\alpha)(1-\tau)\Phi'(t)+\tau\right]\left\{(1-\alpha)\left[\alpha(1-\beta)(1-\tau)+\beta\right]\Phi'(t)+\alpha\tau\right\}^{-\alpha}, \\
D(t) &= \frac{\theta}{(1+\theta)f}(1-\alpha)\gamma(\Phi(t))\Phi'(t)R(t)^{-\beta}W(t)^{-(1-\beta)}
\end{aligned}$$

for all $t \in [0,1]$, where e^{c_0} is computed by integrating $N(t)$ over $[0,1]$ and using the normalization condition, i.e., $\int_0^1 N(t)dt = 1$. The implication of spatial equilibrium is that the first term in the above equation is independent of location or task t such that utility is equalized across locations.[28]

Panel (a) in Figure 1.5 then depicts the consequent relationship between tax rate τ and the natural logarithm of the equilibrium utility $\ln \bar{U}$, which contains a unique peak at the optimal income tax rate of about 0.6%.[29] Therefore, the laissez-faire outcome, i.e., the case of $\tau = 0$, is not efficient. Its allocation is Pareto-dominated by the sorting equilibrium with a positive but sufficiently low income tax rate. The optimal tax rate achieves about 0.09% higher equilibrium utility as reported in Table

[28] Using the normalization condition on the population size, i.e., $\int_0^1 N(t)dt = 1$, it is easily shown that $E_a(t)/[N(t)R(t)^{\alpha}] = [1-\alpha(1-\tau)]^{-(1-\alpha)}e^{-c_0}$ for all $t \in [0,1]$.

[29] This small number is not a computational error as suggested by Panel (a) in Figure 1.5 and Figure 1.6, the latter of which shows the optimal tax rate for each pair (σ, Γ_s) of the elasticity of substitution and the average skill intensity.

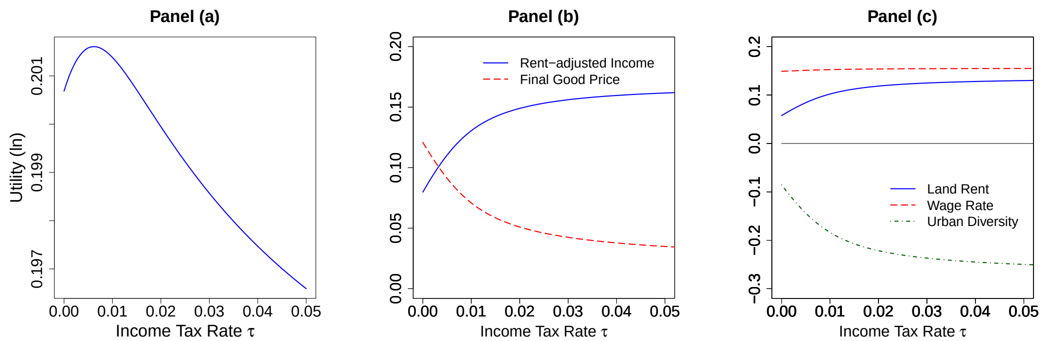

Figure 1.5: Income Tax τ and the Equilibrium Utility (ln): Benchmark

Note: The vertical axes of all panels are measured in terms of utility. For ease of visualization, the horizontal axis covers only a part of τ, i.e., $[0, 0.05]$. For higher τ's, the relative magnitude of each term does not change.

Item	Value (%)
Welfare Gain	0.09
Rent-adjusted Income	3.63
Final Good Price	-3.53
Land Rent	3.18
Wage Rate	0.24
Urban Diversity	-6.95

Table 1.3: Welfare Gain from Optimal Income Redistribution Policy: Benchmark

Note: The welfare effect of each item is measured in terms of % increase in utility from the laissez-faire case.

1.3, which corresponds to about 2.25% increase in life-time utility with log utility and the discount factor of 0.96.

The unique peak is a result of the combination of two relationships: (i) the welfare effect of the land-rent-adjusted income is monotonically increasing in the tax rate τ; and (ii) the welfare effect of the final good price is monotonically decreasing in the tax rate τ. Panel (a) and (b) in Figure 1.5 suggest that for a sufficiently low tax rate, the former effect dominates, while the latter does for a sufficiently high tax rate. As the decomposition of the welfare effect of the final good price shows (Panel (c) in Figure 1.5), the latter relationship derives from the fact that the welfare effect of urban diversity increases as the spatial distribution of economic activity concentrates, i.e., as the tax rate τ decreases. The wage rate and the land rent both become lower when economic activity is dispersed, i.e., τ is higher, suggesting that the former relationship simply reflects the relationship between the welfare effect of the land rent and the tax rate. The decomposition of the welfare gain from the optimal income tax reported in Table 1.3 confirms such relative magnitudes of welfare effects.

Robustness

Numerical comparative statics shows that the laissez-fair outcome is inefficient in the most of an empirically plausible parameter range. The left panel in Figure 1.6 depicts the optimal income tax rate corresponding to each pair (σ, Γ_s) of the elasticity of substitution and the average skill intensity in percentage terms.[30] It shows that the optimal tax rate is positive in the range, implying that the laissez-fair is characterized as excess agglomeration. This result agree with the one reported by Pflüger and Tabuchi (2010), who, using a New Economic Geography model, argue that with land use in both consumption and production, the spatial configuration is characterized by either efficient or excess agglomeration.

Although the optimal tax rate changes depending on parameter values, we observe systematic variations: For a fixed elasticity σ of substitution, the optimal tax rate is monotonically decreasing in the average skill intensity Γ_s and approaching to zero as Γ_s increases to one; and, for a fixed average skill intensity Γ_s, the optimal tax rate is monotonically increasing in the elasticity σ of substitution if Γ_s is sufficiently high and has an inverted-U-shaped relationship otherwise (the right panel in Figure 1.6). The former result is interpreted as the combination of two key results. On one hand, if the average skill intensity Γ_s is sufficiently close to unity, the density of skill intensity is highly skewed to the right (Figure 1.3), making almost every production processes to use inputs produced by the monopolistically competitive sector and thus suggesting that the laissez-faire is close to an efficient allocation.[31] On the other hand, as the average skill intensity Γ_s decreases, production technology exhibits constant returns more, and, given limited land areas generating a dispersion force, it is implied that the optimal allocation is more dispersed than the laissez faire.

The first part of the latter result is interpreted in a similar way to the former case. That is, as the elasticity σ of substitution decreases, production technology exhibits constant returns to a greater extent, favoring more dispersed economic activity as the optimal. In the second part, where the average skill intensity Γ_s is not close to unity, the laissez-faire outcome is distant from an efficient allocation, implying more room for income redistribution. This is especially relevant for lower σ which is associated with a greater concentration of economic activity and thus sever negative effects of higher land rents on welfare. Then income redistribution can mitigate those negative effects even if losing opportunities urban diversity provides. This interpretation is consistent with the result that the region of the elasticity of substitution, where the optimal tax rate is decreasing in σ, expands as the average skill intensity Γ_s decreases (the left panel of Figure 1.6). The analysis here suggests that

[30] We note that the analysis here is not exhaustive in the sense that we focus on a range of (σ, Γ_s), where computational errors associated with numerical solutions to the modified fundamental equation and numerical integration used when calculating equilibrium utility are sufficiently small such that the objective function, i.e., equilibrium (log) utility as a function of the income tax rate τ, is relatively well approximated with cubic spline interpolation, an approximation method used together with the golden section search when computing the optimal tax rate.

[31] If $\gamma(t) = 1$ for all $t \in [0,1]$, there is no specialization of cities, and the associated market outcome is equivalent to an autarkic outcome.Then, according to the literature (Dixit and Stiglitz, 1977; Dhingra and Morrow, 2012), an equilibrium allocation is efficient.

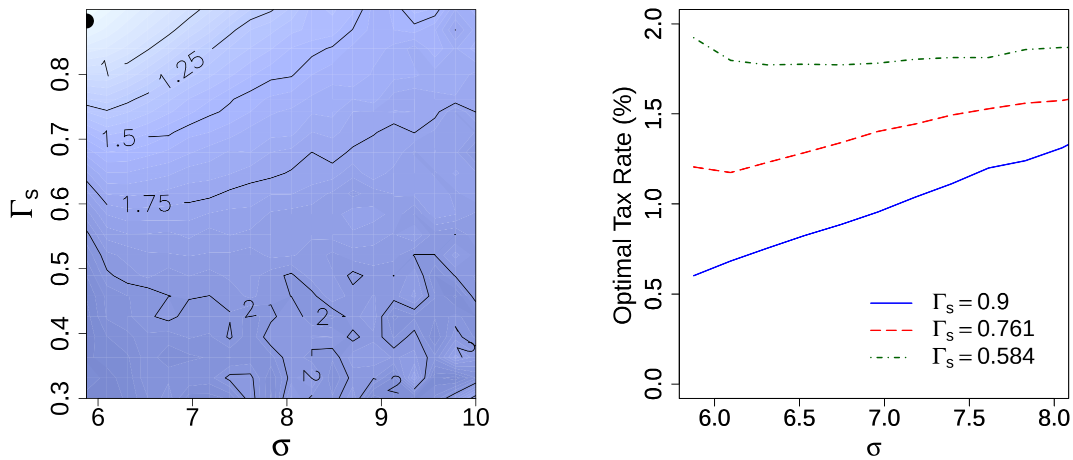

Figure 1.6: Optimal Income Tax Rate, $100 \times \tau$ (%)

Note: Left panel: A darker color means a higher tax rate; The point $(5.877, 0.882)$ represents the benchmark case in which case the optimal tax rate is about 0.6%; In the south east region, computational errors are relatively large due to a smaller role of monopolistic competition implying equilibrium utility is less elastic with respect to parameters (σ, Γ_s). Right panel: The range of σ above 8 is excluded because computational errors are relatively large.

the optimal income tax rate is determined by the interplay between two opposing components: the distance of production technology from the dominance of monopolistic competition in terms of the distribution of skill intensity and that from the dominance of constant returns to scale in terms of the elasticity of substitution.

The above results then provide an important implication for a modern developed economy: As Michaels et al. (2013) shows, we observe an increasing importance of interactive activities in developed countries such as the United States. In our model, this change in the nature of economic activity can be interpreted in two ways, i.e., a right shift of the density of skill intensity or an increase in the degree of product differentiation. The comparative statics then suggests that the desirable income redistribution policies are not necessarily the same in both case.

As for the welfare gain from income redistribution, numerical results show that the relative magnitude of each decomposed effect is the same as in the benchmark case, implying that although the absolute magnitude of each effect varies depending on parameter values, a positive overall welfare gain from income redistribution is a result of weakened negative effects of land rents compensating lower urban diversity. Panel (a) in Figure 1.7 shows the overall welfare gain from the optimal income redistribution policy, measured in terms of the increase in the equilibrium utility from the laissez-faire, is positive but limited to a similar degree as the benchmark case,[32] being in contrast to decomposed

[32]This result is consistent with that in Desmet and Rossi-Hansberg (2013) who use a totally different, neo-classical framework with Marshallian externalities in both production and preference, showing that eliminating efficiency differences

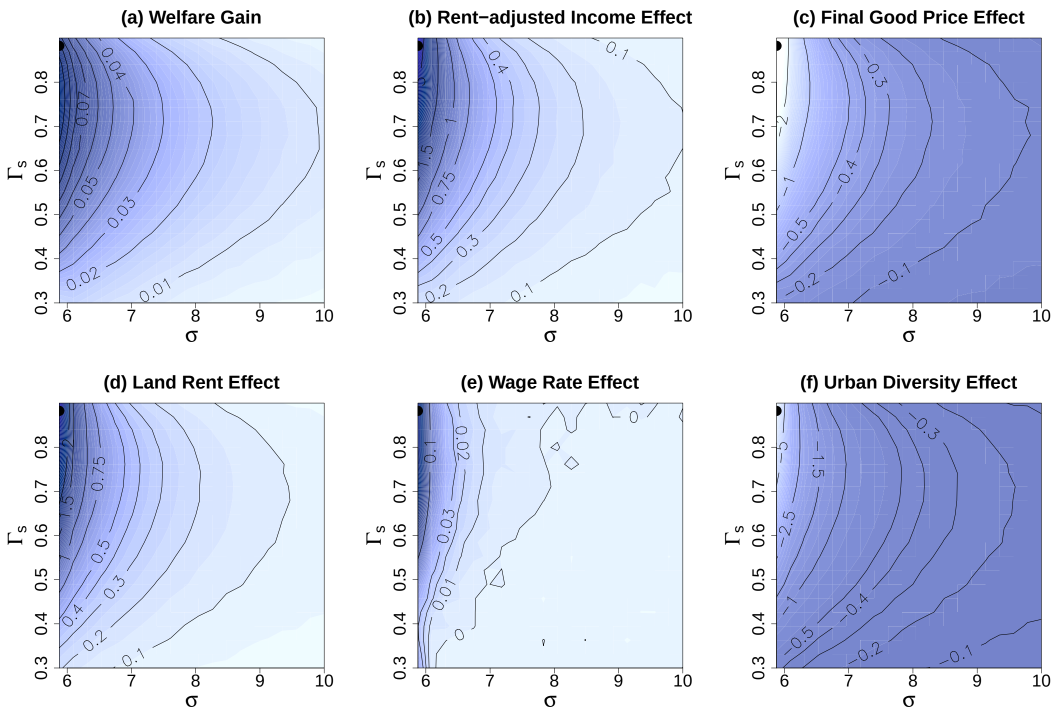

Figure 1.7: Welfare Gain from Optimal Income Redistribution Policy and Optimal Tax Rate τ

Note: Panels show the welfare effects of an optimal tax in terms of % increase in utility from the laissez-faire. A darker color within a fixed panel means a higher welfare gain or a higher tax rate. Colors in different panels are not comparable. The point $(5.877, 0.882)$ represents the benchmark case. For a given triplet $(\alpha, \beta, \tilde{a})$, there is a negative relationship between η and the average skill intensity Γ_s.

effects, Panel (b)-(f), which vary widely depending on parameter values. However, signs and relative magnitudes of those effects are the same as in the benchmark case. We also note that the welfare gain from income redistribution has the same property as that from functional specialization, i.e., gains tend to be larger when the distribution of skill intensity is close to a uniform.

1.5 Conclusion

To formalize functional specialization of cities, this chapter develops a static equilibrium model of a system of cities in which ex ante identical locations specialize ex post in different sets of stages of production, resulting in a unique, non-degenerate size distribution of cities with the comovement of income, population, the wage rate, the land rent, the average establishment size in the the non-routine local services sector, and urban diversity as observed for the U.S. cities. The model is fairly tractable in that the necessary and sufficient condition for the city size distribution to obey a power

across cities results in 1.2% increase in the U.S.

law is analytically obtained. The analysis then takes a step further to analyses of welfare gains from functional specialization and optimal income redistribution. Even if focusing on a static environment, the welfare gain from functional specialization is large, but that from income redistribution is limited. The latter analysis also provides an important implication of an increasing importance of interactive activities in a modern developed economy for desirable income redistribution.

Given its simplicity, we hope that the model serves as a tool for further analyses of issues related to specialization in a system of cities. Although not pursued in this chapter, the model seems to be easily extended to introduce externalities both in preference and production, elements often employed by recent quantitative studies such as Desmet and Rossi-Hansberg (2013). In addition, the model can also accommodate exogenous amenity or productivity differentials, which are also assumed in previous studies, at least in two extreme cases: the ordering of cities in terms of skill intensity is exactly the same as in the case of exogenous differentials; or it is exactly the contrary. The former case is used as an approximation if exogenous amenities are important in determining the city size distribution, e.g., Behrens et al. (2014b), while the latter could be used if interested in the "first nature v.s. second nature" type of argument. Intermediate cases are complicated to analyze because there are multiple distribution in equilibrium.

In addition to the above extensions, conducting an empirical analysis is also one direction of future research. Examples include investigating the model's prediction for rankings of variables more rigorously by controlling individual heterogeneity and differences across cities in the composition of industries so as to ask whether the rankings are still observed under an environment consistent with the model.

In any case, extensions allow us to more deeply understand the relationship between optimal income redistribution and the above elements as well as that between comparative advantage and agglomeration and dispersion forces.

Appendices

.1 Why the ordering, $0 = T_1 < T_2 < \cdots < T_J = 1$, holds

The following argument shows that the ordering of thresholds assumed in the text is only the relevant case.

1. Suppose without loss of generality that $|\mathbb{T}_j| < |\mathbb{T}_{j+1}|$ for all j.[33] Then, we can show that $\mathbb{T}_j \cap \mathbb{T}_{j'} = \emptyset$ for all $j \neq j'$. This is because if $\mathbb{T}_j \cap \mathbb{T}_{j'} \neq \emptyset$ for some $j \neq j'$, then it must hold that

$$\frac{P_j(t)}{P_{j'}(t)} = \left(\frac{P_{n,j}}{P_{n,j'}}\right)^{\gamma(t)} \left(\frac{P_{r,j}}{P_{r,j'}}\right)^{1-\gamma(t)} = 1 \qquad \forall t \in \mathbb{T}_j \cap \mathbb{T}_{j'}.$$

   Since $\gamma(t)$ is strictly increasing, this is possible only if

$$\frac{P_{n,j}}{P_{n,j'}} = \frac{P_{r,j}}{P_{r,j'}} = 1.$$

   However, this contradicts the fact that $|\mathbb{T}_j| \neq |\mathbb{T}_{j'}|$, given that

$$\frac{P_{r,j}}{P_{r,j'}} = \left(\frac{R_j}{R_{j'}}\right)^{\beta} \left(\frac{W_j}{W_{j'}}\right)^{1-\beta} = \left(\frac{|\mathbb{T}_j|}{|\mathbb{T}_{j'}|}\right)^{\alpha(1-\beta)+\beta}.$$

2. Furthermore, it holds that $\cup_j \mathbb{T}_j = [0,1]$. This is simply due to the fact that for any fixed $t \in [0,1]$, there exists j that minimizes the costs of processing stage t.

3. Therefore, we can say that there exists a sequence of thresholds $\{T_j\}$, implying that we can use the condition for comparative advantage, i.e., Condition 5 in the definition of an equilibrium. It is then implied that $\Gamma_j < \Gamma_{j+1}$ for all j.

4. Since $\mathbb{T}_j \cap \mathbb{T}_{j'} = \emptyset$ for any $j \neq j'$ and it must hold that $|\mathbb{T}_j| < |\mathbb{T}_{j+1}|$ and $\Gamma_j < \Gamma_{j+1}$ for all j, the sequence $\{T_j\}$ must be increasing. Otherwise, there exists $j < j'$, i.e., $|\mathbb{T}_j| < |\mathbb{T}_{j'}|$, such that $\Gamma_j > \Gamma_{j'}$, which contradicts $|\mathbb{T}_j| < |\mathbb{T}_{j'}|$ under the assumption that $(1-\alpha)(1-\beta)\theta\gamma(t) < \alpha(1-\beta)+\beta$ for all $t \in [0,1]$.

.2 Derivation of the Fundamental Equation

Reproduce the fundamental equation corresponding to the discrete case:

$$\left(\frac{T_{j+1} - T_j}{T_j - T_{j-1}}\right)^{\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma(T_j)} = \left(\frac{\Gamma_{j+1}}{\Gamma_j}\right)^{\theta\gamma(T_j)}, \tag{19}$$

[33] Of course, it is possible that $|\mathbb{T}_j| = |\mathbb{T}_{j'}|$ for some j and j'. However, if a shock to the economy is drawn from some continuous distribution, such a case occurs with probability zero.

where

$$\Gamma_j \equiv \frac{1}{T_j - T_{j-1}} \int_{T_{j-1}}^{T_j} \gamma(t)dt.$$

In the limit, i.e., $J \to \infty$, each location specializes in a single stage of production. Therefore, each location is characterized by its stage $t \in [0,1]$. So, let $\Phi(t)$ denote the income Lorenz curve, i.e., $\Phi(t)$ is equal to the accumulated income of locations that belong to $[0,t]$.

Using this Lorenz curve, we then have the following approximation:

$$\frac{T_{j+1} - T_j}{T_j - T_{j-1}} = \frac{\Phi(t+\Delta t) - \Phi(t)}{\Phi(t) - \Phi(t-\Delta t)}, \tag{20}$$

where t and Δt correspond to j/J and $1/J$, respectively. Using the asymptotic expansion, the right hand side is expressed by

$$\frac{\Phi(t+\Delta t) - \Phi(t)}{\Phi(t) - \Phi(t-\Delta t)} = 1 + \frac{\Phi''(t)}{\Phi'(t)}\Delta t + o(|\Delta t|), \tag{21}$$

where $o(\cdot)$ denotes the *little o*, i.e., as $|\Delta t| \to 0$, $o(|\Delta t|)/|\Delta t| \to 0$.

We also have the following approximation:

$$\begin{aligned}
\Gamma_{j+1} &= \frac{\int_{\Phi(t)}^{\Phi(t+\Delta t)} \gamma(\tau)d\tau}{\Phi(t+\Delta t) - \Phi(t)} = \gamma(\Phi(t)) + \frac{1}{2}\gamma'(\Phi(t))\Phi'(t)\Delta t + o(|\Delta t|), \\
\Gamma_j &= \frac{\int_{\Phi(t-\Delta t)}^{\Phi(t)} \gamma(\tau)d\tau}{\Phi(t) - \Phi(t-\Delta t)} = \gamma(\Phi(t)) - \frac{1}{2}\gamma'(\Phi(t))\Phi'(t)\Delta t + o(|\Delta t|).
\end{aligned}$$

Then,

$$\frac{\Gamma_{j+1}}{\Gamma_j} = 1 + \frac{\gamma'(\Phi(t))\Phi'(t)}{\gamma(\Phi(t))}\Delta t + o(|\Delta t|). \tag{22}$$

Thus, substituting (20)-(22) into (19), we obtain

$$\begin{aligned}
& \left[1 + \frac{\Phi''(t)}{\Phi'(t)}\Delta t + o(|\Delta t|)\right]^{\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma(\Phi(t))} = \left[1 + \frac{\gamma'(\Phi(t))\Phi'(t)}{\gamma(\Phi(t))}\Delta t + o(|\Delta t|)\right]^{\theta\gamma(\Phi(t))} \\
\Longrightarrow \quad & 1 + [\alpha(1-\beta) + \beta - (1-\alpha)(1-\beta)\theta\gamma(\Phi(t))]\frac{\Phi''(t)}{\Phi'(t)}\Delta t + o(|\Delta t|) \\
& \qquad = 1 + \theta\gamma(\Phi(t))\frac{\gamma'(\Phi(t))\Phi'(t)}{\gamma(\Phi(t))}\Delta t + o(|\Delta t|) \\
\Longrightarrow \quad & [\alpha(1-\beta) + \beta - (1-\alpha)(1-\beta)\theta\gamma(\Phi(t))]\frac{\Phi''(t)}{\Phi'(t)} + \frac{o(|\Delta t|)}{\Delta t} = \theta\gamma'(\Phi(t))\Phi'(t) + \frac{o(|\Delta t|)}{\Delta t} \\
\Longrightarrow \quad & [\alpha(1-\beta) + \beta - (1-\alpha)(1-\beta)\theta\gamma(\Phi(t))]\frac{\Phi''(t)}{\Phi'(t)} = \theta\gamma'(\Phi(t))\Phi'(t) \quad (\Delta t \to 0)
\end{aligned}$$

.3 Derivation of the Inverse Lorenz Curve H

As in the text, let $H : z \to t$ denote the inverse Lorenz curve, i.e., $H(z) \equiv \Phi^{-1}(z)$. By definition, it satisfies

$$t = H(\Phi(t)),$$

which in turn implies that

$$1 = H'(\Phi(t))\Phi'(t) \equiv H'(z)\Phi'(t). \tag{23}$$

Meanwhile, integrating the fundamental equation, we obtain

$$\Phi'(t) \propto [\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma(z)]^{-\frac{1}{(1-\alpha)(1-\beta)}}. \tag{24}$$

Substituting (24) into (23), we then obtain

$$\begin{aligned} & H'(z) \propto [\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma(z)]^{\frac{1}{(1-\alpha)(1-\beta)}} \\ \Longrightarrow \quad & H(z) = c_0 + c_1 \int_0^z G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du, \end{aligned} \tag{25}$$

where c_0 and c_1 are undetermined constants. Using the boundary conditions, $H(0) = \Phi(0) = 0$ and $H(1) = \Phi(1) = 1$, c_0 and c_1 are then determined by

$$\begin{aligned} c_0 &= 0, \\ c_1 &= \left[\int_0^1 G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du\right]^{-1}, \end{aligned}$$

implying the desired expression:

$$H(z) = \frac{\int_0^z G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du}{\int_0^1 G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du} \qquad \forall z \in [0,1].$$

.4 Proof of Proposition 2

Here, we only show that (i) $\gamma'[\gamma^{-1}(B(\lambda))] \propto \lambda^{\tilde{\eta}}$ if and only if $\gamma(t)$ is given as in Proposition 2; and that (ii) the size distribution of cities is consistent with Zipf's law if and only if $\eta = -\alpha/[(1-\alpha)(1-\beta)]$.

.4.1 Determination of the Functional Form of $\gamma(t)$

Define x by

$$x \equiv \frac{\alpha(1-\beta)+\beta-\lambda^{-1}}{(1-\alpha)(1-\beta)\theta},$$

with which we can rewrite $\gamma'[\gamma^{-1}(B(\lambda))] \propto \lambda^{\tilde{\eta}}$ as follows:

$$\gamma'[\gamma^{-1}(x)] \propto (a-x)^{-\tilde{\eta}}, \quad \text{or} \quad \frac{1}{\gamma'[\gamma^{-1}(x)]} \propto (a-x)^{\tilde{\eta}}.$$

Defining a function g by $t = g(x) = \gamma^{-1}(x)$, we can interpret this as

$$g'(x) = c_1(a-x)^{\tilde{\eta}}, \qquad c_1 > 0. \tag{26}$$

Note that the function g must satisfy

$$g(0) = 0 \quad \text{and} \quad g(1) = 1 \tag{27}$$

since $\gamma(0) = 0$ and $\gamma(1) = 1$.

Now, suppose that $\tilde{\eta} = -1$. We then obtain

$$t = g(x) = c_0 - c_1 \ln(a-x) \quad \text{and thus} \quad \gamma(t) = a - \exp\left(\frac{c_0 - t}{c_1}\right). \tag{28}$$

Using the terminal condition (27), it is immediate to obtain

$$c_0 = \frac{\ln a}{\ln a - \ln(a-1)} \quad \text{and} \quad c_1 = [\ln a - \ln(a-1)]^{-1},$$

which together with (28) imply the desired result for $\eta = \tilde{\eta} + 1 = 0$.

As for the case of $\tilde{\eta} \neq -1$, we can apply the similar method and obtain the desired result.

.4.2 Determination of the Constraint on the Parameters for Zipf's Law

Under the assumption that $\eta \neq 0$, the density function of λ is given by

$$f_\Lambda(\lambda) \propto \lambda^{-\left[\frac{1}{(1-\alpha)(1-\beta)}+\eta+1\right]}.$$

This implies that λ obeys a Pareto distribution with coefficient of $1/[(1-\alpha)(1-\beta)]+\eta$.

Then, since $\Phi'(t)$ is related with λ by

$$\Phi'(t) \propto \lambda^{\frac{1}{(1-\alpha)(1-\beta)}},$$

$\Phi'(t)$ obeys a Pareto distribution with coefficient of $(1-\alpha)(1-\beta)\eta+1$.

Furthermore, since $N(t) \propto \Phi'(t)^{1-\alpha}$, $N(t)$ obeys a Pareto distribution with coefficient of $(1-\beta)\eta+1/(1-\alpha)$.

Finally, the definition of Zipf's law, i.e., a Pareto distribution with unit coefficient, pins down the desired value of η.

.5 Derivation of the Modified Fundamental Equation

Reproduce the modified fundamental equation:

$$\left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{1-\beta[1+\theta\gamma(T_j)]}\left\{\frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_{j+1}|+\alpha\tau J^{-1}}{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]|\mathbb{T}_j|+\alpha\tau J^{-1}}\right\}^{[\alpha(1-\beta)+\beta][1+\theta\gamma(T_j)]}$$
$$\times\left[\frac{(1-\alpha)(1-\tau)|\mathbb{T}_{j+1}|+\tau J^{-1}}{(1-\alpha)(1-\tau)|\mathbb{T}_j|+\tau J^{-1}}\right]^{-(1-\beta)[1+\theta\gamma(T_j)]} = \left(\frac{\Gamma_{j+1}}{\Gamma_j}\right)^{\theta\gamma(T_j)} \qquad \forall j \in \{1,2,\cdots,J-1\},$$

which contains

$$\frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_i J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_i J^{-1}}, \qquad i=1,2.$$

Because, replacing J^{-1} with Δt, we have

$$\begin{aligned}
|\mathbb{T}_{j+1}|+\tilde{\tau}_i J^{-1} &= \Phi(t+\Delta t)-\Phi(t)+\tilde{\tau}_i\Delta t = \left[\Phi'(t)+\tilde{\tau}_i\right]\Delta t+\frac{1}{2}\Phi''(t)\Delta t^2+o(|\Delta t|)\\
|\mathbb{T}_j|+\tilde{\tau}_i J^{-1} &= \Phi(t)-\Phi(t-\Delta t)+\tilde{\tau}_i\Delta t = \left[\Phi'(t)+\tilde{\tau}_i\right]\Delta t-\frac{1}{2}\Phi''(t)\Delta t^2+o(|\Delta t|),
\end{aligned}$$

with the help of the asymptotic expansion, we can approximate the above term as follows:

$$\frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_i J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_i J^{-1}} = 1+\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_i}\Delta t+o(|\Delta t|).$$

Then, substituting the following approximations into the modified fundamental equation,

$$
\begin{aligned}
\left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{1-\beta[1+\theta\gamma(T_j)]} &= \left[1+\frac{\Phi''(t)}{\Phi'(t)}\Delta t+o(|\Delta t|)\right]^{1-\beta[1+\theta\gamma(\Phi(t))]} \\
&= 1+\{1-\beta[1+\theta\gamma(\Phi(t))]\}\frac{\Phi''(t)}{\Phi'(t)}\Delta t+o(|\Delta t|), \\
\left(\frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_1 J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_1 J^{-1}}\right)^{[\alpha(1-\beta)+\beta][1+\theta\gamma(T_j)]} &= \left[1+\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1}\Delta t+o(|\Delta t|)\right]^{[\alpha(1-\beta)+\beta][1+\theta\gamma(\Phi(t))]} \\
&= 1+[\alpha(1-\beta)+\beta][1+\theta\gamma(\Phi(t))]\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1}\Delta t+o(|\Delta t|), \\
\left(\frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_2 J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_2 J^{-1}}\right)^{-(1-\beta)[1+\theta\gamma(T_j)]} &= \left[1+\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_2}\Delta t+o(|\Delta t|)\right]^{-(1-\beta)[1+\theta\gamma(\Phi(t))]} \\
&= 1-(1-\beta)[1+\theta\gamma(\Phi(t))]\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_2}\Delta t+o(|\Delta t|), \\
\left(\frac{\Gamma_{j+1}}{\Gamma_j}\right)^{\theta\gamma(T_j)} &= \left[1+\frac{\gamma'(\Phi(t))\Phi'(t)}{\gamma(\Phi(t))}\Delta t+o(|\Delta t|)\right]^{\theta\gamma(\Phi(t))} \\
&= 1+\theta\gamma(\Phi(t))\frac{\gamma'(\Phi(t))\Phi'(t)}{\gamma(\Phi(t))}\Delta t+o(|\Delta t|),
\end{aligned}
$$

and rearranging the result, we obtain the modified fundamental equation.

.6 Derivation of Profiles

.6.1 Land Rent $R(t)$

From the text, we have

$$
\begin{aligned}
\frac{R_{j+1}}{R_j} &= \frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_1 J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_1 J^{-1}}, \\
\sum_{j=1}^{J} R_j &= \frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]+\alpha\tau}{1-\alpha(1-\tau)},
\end{aligned}
$$

where we invoke the normalization of the total income, i.e., $E=1$, as in the text.

Using the asymptotic expansion, the former can be written as follows:

$$\begin{aligned}
& \frac{R(t+\Delta t)}{R(t)} = 1 + \frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1}\Delta t + o(|\Delta t|) \\
\Longrightarrow \quad & \frac{1}{R(t)}\frac{R(t+\Delta t)-R(t)}{\Delta t} = \frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1} + \frac{o(|\Delta t|)}{\Delta t} \\
\Longrightarrow \quad & \frac{R'(t)}{R(t)} = \frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1} \quad (\Delta t \to 0) \\
\Longrightarrow \quad & \ln R(t) = c_0 + \ln[\Phi'(t)+\tilde{\tau}_1] \\
\Longrightarrow \quad & R(t) = \tilde{c}_0[\Phi'(t)+\tilde{\tau}_1].
\end{aligned}$$

Integrating the last expression, we then obtain

$$\begin{aligned}
& \int_0^1 R(t)dt = \sum_{j=1}^{J} R_j = \frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]+\alpha\tau}{1-\alpha(1-\tau)} = \tilde{c}_0\left[\int_0^1 \Phi'(t)dt + \tilde{\tau}_1\right] = \tilde{c}_0(1+\tilde{\tau}_1) \\
& \tilde{c}_0 = \frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]}{1-\alpha(1-\tau)},
\end{aligned}$$

where the last equality of the first line uses the fact that $\Phi(0) = 0$ and $\Phi(1) = 1$.

Therefore, we have

$$\begin{aligned}
R(t) &= \frac{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]}{1-\alpha(1-\tau)}[\Phi'(t)+\tilde{\tau}_1] \\
&= [1-\alpha(1-\tau)]^{-1}\left\{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]\Phi'(t)+\alpha\tau\right\}.
\end{aligned}$$

.6.2 Population $N(t)$

From the text, we have

$$\frac{N_{j+1}}{N_j} = \frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_2 J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_2 J^{-1}}\left(\frac{|\mathbb{T}_{j+1}|+\tilde{\tau}_1 J^{-1}}{|\mathbb{T}_j|+\tilde{\tau}_1 J^{-1}}\right)^{-\alpha},$$

which is in turn approximated as follows:

$$\begin{aligned}
\frac{N(t+\Delta t)}{N(t)} &= \left[1+\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_2}\Delta t + o(|\Delta t|)\right]\left[1+\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1}\Delta t + o(|\Delta t|)\right]^{-\alpha} \\
&= \left[1+\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_2}\Delta t + o(|\Delta t|)\right]\left[1-\alpha\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1}\Delta t + o(|\Delta t|)\right] \\
&= 1+\left[\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_2} - \alpha\frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1}\right]\Delta t + o(|\Delta t|).
\end{aligned}$$

Integrating both sides, arranging the result and taking the limit ($\Delta t \to 0$), we then obtain

$$
\begin{aligned}
& \frac{N'(t)}{N(t)} = \frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_2} - \alpha \frac{\Phi''(t)}{\Phi'(t)+\tilde{\tau}_1} \\
\Longrightarrow \quad & \ln N(t) = \hat{c}_0 + \ln[\Phi'(t)+\tilde{\tau}_2] - \alpha \ln[\Phi(t)+\tilde{\tau}_1] \\
\Longrightarrow \quad & N(t) = e^{\hat{c}_0}\left[\Phi'(t)+\tilde{\tau}_2\right]\left[\Phi'(t)+\tilde{\tau}_1\right]^{-\alpha}, \quad \text{or} \\
& N(t) = e^{c_0}\left[(1-\alpha)(1-\tau)\Phi'(t)+\tau\right]\left\{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]\Phi'(t)+\alpha\tau\right\}^{-\alpha},
\end{aligned}
$$

where c_0 satisfies

$$
e^{\hat{c}_0} = e^{c_0}(1-\alpha)(1-\tau)\left\{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]\right\}^{-\alpha}.
$$

Here, c_0 should be consistent with the normalized total population, i.e., $\int_0^1 N(t)dt = 1$.

.6.3 Wage Rate $W(t)$

From the labor market clearing condition in the text, we have

$$
\begin{aligned}
\frac{W_{j+1}N_{j+1}}{W_j N_j} &= \frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}, \\
\sum_{j=1}^{J} W_j N_j &= (1-\alpha)(1-\beta),
\end{aligned}
$$

where we invoked the normalization of the total income, i.e., $E = 1$.

Defining the local labor income M_j by $M_j = W_j N_j$, we can approximate the above first equation as follows:

$$
\frac{M(t+\Delta t)}{M(t)} = 1 + \frac{\Phi''(t)}{\Phi'(t)}\Delta t + o(|\Delta t|).
$$

Therefore, we have

$$
\begin{aligned}
& \frac{M'(t)}{M(t)} = \frac{\Phi''(t)}{\Phi'(t)} \\
\Longrightarrow \quad & \ln M(t) = c_0 + \ln \Phi'(t) \\
\Longrightarrow \quad & M(t) = \tilde{c}_0 \Phi'(t),
\end{aligned}
$$

where $\tilde{c}_0$ is determined by integrating both sides and using the above result:

$$
\int_0^1 M(t)dt = \sum_{j=1}^{J} W_j N_j = (1-\alpha)(1-\beta) = \tilde{c}_0 \int_0^1 \Phi'(t)dt = \tilde{c}_0.
$$

This implies that

$$M(t) = W(t)N(t) = (1-\alpha)(1-\beta)\Phi'(t) \quad \Longrightarrow \quad W(t) = (1-\alpha)(1-\beta)\frac{\Phi'(t)}{N(t)}.$$

.6.4 Urban Diversity $D(t)$

Reproduce the market clearing condition for the nonroutine service sector:

$$D_j p_{n,j} q = (1-\alpha)\Gamma_j|\mathbb{T}_j| = (1-\alpha)\int_{T_{j-1}}^{T_j} \gamma(t)dt,$$

where

$$p_{n,j} = (1+\theta)mR_j^{\beta}W_j^{1-\alpha}, \qquad q = \frac{f}{\theta m}.$$

This suggests that

$$\begin{aligned} & D(t)\frac{(1+\theta)f}{\theta}R(t)^{\beta}W(t)^{1-\beta}dt = (1-\alpha)\gamma(\Phi(t))\Phi'(t)dt \\ \Longrightarrow \quad & D(t) = \frac{\theta}{(1+\theta)f}(1-\alpha)\gamma(\Phi(t))\Phi'(t)R(t)^{-\beta}W(t)^{-(1-\beta)}. \end{aligned}$$

.6.5 After-tax Local Income $E_a(t)$

From the text, we have

$$E_{a,j} = (1-\tau)E_{b,j} + \tau J^{-1} = (1-\tau)(W_jN_j + R_j) + \tau J^{-1},$$

where the normalization of the total income, i.e., $E = 1$, is used.

This then suggests that

$$\begin{aligned} E_a(t)dt &= (1-\tau)[W(t)N(t) + R(t)]dt + \tau dt \\ &= (1-\tau)[M(t) + R(t)]dt + \tau dt \\ \Longrightarrow \quad E_a(t) &= (1-\tau)[M(t) + R(t)] + \tau. \end{aligned}$$

Substituting

$$\begin{aligned} M(t) + R(t) &= (1-\alpha)(1-\beta)\Phi'(t) + [1-\alpha(1-\tau)]^{-1}\left\{(1-\alpha)[\alpha(1-\beta)(1-\tau)+\beta]\Phi'(t) + \alpha\tau\right\} \\ &= [1-\alpha(1-\tau)]^{-1}[(1-\alpha)\Phi'(t) + \alpha\tau] \end{aligned}$$

into the above equation, we obtain

$$E_a(t) = [1-\alpha(1-\tau)]^{-1}[(1-\alpha)(1-\tau)\Phi'(t)+\tau].$$

.6.6 Natural Logarithm of the Price Index $\ln P$ of the Final Good

In a sorting equilibrium, each location t specializes in a stage of production with skill intensity of $z=\Phi(t)$. Then, the Cobb-Douglas technology implies that

$$\ln P = \int_0^1 \ln[P(\Phi^{-1}(z))]dz,$$

where $P(t)$ denotes the location-t unit cost of processing stage with $z=\Phi(t)$. The change of the variable then implies

$$\ln P = \int_0^1 \ln[P(t)]\Phi'(t)dt.$$

Meanwhile, the unit cost $P(t)$ of processing a stage with skill intensity of $\Phi(t)$ is given by

$$P(t) = P_n(t)^{\gamma(\Phi(t))}P_r(t)^{1-\gamma(\Phi(t))}.$$

Substituting the unit costs, $P_n(t)$ and $P_r(t)$, of local service sectors

$$\begin{aligned} P_n(t) &= (1+\theta)mR(t)^{\beta}W(t)^{1-\beta}D(t)^{-\theta}, \\ P_r(t) &= R(t)^{\beta}W(t)^{1-\beta} \end{aligned}$$

into the equation, it follows that

$$P(t) = [(1+\theta)m]^{\gamma(\Phi(t))}R(t)^{\beta}W(t)^{1-\beta}D(t)^{-\theta\gamma(\Phi(t))},$$

which, with the help of normalization, $(1+\theta)m=1$, becomes

$$P(t) = R(t)^{\beta}W(t)^{1-\beta}D(t)^{-\theta\gamma(\Phi(t))}.$$

Therefore, we have

$$\ln P = \int_0^1 [\beta\ln R(t) + (1-\beta)W(t) - \theta\gamma(\Phi(t))D(t)]\,\Phi'(t)dt.$$

.7 Calibration

In this section, we calibrate the model to the U.S. data. We describe our procedure and the data we use in Subsection .7.1 and .7.2, respectively. The result of the calibration is then presented in Subsection .7.3. Our purpose is not testing the model. Instead, we simply intend to pick up an example of values of parameters used as a benchmark in numerical exercises conducted in Subsections 1.3.4 and 1.4.3 in such a way that the model roughly matches the data.

.7.1 Procedure

In order to compute the equilibrium numerically, we need to specify the values of the parameters $(\alpha, \beta, \sigma, \tilde{a}, \eta)$. The first two are calibrated independently, and the other by matching the model with the U.S. data in terms of the upper tail of the distribution of market sizes of MSAs. Without loss of generality, $f = 1$ and $(1+\theta)m = 1$.

Expenditure Share α of Land

Let A, B, and C denote the expenditure share of land, housing (excluding land), and goods and services, respectively, i.e., $A + B + C = 1$. According to Davis and Ortalo-Magne (2011), the housing expenditure (including land) $A + B = 0.24$. In addition, Albouy and Ehrlich (2012) report that the one-third of housing costs are due to land, implying $A = 1/3 \times (A + B)$. Since there is no housing in the model, the expenditure share α of land is then calculated by $\alpha = A/(A + C) \approx 0.095$.

Cost Share β of Land

Valentinyi and Herrendorf (2008) report sectoral income shares of land, structures, and equipments, the summation of which corresponds to the capital share. Let $\hat{A}$, $\hat{B}$, and $\hat{C}$ denote the income shares of land, structures and equipments, and labor, respectively, i.e., $\hat{A} + \hat{B} + \hat{C} = 1$. If we focus on the services sector, $\hat{A} = 0.06$, and $\hat{C} = 0.65$ according to Table 5 in Valentinyi and Herrendorf (2008). Since production factors in our model consist of land and labor only, we can calculate β by $\beta = \hat{A}/(\hat{A} + \hat{C}) \approx 0.085$.

Other Parameters $(\sigma, \tilde{a}, \eta)$

The elasticity $\sigma = 1/\theta + 1$ of substitution is calibrated in such a way that the model can exactly match the natural logarithm of the observed max-min ratio of the market size, which is given by

$$\ln\left[\frac{\Phi'(1)}{\Phi'(0)}\right]\Bigg|_{\text{data}} = \frac{1}{(1-\alpha)(1-\beta)} \ln\left[\frac{\alpha(1-\beta)+\beta}{\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta}\right].$$

Parameter	Meaning	Restriction	Target	Source
α	Expenditure share of land	$(0,1)$	Expenditure share of land	Albouy and Ehrlich (2012)
				Davis and Ortalo-Magne (2011)
β	Cost share of land	$(0,1)$	Cost share of land	Valentinyi and Herrendorf (2008)
σ	Elasticity of substitution	$(\max\{1,[\alpha(1-\beta)+\beta]^{-1}\},+\infty)$	Max-Min ratio of labor compensation	*Occupational Employment Statistics*
$\tilde{a}$	Parameter of $\gamma(t)$	$(1,+\infty)$	Distribution of labor compensations	*Occupational Employment Statistics*
η	Parameter of $\gamma(t)$	$(-\infty,+\infty)$	do.	do.

Table 4: Restrictions on and Targets of the Calibrated Parameters

As for $(\tilde{a},\eta)$, we solve the following constrained minimization problem in a brute force manner with a discretized parameter space:

$$\min_{\tilde{a},\eta} \max_{i\in\{1,2,\cdots,\hat{N}\}} \left|\hat{F}(m_i)-F(m_i;\tilde{a},\eta)\right| \quad s.t. \quad \tilde{a}>1,$$

where $\hat{F}$ and F are the distribution functions of the natural logarithm of the market size for the data and the model, respectively. $\hat{N}$ is the sample size of MSAs, and m_i is the actual normalized market size of ith MSA. Table 4 summarizes the restrictions on and targets of the calibrated parameters, where the lower bound for the elasticity σ of substitution takes account of the assumption that $G(1)>0$ in Proposition 1.1.

.7.2 Data

The data that we use in the calibration are taken from the *May 2011 Occupational Employment Statistics* compiled by the Bureau of Labor Statistics, which reports the number of employments and the average annual wage rates for occupations listed in the *2010 Standard Occupational Classification System* for each of the Metropolitan Statistical Areas (MSAs). We then exploit the equivalence between the normalized market size of an MSA and the normalized total labor compensation due to the constancy of the labor cost share $1-\beta$ in order to construct the distribution of the natural logarithm of the market size from the data on labor compensation. Focusing on the upper tail of the distribution results in the sample size of 342 MSAs which is not particularly different from those in previous studies such as Rossi-Hansberg and Wright (2007). The implied max-min ration of market sizes is 5.304.

.7.3 Result

The results of the calibration are reported in Table 1.1. The elasticity σ of substitution falls in the typical range, $[5,10]$, in the (international) literatures such as Anderson and van Wincoop (2004). As shown in Figure 8, the model replicates the observed distribution of the natural logarithm of market sizes with the maximal deviation of 2.0%.

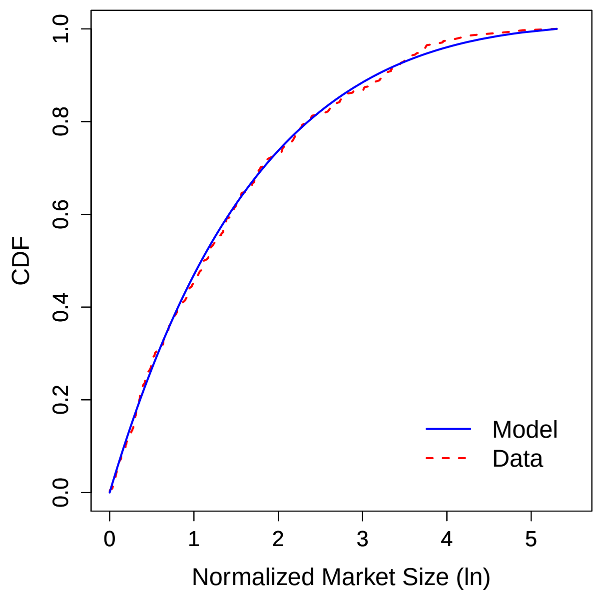

Figure 8: Cumulative Distribution Function of the Normalized Market Size

Note: The cumulative distribution of the data is constructed using labor compensations of the top 342 MSAs.

Source: May 2011 Occupational Employment Statistics.

Chapter 2

Comparative Advantage and Skill Premium of Regions[1]

2.1 Introduction

This chapter provides one explanation for why a positive correlation between the size and skill premium of a region emerges by providing a comparative advantage model with a continuum of mobile heterogeneous individuals as well as a continuum of final goods sectors that are different in terms of their skill intensities of intermediate goods. All individuals choose their occupations depending on their productivity, and any occupation can freely migrate across regions unlike footloose entrepreneur models such as Forslid and Ottaviano (2003). This location-occupation choice then interacts with the regional comparative advantage in final goods sectors which depends on the regional offer prices of two different types of intermediate goods, one of which features monopolistic competition à la Dixit and Stiglitz (1977). Although regions are ex-ante identical, interactions between individuals' location-occupation choices and regional comparative advantage result in a self-organized positive correlation between the skill premium and income of regions. The theory can also accommodate the interpretation that the regional difference in skill premium is caused by specialization in task trade within firms, not industries.

The basic mechanism is simple: Since regions are ex-ante identical in their environment including land, some initial shock or history which reallocates the economy's expenditure across regions unevenly results in a cross-region variation in land rents. Free migration of workers is then associated with a compensating differential, i.e., wage rates in regions with higher land rents must be associated with higher wage rates in order for workers to reside in such regions. Because of cross variation in factor prices, regions with higher prices have no comparative advantage in producing non-differentiated intermediate goods. However, by making the average productivity of high-skilled

[1]The content of this chapter was published in a journal (Nagamachi, 2013).

workers higher through sorting, higher land rents give such regions a comparative advantage in producing skill-intensive intermediate goods. Reflecting this regional comparative advantage, final goods sectors relocate across regions, and such relocation of industries makes the initial reallocation of expenditures sustainable. Thus, a positive correlation between skill premium and the size of regions is observed.

This chapter is related to at least two lines of research. The first concerns trade models with Ricardian comparative advantage. The current model is an application of Matsuyama (2013) to the regional context. His model is basically an extension of Dornbusch et al. (1977), where the comparative advantage of countries is determined endogenously through firms' entry into a monopolistically competitive sector and the number of countries is increased arbitrarily. Although one of his motivations is to construct a theory of income distribution across a large number of countries, I focus on a two-region case. Unlike the international economy, the regional economy is more complicated in that individuals are mobile across regions, which makes it difficult to derive the distribution of regional income explicitly. In addition to individuals' mobility, the current model differs from Matsuyama's (2013) in that individuals choose their occupations, to be either workers or entrepreneurs; that there are two types of intermediate goods sectors, one of which is characterized by monopolistic competition; and that land, which is one of the usual elements in the urban economics literature, is introduced.

The second line of research involves models of the spatial sorting of individuals. Amongst these, Davis and Dingel (2012) is the most related in the sense that it shares the same motivation and the assumption of identical cities or regions and zero trade costs of some goods. Although both studies feature a self-organized positive correlation between the skill premium and the size of regions, the key mechanism is quite different. In their paper, knowledge exchange works as an agglomeration force, while this force is regional specialization in different industries or tasks in mine.

The remainder of this chapter is organized as follows: I first introduce the model in Section 2.2. I summarize the basic mechanism working through general equilibrium and conduct a numerical exercise in order to verify it in Section 2.3. In Section 2.4, I conclude this chapter.

2.2 The Model

The economy consists of two ex-ante identical regions: Region 1 and Region 2. Each region is endowed with one unit of land, which is owned by a competitive landowner outside the economy. Individuals, the mass of which is normalized to unity, are ex-ante heterogeneous in their entrepreneurial productivity, and they choose their occupation, to be a worker or entrepreneur, depending on their productivity as well as their residential choice. There is a $[0,1]$-continuum of final goods sectors, each of which is different in its share parameters of two types of intermediate goods: labour- and skill- intensive intermediate goods.

2.2.1 Final Goods Sectors

Competitive final goods sectors exist on a $[0,1]$ interval. Each sector $s \in [0,1]$ uses a Cobb-Douglas production technology with constant returns to scale, inputs of which are local differentiated skill-intensive intermediate goods and local homogeneous labour-intensive intermediate goods. $\gamma(s) \in [0,1]$ is the share parameter of the former, and $\gamma'(s) > 0$.

The location of each sector s is determined through competition, resulting in the price $P(s)$ of sector-s final good which is equal to the lowest unit cost of production. Letting $\chi_j(s)$ and $\mathbb{S}_j \subseteq [0,1]$ denote the unit cost of production of sector s active in Region j and the set of sectors active in Region j, respectively, it holds that $P(s) = \chi_j(s)$ if $s \in \mathbb{S}_j$.

Formally, a typical firm in sector s residing in Region j, i.e., $s \in \mathbb{S}_j$, solves

$$\begin{aligned}
&\max_{\{M_{i,j}(s)\}_{i\in\{E,L\}}\{m_{i,j}(\varphi,s)\}_{\varphi}} P(s)M_{E,j}(s)^{\gamma(s)}M_{L,j}(s)^{1-\gamma(s)} \\
&\qquad\qquad - \int_0^\infty p_{E,j}(\varphi)m_{E,j}(\varphi,s)N_{E,j}g_j^*(\varphi)d\varphi - P_{L,j}M_{L,j}(s) \\
&s.t. \\
&M_{E,j}(s) = \left[\int_0^\infty m_{E,j}(\varphi,s)^{\frac{\sigma-1}{\sigma}}N_{E,j}g_j^*(\varphi)d\varphi\right]^{\frac{\sigma}{\sigma-1}}.
\end{aligned}$$

$m_{E,j}(\varphi,s)$ is the sector-s demand for a variety of skill-intensive intermediate goods produced locally by a firm with productivity φ (hereafter variety-φ skill-intensive intermediate good).[2] $N_{E,j}$ is the mass of skill-intensive intermediate goods or entrepreneurs in Region j. $g_j^*(\varphi)$ denotes the density function of productivity conditional on location. These differentiated goods are aggregated to $M_{E,j}(s)$ by technology with constant elasticity $\sigma > 1$ of substitution. $M_{L,j}(s)$ is the sector-s demand for homogeneous labor-intensive intermediate goods produced locally. Prices are denoted by $p_{E,j}(\varphi)$ and $P_{L,j}$ for variety-φ skill-intensive intermediate good and homogeneous labor-intensive intermediate goods, respectively.

Profit maximization implies the following demand for variety-φ skill-intensive good:

$$m_{E,j}(\varphi,s) = \left[\frac{p_{E,j}(\varphi)}{P_{E,j}}\right]^{-\sigma} M_{E,j}(s), \tag{2.1}$$

where $P_{E,j}$ is the price index of Region-j skill-intensive intermediate goods defined by

$$P_{E,j} \equiv N_{E,j}^{-\theta}\left[\int_0^\infty p_{E,j}(\varphi)^{-\frac{1}{\theta}}g_j^*(\varphi)d\varphi\right]^{-\theta}, \quad \theta \equiv 1/(\sigma-1). \tag{2.2}$$

[2] Subscripts E and L are used when making it explicit that variables or parameters with subscripts E and L are specific to the skill- and labor- intensive intermediate goods sectors, respectively. Subscript E is used because the skill-intensive intermediate goods are produced by entrepreneurs.

2.2.2 Labour-intensive Intermediate Goods Sectors

The local labour-intensive intermediate goods sector in each region is competitive. Firms can access a Cobb-Douglas production technology with constant returns to scale, the inputs of which consist of workers' labour services $L_{L,j}$ and land $T_{L,j}$.

$$\max_{L_{L,j},T_{L,j}} P_{L,j}BL_{L,j}^{\beta_L}T_{L,j}^{1-\beta_L} - W_jL_{L,j} - R_jT_{L,j}, \quad B \equiv \beta_L^{-\beta_L}(1-\beta_L)^{-(1-\beta_L)},$$

where $\beta_L \in (0,1)$ and R_j are the share parameter of labour and land price in Region j.

2.2.3 Skill-intensive Intermediate Goods Sectors

The local skill-intensive intermediate goods sector is characterized by monopolistic competition à la Dixit and Stiglitz (1977), where each entrepreneur produces one variety of goods using workers' labour services and land as production inputs. Specifically, each entrepreneur must rent f units of land for her office and then use workers' labour services and land as variable inputs.

Therefore, the income $\pi_j(\varphi)$ of an entrepreneur residing in Region j with productivity φ is given by her sales net of input costs:

$$\begin{aligned}
\pi_j(\varphi) \quad = \quad & \max_{p_{E,j}(\varphi),q_{E,j}(\varphi)} \left[p_{E,j}(\varphi) - W_j^{\beta_E}R_j^{1-\beta_E}\varphi^{-1}\right]q_{E,j}(\varphi) - R_jf \\
& s.t. \\
& q_{E,j}(\varphi) = \int_{\mathbb{S}_j} m_{E,j}(\varphi,s)ds = \int_{\mathbb{S}_j}\left[\frac{p_{E,j}(\varphi)}{P_{E,j}}\right]^{-\sigma}M_{E,j}(s)ds,
\end{aligned}$$

where $q_{E,j}(\varphi)$ is the output of variety-φ skill-intensive intermediate good produced in Region j. Here, it is assumed that the unit cost of production is some amount of the Cobb-Douglas composite of workers' labour services and land, in which β_E governs the labour cost share.

The associated optimal pricing rule is then $p_{E,j}(\varphi) = (1+\theta)W_j^{\beta_E}R_j^{1-\beta_E}\varphi^{-1}$. Substituting this into (2.2) results in

$$P_{E,j} = (1+\theta)W_j^{\beta_E}R_j^{1-\beta_E}\left(\tilde{\varphi}_jN_{E,j}^{\theta}\right)^{-1}, \tag{2.3}$$

where $\tilde{\varphi}_j$ is the average productivity in Region j defined by

$$\tilde{\varphi}_j = \left[\int_0^\infty \varphi^{\frac{1}{\theta}}g_j^*(\varphi)d\varphi\right]^{\theta}. \tag{2.4}$$

In the following, I assume that the entrepreneurial productivity φ follows a Pareto distribution with coefficient δ and a lower bound $\underline{\varphi}$. Under the assumption that $\delta > 1/\theta$, the individual variable profit $\pi_j^V(\varphi)$ and output $q_j(\varphi)$ are expressed as functions of the productivity ratio $\varphi/\tilde{\varphi}_j$ and the average variables as in Melitz (2003):

$$\pi_j^V(\varphi) = \left(\frac{\varphi}{\tilde{\varphi}_j}\right)^{\frac{1}{\theta}} \pi_j^V(\tilde{\varphi}_j), \qquad \pi_j^V(\tilde{\varphi}_j) = \theta W_j^{\beta_E} R_j^{1-\beta_E} \tilde{\varphi}_j^{-1} N_{E,j}^{-(1+\theta)} \int_{\mathbb{S}_j} M_{E,j}(s)ds, \tag{2.5}$$
$$q_j(\varphi) = \left(\frac{\varphi}{\tilde{\varphi}_j}\right)^{\sigma} q_j(\tilde{\varphi}_j), \qquad q_j(\tilde{\varphi}_j) = N_{E,j}^{-(1+\theta)} \int_{\mathbb{S}_j} M_{E,j}(s)ds.$$

2.2.4 Individuals

Individuals are ex-ante heterogeneous in their entrepreneurial productivity φ. Depending on this productivity, each individual chooses her occupation and location freely in order to maximize her utility. Let $U_j(\varphi)$ and $e_j(\varphi)$ denote the utility and income of an individual having the productivity of φ and residing in Region j.

Occupational Choice

Suppose that an individual chooses to reside in Region j. Then she chooses the occupation which maximizes her income. Thus, her income $e_j(\varphi)$ is given by $e_j(\varphi) = \max\{\pi_j(\varphi), W_j\}$. This suggests that there exists a cut-off level φ_j^* such that

$$W_j = \sigma^{-1}\tilde{A}_j \varphi_j^{*\frac{1}{\theta}} - R_j f, \tag{2.6}$$

where $\tilde{A}_j$ denotes the per-capita market size of the skill-intensive intermediate goods sector in Region j normalized by the regional average productivity, i.e., $\alpha\Gamma_j|\mathbb{S}_j|E/N_{E,j}$ divided by $\tilde{\varphi}^{\frac{1}{\theta}}$, which is derived in Appendix .1.2.

For the given income $e_j(\varphi)$ as well as the given location j, each individual then consumes final goods and housing services:

$$\begin{aligned} U_j(\varphi) \quad &= \quad \max_{\{c_j(s,\varphi)\}_{s\in[0,1]}, h_j(\varphi)} \exp\left[\alpha \int_0^1 \ln(c_j(s,\varphi))ds\right] h_j(\varphi)^{1-\alpha}, \quad \alpha \in (0,1), \\ & \qquad s.t. \\ & \qquad \int_0^1 P(s)c_j(s,\varphi)ds + R_j h_j(\varphi) = e_j(\varphi), \end{aligned}$$

where α is the expenditure share of the consumption goods. $c_j(s,\varphi)$ and $h_j(\varphi)$ denote the quantities of goods and housing services, respectively, consumed by an individual with φ residing in Region j.

Residential Choice

Finally, each individual chooses her location in order to maximize her utility.

The following result then states that if both regions host a positive measure of production activities, or, stated more weakly, if there exists a threshold $\bar{\varphi}$ such that entrepreneurs with productivity of

$\bar{\varphi}$ are indifferent between the two location choices, the sorting of entrepreneurs is always associated:[3]

Proposition 2.1. *Suppose that there exists $\bar{\varphi}$ such that $\bar{\varphi} > \max_j\{\varphi_j^*\}$ and*

$$u(\bar{\varphi}) = \frac{\pi_2(\bar{\varphi})/(P^\alpha R_2^{1-\alpha})}{\pi_1(\bar{\varphi})/(P^\alpha R_1^{1-\alpha})} = \frac{\pi_2(\bar{\varphi})/\pi_1(\bar{\varphi})}{(R_2/R_1)^{1-\alpha}} = 1.$$

Then, if $R_1 < R_2$, it holds that

$$\left(\frac{R_2}{R_1}\right)^{1-\alpha} < \frac{\tilde{A}_2}{\tilde{A}_1} < \frac{R_2}{R_1}, \text{ and}$$

$\pi_2(\varphi)/\pi_1(\varphi)$ is monotonically increasing in a well-defined region. Or, if $R_1 = R_2$, it must hold that $\tilde{A}_1 = \tilde{A}_2$ and thus $u(\varphi) = 1$ for all $\varphi \geq \underline{\varphi}$.

2.3 Equilibrium Analysis

2.3.1 Basic Mechanism

In the following, I focus on the case where regions are ex-post heterogeneous. Specifically, without loss of generality, I focus on equilibria in which the land rent in Region 2 is greater than that in Region 1, i.e., $R_1 < R_2$. Therefore, given Proposition 2.1, an interior equilibrium is associated with a unique threshold $\bar{\varphi}$ such that $\bar{\varphi} > \max_j\{\varphi_j^*\}$ and there is a spatial sorting of entrepreneurs. In this case, individuals with φ higher than or equal to $\bar{\varphi}$ reside in Region 2 and work as entrepreneurs. Those with φ less than $\bar{\varphi}$ but higher than or equal to φ_1^* reside in Region 1 and also work as entrepreneurs. Workers consist of individuals with φ less than φ_1^*. Since workers' income is independent of φ, the following free-migration condition or compensated differential for workers must be satisfied:

$$\frac{W_1}{P^\alpha R_1^{1-\alpha}} = \frac{W_2}{P^\alpha R_2^{1-\alpha}}, \quad \text{or} \quad \frac{W_2}{W_1} = \left(\frac{R_2}{R_1}\right)^{1-\alpha}, \tag{2.7}$$

which states that utility levels are equalized across regions.

In order to compute an equilibrium, I use the next result which is obtained immediately:

Proposition 2.2. *Suppose an asymmetric interior equilibrium exists. Then, the spatial distribution of final goods sectors is summarized by a threshold $S_1 \in (0,1)$ such that $\mathbb{S}_1 = [0, S_1)$ and $\mathbb{S}_2 = [S_1, 1]$.*

The implication of this result for the computation of an equilibrium is that the system of an equilibrium can be now interpreted as a fixed-point problem of S_1. The discussion, which is described in Appendix .1, proceeds in two steps. (i) Given the spatial distribution S_1, the system of an equilibrium is consolidated into two simultaneous equations with two unknowns: the ratio of $\bar{\varphi}$ to φ_1^* and the ratio

[3] The proof is straightforward and is thus omitted.

of φ_1^* to $\underline{\varphi}$. All other variables except S_1 are given as functions of these two and S_1. (ii) The rest of the computation is to search for an S_1 that is consistent with the comparative advantage of regions. Stated differently, S_1 must be a solution to the nonlinear equation

$$\frac{\chi_2(S_1)}{\chi_1(S_1)} = \left(\frac{P_{L,2}}{P_{L,1}}\right)^{1-\gamma(S_1)} \left(\frac{P_{E,2}}{P_{E,1}}\right)^{\gamma(S_1)} = 1 \tag{2.8}$$

and the Region 2-1 ratio $\chi_2(s)/\chi_1(s)$ of offer prices must be decreasing in s. If the latter condition does not hold, all ss greater than or equal to S_1 reside in Region 1, not Region 2, clearly contradicting the assumption.

The intuitive mechanism which can work in the model is summarized as follows: Suppose that some shock hits the economy consisting of two ex-ante identical regions in a way that expenditures concentrate on one of the regions (here Region 2), i.e., $|\mathbb{S}_1| < |\mathbb{S}_2|$. Since both regions have the same amount of land, it then holds that the land rent in Region 2 becomes higher than in Region 1, i.e., $R_1 < R_2$.[4] Because of the free migration of workers or the compensating differential, i.e., (2.7), the wage rate in Region 2 also becomes higher than that in Region 1, i.e., $W_1 < W_2$. Thus, unit costs or prices of non-differentiated goods are higher in Region 2 than in Region 1, i.e., $P_{L,1} < P_{L,2}$.[5] Instead, because of the sorting of entrepreneurs, i.e., (2.3), Region 2 has a comparative advantage in producing skill-intensive intermediate goods, i.e., $P_{E,1} > P_{E,2}$. Reflecting these regional advantages, the spatial distribution of final goods sectors settles down in such a way that the reallocation of expenditures caused by the initial shock is actually preserved as an equilibrium outcome.

2.3.2 Numerical Exercise

In order to verify the mechanism in the previous subsection, I resort to a numerical exercise. The result shows that an equilibrium with such a mechanism actually exists. It is also verified that the equilibrium is unique in the sense that there is only one interior sorting equilibrium with the assumed regional rankings of variables.

In this exercise, parameters are set as follows: the elasticity of substitution σ between skill-intensive intermediate goods is set to 3. The expenditure share α of final goods is set to 0.7. The lower bound $\underline{\varphi}$ of the Pareto distribution of entrepreneurial productivities is set to 1. The coefficient δ of the Pareto distribution is set to 4.2. The labour share parameter of the labour-intensive intermediate goods sector is set to 0.6. The same number is used for the labour share in variable costs of the skill-intensive intermediate goods sector. The fixed requirement f of land is set to 1. This value is chosen in a way that the demand of entrepreneurs for land does not substantially affect land prices, and these prices are mainly determined by housing expenditures and housing demands associated with variable

[4] Strictly speaking, this relationship between the two rankings holds only if expenditures are the most important determinant of land rents, as suggested by the land market clearing condition (18) derived in Appendix .1.

[5] Note that the price of homogeneous labour-intensive intermediate goods is a weighted geometric mean of the wage rate and land rent.

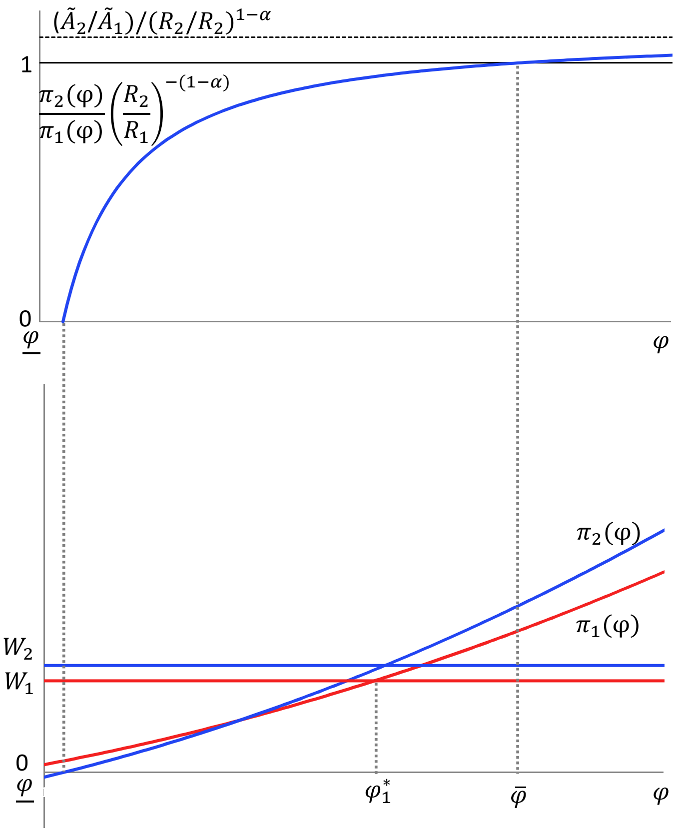

Figure 2.1: Region 2-1 Ratio Utility Conditional on Choosing to Become an Entrepreneur (The Upper Panel) and Entrepreneurs' and Workers' Incomes (The Lower Panel)

inputs. As for the specification of $\gamma(s)$, I simply assume that $\gamma(s) = s$ for $s \in [0,1]$.

The equilibrium is summarized by Figures 2.1 and 2.2. The lower panel of Figure 2.1 depicts the relationship between the entrepreneurial productivity φ and the wage and entrepreneurs' profit schedules for each region. As already mentioned, the wage schedule is flat since workers' income is independent of their entrepreneurial productivity φ. Meanwhile, entrepreneurs' income $\pi_j(\varphi)$ is monotonically increasing in φ. That $W_1 < W_2$ is implied by the free-migration condition for workers together with the ranking $R_1 < R_2$. As for $\pi_j(\varphi)$, it is not always the case that $\pi_1(\varphi) < \pi_2(\varphi)$ for all φ. What is important here is that $u(\varphi)$ is monotonically increasing in φ (Proposition 2.1) and that $u(\varphi) = 1$ at $\bar{\varphi}$, which are shown in the upper panel of Figure 2.1.

Figure 2.2 shows that the Region 2-1 ratio $\chi_2(s)/\chi_1(s)$ of offer prices is monotonically decreasing in s, and there actually exists a threshold S_1 which summarizes the spatial distribution of final goods sectors. Since $|\mathbb{S}_j|$ is proportional to the regional GDP, the result that $|\mathbb{S}_1| < |\mathbb{S}_2|$ implies that the size of Region 2 is greater than that of Region 1 in terms of income. Importantly, the monotonicity of $\chi_2(s)/\chi_1(s)$ is the consequence of two results: $P_{L,2}/P_{L,1} > 1$ and $P_{E,2}/P_{E,1} < 1$. The former result is

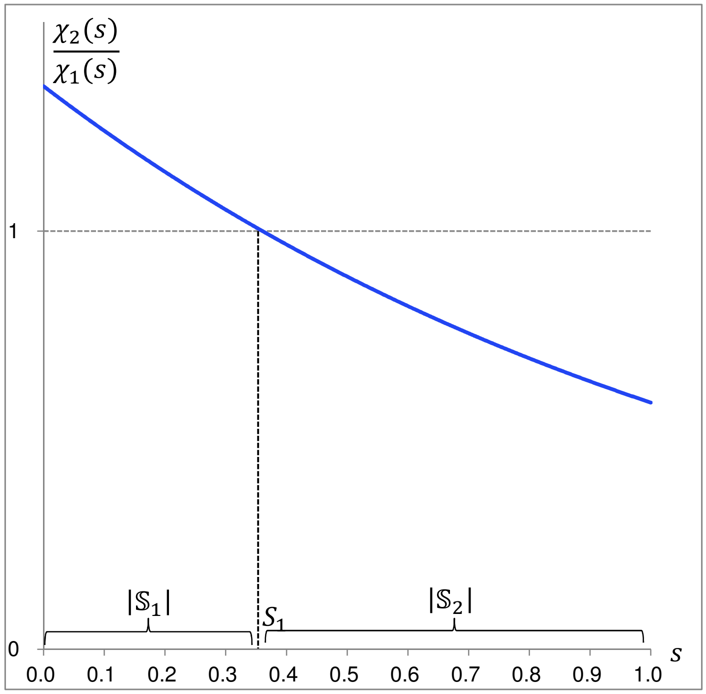

Figure 2.2: Region 2-1 Ratio of Offer Prices of Final Goods

simply due to the fact that $R_1 < R_2$ and $W_1 < W_2$ as discussed before. The latter result suggests that there actually exists a case where the cost-reducing effect of product differentiation and sorting on the aggregate price level dominates the cost push because of higher land rents and thus higher wage rates.[6]

2.4 Conclusion

A positive correlation is observed between skill premium and the size of regions, which are measured by the income ratio of high-skilled and low-skilled workers and regional income, respectively. This chapter theoretically investigates one possible explanation for this fact by providing a model with heterogeneous individuals and final and intermediate goods sectors, in which ex-ante identical regions specialize in different sectors, and interactions between individuals' location-occupation choices and regional comparative advantage result in the positive correlation between the skill premium and income of regions. The theory can also accommodate the interpretation that the regional difference in skill premium is caused by specialization in task trade, not industries. Although perfect sorting fea-

[6] It should be noted that the numerical exercise also shows that in equilibrium, Region 1 has a greater number of entrepreneurs than Region 2, i.e., $N_{E,1} > N_{E,2}$. This suggests that for the theory considered in this chapter, what is important for Region 2 to have a comparative advantage in producing skill-intensive goods is not the number of entrepreneurs, but the average productivity.

turing the equilibrium itself is not a crucial element of the theory, filling the gap between the model and reality could be an important direction for future research.

Appendices

.1 Equilibrium System as a Fixed-Point Problem of S_1

In this section, I show that the equilibrium system of an equilibrium of interest is summarized as a fixed-point problem of S_1.

.1.1 Number of Entrepreneurs, Conditional Densities, and Average Productivities

First of all, given the Pareto distribution $g(\varphi)$ of entrepreneurial productivity φ and the ranking of thresholds, i.e., $\varphi_1^* < \bar{\varphi}$, the number $\{N_{E,j}\}_{j=1}^2$ of entrepreneurs in each region, densities $\{g_j^*\}$ of productivity conditional on sorting, and average productivities $\{\varphi_j^*\}_{j=1}^2$ of regions are given as functions of thresholds $(\varphi_1^*, \bar{\varphi})$:

$$
\begin{aligned}
N_{E,1} &= G(\bar{\varphi}) - G(\varphi_1^*) = \left(\varphi_1^{*-\delta} - \bar{\varphi}^{-\delta}\right)\underline{\varphi}^{\delta}, && (9)\\
N_{E,2} &= 1 - G(\bar{\varphi}) = \left(\frac{\underline{\varphi}}{\bar{\varphi}}\right)^{\delta}, && (10)\\
g_1^*(\varphi) &= \frac{1}{N_{E,1}}\mathbf{1}\{\varphi_1^* \le \varphi < \bar{\varphi}\}g(\varphi) = \left(\varphi_1^{*-\delta} - \bar{\varphi}^{-\delta}\right)^{-1}\mathbf{1}\{\varphi_1^* \le \varphi < \bar{\varphi}\}\delta\varphi^{-(\delta+1)}, && (11)\\
g_2^*(\varphi) &= \frac{1}{N_{E,2}}\mathbf{1}\{\bar{\varphi} \le \varphi\}g(\varphi) = \bar{\varphi}^{\delta}\mathbf{1}\{\bar{\varphi} \le \varphi\}\delta\varphi^{-(\delta+1)}, && (12)\\
\tilde{\varphi}_1 &= \left(\frac{\delta}{\delta - 1/\theta}\right)^{\theta}\left[\frac{(\varphi_1^*/\bar{\varphi})^{\frac{1}{\theta}-\delta} - 1}{(\varphi_1^*/\bar{\varphi})^{-\delta} - 1}\right]^{\theta}\bar{\varphi}, \text{ or } \left(\frac{\delta}{\delta - 1/\theta}\right)^{\theta}\left[\frac{1 - (\bar{\varphi}/\varphi_1^*)^{\frac{1}{\theta}-\delta}}{1 - (\bar{\varphi}/\varphi_1^*)^{-\delta}}\right]^{\theta}\varphi_1^*, && (13)\\
\tilde{\varphi}_2 &= \left(\frac{\delta}{\delta - 1/\theta}\right)^{\theta}\bar{\varphi}, && (14)
\end{aligned}
$$

where $\mathbf{1}\{\cdot\}$ is the indicator function which is equal to one if the statement in the braces is true and zero otherwise.

.1.2 Factor Prices as Functions of Three Thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$

Next, $\tilde{A}_j$ is computed as a function of thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$ with the help of market clearing conditions: the Cobb-Douglas preference suggests that the economy-wide expenditure for final goods is given by αE, where E denotes the economy-wide income excluding land rents. If the $\mathbb{S}_j$ set of industries locates in Region j, equal weights of industries in preference and the production technology of the final goods sectors then imply that two market clearing conditions, one for the final goods and the other for the skill-intensive intermediate goods, are consolidated into

$$
\int_{\mathbb{S}_j} P_{E,j}M_{E,j}(s)ds = \alpha\Gamma_j|\mathbb{S}_j|E, \quad \text{or} \quad \int_{\mathbb{S}_j} M_{E,j}(s)ds = P_{E,j}^{-1}\alpha\Gamma_j|\mathbb{S}_j|E,
$$

where $\Gamma_j \equiv |\mathbb{S}_j|^{-1} \int_{\mathbb{S}_j} \gamma(s)ds$, implying that $\alpha\Gamma_j|\mathbb{S}_j|E = \int_{\mathbb{S}_j} \alpha E\gamma(s)ds$, the sum of expenditures for the skill-intensive intermediate goods in Region j, which in turn is related to consumers' demand. Substituting (2.3) into this equation, I get

$$\int_{\mathbb{S}_j} M_{E,j}(s)ds = (1+\theta)^{-1}\left(W_j^{\beta_H} R_j^{1-\beta_H}\right)^{-1} \tilde{\varphi}_j N_{E,j}^{\theta} \alpha\Gamma_j|\mathbb{S}_j|E.$$

Finally, substituting this equation into (2.5) results in

$$\pi_j^V(\tilde{\varphi}) = \sigma^{-1}\frac{\alpha\Gamma_j|\mathbb{S}_j|E}{N_{E,j}},$$

which gives

$$\tilde{A}_j = \frac{\alpha\Gamma_j|\mathbb{S}_j|E}{\tilde{\varphi}_j^{\frac{1}{\theta}} N_{E,j}}. \tag{15}$$

That is, $\tilde{A}_j$ is the normalized average market size of skill-intensive intermediate goods in Region j. Note that given (9)-(14) and Proposition 2.2, $\tilde{A}_j$ is a function of three thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$.

This derivation of $\tilde{A}_j$ is useful for the computation of factor prices $\{(W_j, R_j)\}_{j=1}^2$ in relating labour and land market clearing conditions with thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$, to which I turn next.

Since the sales of an entrepreneur with productivity φ are $\tilde{A}_j\varphi^{\frac{1}{\theta}}$ and since the variable profit $\pi_j^V(\varphi) = \sigma^{-1}\tilde{A}_j\varphi^{\frac{1}{\theta}}$, the variable cost is equal to $(1+\theta)^{-1}\tilde{A}_j\varphi^{\frac{1}{\theta}}$. Thus the Cobb-Douglas technology implies that the associated variable labour and land costs are given by $\beta_h(1+\theta)^{-1}\tilde{A}_j\varphi^{\frac{1}{\theta}}$ and $(1-\beta_H)(1+\theta)^{-1}\tilde{A}_j\varphi^{\frac{1}{\theta}}$, respectively. Factor market clearing conditions, which aggregate these firm-level costs, then pin down factor prices and the spatial distribution of workers.

The labour market clearing condition for each region is given as follows:

$$N_{E,j}\int_{\underline{\varphi}}^{\infty} \beta_H(1+\theta)^{-1}\tilde{A}_j\varphi^{\frac{1}{\theta}} g_j^*(\varphi)d\varphi + \beta_L\alpha(1-\Gamma_j)|\mathbb{S}_j|E = W_j\lambda_j N_W,$$

where $\lambda_j \in (0,1)$ denotes the share of Region j in workers, and N_W the total number of workers, i.e., $N_W = G(\varphi_1^*) = 1-(\underline{\varphi}/\varphi_1^*)^{\delta}$. Together with (2.4) and (15), the first term becomes $\beta_H(1+\theta)^{-1}\alpha\Gamma_j|\mathbb{S}_j|E$, i.e., the clearing condition simplifies to

$$\beta_H(1+\theta)^{-1}\alpha\Gamma_j|\mathbb{S}_j|E + \beta_L\alpha(1-\Gamma_j)|\mathbb{S}_j|E = W_j\lambda_j N_W \quad \text{for all } j = 1,2. \tag{16}$$

The second term on the left-hand side is the demand from the labour-intensive sector, where the total sales $\alpha(1-\Gamma_j)|\mathbb{S}_j|E$ are derived in a similar way as in the case of the skill-intensive intermediate goods sector, and the Cobb-Douglas technology then implies that a β_L fraction of these must be distributed to workers.

Thus, noting that N_W is a function of φ_1^* and that both Γ_j and $|\mathbb{S}_j|$ are functions of S_1, the labour market clearing condition together with the free-migration condition for workers, i.e., (2.7), gives the wage rate and the spatial distribution of workers as functions of two thresholds (φ_1^*, S_1) and the land rents ratio R_2/R_1:

$$W_1 = \frac{\tilde{\beta}_1 A_1}{(1-\lambda_2)N_W}, \quad W_2 = W_1\left(\frac{R_2}{R_1}\right)^{1-\alpha}, \quad \lambda_2 = \frac{\frac{\tilde{\beta}_2 A_2}{\tilde{\beta}_1 A_1}}{\frac{\tilde{\beta}_2 A_2}{\tilde{\beta}_1 A_1} + \left(\frac{R_2}{R_1}\right)^{1-\alpha}}, \quad \lambda_1 = 1-\lambda_2,$$

where

$$A_j \equiv \alpha|\mathbb{S}_j|E, \quad \tilde{\beta}_j \equiv \Gamma_j \frac{\beta_H}{1+\theta} + (1-\Gamma_j)\beta_L.$$

As for the land market clearing condition, an argument similar to that in the case of the labour market gives land prices as a function of three thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$: the demands for land consists of not only those from firms in both skill-intensive and labour-intensive sectors but also those from individuals, i.e., $(1-\alpha)E_j$, where E_j is Region-j income excluding land rents given by

$$E_j = N_{E,j}\int_{\underline{\varphi}}^{\infty} \pi_j(\varphi) g_j^*(\varphi) d\varphi + W_j \lambda_j N_W = \frac{\theta}{1+\theta}\alpha\Gamma_j|\mathbb{S}_j|E - R_j f N_{E,j} + W_j\lambda_j N_W. \tag{17}$$

Noting that the demands from the skill-intensive sector are further divided into those related to variables costs and those related to fixed costs, the market clearing condition is specified by

$$\begin{aligned} R_j &= N_{E,j}\int_{\underline{\varphi}}^{\infty} (1-\beta_H)(1+\theta)^{-1}\tilde{A}_j \varphi^{\frac{1}{\theta}} g_j^*(\varphi) d\varphi + R_j f N_{E,j} + (1-\beta_L)(1-\Gamma_j)\alpha|\mathbb{S}_j|E + (1-\alpha)E_j, \\ &= (1-\beta_H)(1+\theta)^{-1}\alpha\Gamma_j|\mathbb{S}_j|E + R_j f N_{E,j} + (1-\beta_L)(1-\Gamma_j)\alpha|\mathbb{S}_j|E + (1-\alpha)E_j, \end{aligned}$$

where the second equation follows from the definitions of $\tilde{\varphi}_j$ and $\tilde{A}_j$, i.e., (2.4) and (15). Together with the labour market clearing condition (16) and equation (17) for the local income E_j, this equation is solved for R_j in order to interpret R_j as a function of three thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$:

$$R_j = \frac{1}{1-\alpha f N_{E,j}}\eta_j A_j, \quad \text{where} \quad \eta_j \equiv \Gamma_j \frac{1-\alpha\beta_H + (1-\alpha)\theta}{1+\theta} + (1-\Gamma_j)(1-\alpha\beta_L). \tag{18}$$

Given this result, wage rates $\{W_j\}_{j=1}^2$ and the spatial distribution $\{\lambda_j\}_{j=1}^2$ of workers are now functions of three thresholds $(\varphi_1^*, \bar{\varphi}, S_1)$.

.1.3 Productivity Thresholds $(\varphi_1^*, \bar{\varphi})$ as Functions of S_1

I now show that two productivity thresholds $(\varphi_1^*, \bar{\varphi})$ are functions of S_1. For this purpose, two conditions are used: one is for φ_1^* and the other for $\bar{\varphi}$. The first condition states that an individual with productivity φ_1^* is indifferent between becoming a worker and working as an entrepreneur in Region

1, i.e., (2.6) with $j=1$. Together with (13) and (15), this reduces to

$$\frac{W_1+R_1f}{\sigma^{-1}\alpha\Gamma_1|\mathbb{S}_1|E/N_{E,1}}\frac{\delta}{\delta-1/\theta}\frac{1-(\bar{\varphi}/\varphi_1^*)^{-(\delta-\frac{1}{\theta})}}{1-(\bar{\varphi}/\varphi_1^*)^{-\delta}}\varphi_1^*=1.$$

Further, substituting the labour and land market clearing conditions, (16) and (18), into this equation results in

$$\frac{\delta}{\delta-1/\theta}\frac{\sigma}{\Gamma_1}\frac{1-(\bar{\varphi}/\varphi_1^*)^{-\left(\delta-\frac{1}{\theta}\right)}}{1-(\bar{\varphi}/\varphi_1^*)^{-\delta}}\left[\frac{\tilde{\beta}_1N_{E,1}}{(1-\lambda_2)N_W}+\eta_1\frac{fN_{E,1}}{1-\alpha fN_{E,1}}\right]=1.$$

Finally, using (9), the first condition is written as follows

$$\frac{\delta}{\delta-1/\theta}\frac{\sigma}{\Gamma_1}\frac{1-(\varphi_1^*/\bar{\varphi})^{\delta-\frac{1}{\theta}}}{1-(\varphi_1^*/\bar{\varphi})^{\delta}}\left\{\frac{\tilde{\beta}_1}{1-\lambda_2}\frac{1}{1-\left(\frac{\underline{\varphi}}{\varphi_1^*}\right)^{\delta}}+\eta_1\frac{f}{1-\alpha f\left[1-\left(\frac{\varphi_1^*}{\bar{\varphi}}\right)^{\delta}\right]\left(\frac{\underline{\varphi}}{\varphi_1^*}\right)^{\delta}}\right\}$$
$$\times\left[1-\left(\frac{\varphi_1^*}{\bar{\varphi}}\right)^{\delta}\right]\left(\frac{\underline{\varphi}}{\varphi_1^*}\right)^{\delta}=1. \tag{19}$$

If I define x and y by $x\equiv\varphi_1^*/\bar{\varphi}\in(0,1)$ and $y\equiv\underline{\varphi}/\varphi_1^*\in(0,1)$, respectively, this equation adds a restriction to the relationship between x and y for a given S_1. Note that λ_2 is a function of $(\varphi_1^*,\bar{\varphi},S_1)$ and that $(\varphi_1^*,\bar{\varphi})$ corresponds to (x,y) equivalently for any given lower bound $\underline{\varphi}$ of productivity.

The second condition is $u(\bar{\varphi})=1$, where $\bar{\varphi}$ is assumed to be greater than $\max_j\{\varphi_j^*\}$, or

$$\frac{\sigma^{-1}\tilde{A}_2\bar{\varphi}^{\frac{1}{\theta}}-R_2f}{\sigma^{-1}\tilde{A}_1\bar{\varphi}^{\frac{1}{\theta}}-R_1f}=\left(\frac{R_2}{R_1}\right)^{1-\alpha}.$$

After some calculations which use (9), (10), (15), and (18), this equation is restated as follows:

$$\frac{\frac{\delta-1/\theta}{\delta}\frac{\Gamma_2}{\eta_2\sigma}\frac{(xy)^{-\delta}-\alpha f}{f}-1}{\frac{\delta-1/\theta}{\delta}\frac{1}{x^{1/\theta}-x^{\delta}}\frac{\Gamma_1}{\eta_1\sigma}\frac{1-\alpha f(1-x^{\delta})y^{\delta}}{fy^{\delta}}-1}=\left[\frac{\frac{1-\alpha f(xy)^{\delta}}{1-\alpha f(1-x^{\delta})y^{\delta}}}{\frac{\eta_2}{\eta_1}\frac{1-S_1}{S_1}}\right]^{\alpha}, \tag{20}$$

which adds another restriction to the relationship between x and y for a given S_1.

Therefore, for a given S_1, there are two unknowns, x and y, and two equations, (19) and (20). This system of equations, if solved, implies that x and y are obtained as functions of S_1. Of course, there might exist multiple solutions for the system, and thus it is more appropriate to state that the system gives x and y as correspondence of S_1. However, in the numerical computation considered in this chapter, the system actually gives a unique solution.

.1.4 Determination of S_1 through Comparative Advantage

In the above discussion, S_1 is fixed at some point. Stated differently, I considered an interior sorting equilibrium where the spatial distribution of final goods industries is fixed in a particular manner. Thus, finally, I discuss how to pin down the value of S_1.

The condition which determines the value of S_1 is the comparative advantage condition, i.e., (2.8), which states that prices of final goods sector S_1, if posted by two regions, are equalized. Focusing on the case considered in the numerical analysis, i.e., $\beta_H = \beta_L$, this condition is written as follows:

$$\frac{\chi_2(S_1)}{\chi_1(S_1)} = \left(\frac{R_2}{R_1}\right)^{1-\alpha\beta} \left[\frac{N_{E,2}^{\theta}\tilde{\varphi}_2}{N_{E,1}^{\theta}\tilde{\varphi}_1}\right]^{-\gamma(S_1)} = 1.$$

Since all the ratios in parentheses and brackets are functions of S_1, as discussed above, this is a single equation determining the value of S_1. Thus, the computation of an equilibrium can be interpreted as a fixed-point problem with respect to S_1, which nests a system of nonlinear equations for (x, y). Once the value of S_1 which satisfies the above equation is found, the values of the other variables are computed. Without loss of generality, the economy-wide income E excluding land rents is normalized to unity.

Chapter 3

Team Production and the Allocation of Creativity across Global and Local Sectors

3.1 Introduction

Functional specialization of cities is a modern view of how a system of cities works in economic activity. As reported by Duranton and Puga (2005), specialization of cities have been transforming from sectoral to functional one. In this form of specialization, larger cities specialize in skill-intensive, non-routine production processes such as management and research and development, while smaller cities in less skill-intensive, more routine ones including line production as an example which are easily replicated by applying ideas and methods developed by the former cities.

Important questions are then how the allocation of creativity is determined; and how policy makers can shift creativity to cities which have functioned as places where routine, repetitive production processes are conducted. Given the increasing importance of non-routine tasks in economic activity (Michaels et al., 2013), creativity and knowledge are clues to understanding urban economic activity and disparities.[1] However, without a comparative advantage in skill-intensive, non-routine production processes, those cities are increasingly faced with difficulties in regional economic development.

Given the above concern about local cities, I focus on the role of team production in attracting creativity, especially connecting creativity to local advantages such as scenery, culture, history, weather and etc. Here, team production is a set of production processes in which the high- and the low-skilled are consolidated in teams, and the former work as managers developing ideas and finding solutions to complex problems in production and the latter as workers mainly conducting production under managers' supervision, respectively. Although this type of team production is observed in the

[1] The fact reported by Moretti (2012) that employment multipliers of creative occupations are relatively large also supports this line of thinking.

whole range of economic activity,[2] relating it to local advantages provides important implications for policies aiming at attracting creativity given that larger cities have a comparative advantage in skill-intensive activities in industries mainly characterized by the second nature, i.e., those less dependent of local advantages.

Therefore, in order to obtain policy implications, I build a stylized two-sector Ricardian comparative advantage model with team production á la Garicano and Rossi-Hansberg (2006b), in which there are two types of agents: those with higher labor productivity (or knowledge in this chapter) and those with lower productivity. Here, the global (local) sector is defined as a sector in which the high-skilled (the low-skilled) have a comparative advantage in production. In both sectors, team production is allowed, providing a device of connecting one high-skilled agent as a manager to some low-skilled agents as workers and allowing for managers to leverage their knowledge more effectively than self-employment.[3] Although there is no cities or space in the model for simplicity, I assume that there is another type of team production in the local sector in addition to that common across sectors under an interpretation that global and local sectors correspond to larger and smaller cities, respectively. Specifically, in the local-sector-specific team production, the high-skilled can learn about local advantages through communications with low-skilled workers within teams who are more knowledgeable about local advantages than managers. The cost and benefit of such learning, the latter is specified as an increase in productivity of managers, are then compared by the high-skilled, who are less willing to learn about local advantages and thus simply tend to apply knowledge applicable to the global sector to the local economic activity when the cost is relatively high.

The implications consist of the following two: The first is that team production changes the nature of comparative advantage. Stated differently, for the relative price level at which the high-skilled do not work in the local sector in the self-employment economy, they optimally choose to work in the local sector via team production. With sufficiently low costs of communication and learning or a sufficiently large productivity gain, the high-skilled still choose to work in the local sector even at the price level at which the low-skilled are willing to work in the global sector in the self-employment economy.

The second is that policy targets should be selected carefully when supporting team production in the local sector in the two senses: First, in general, likelihoods of success in attracting creativity to the local sector are different across policy targets. Reducing communication cost within teams is most effective in the sense that the likelihood is far greater than the other options, reducing cost of learning about local advantages and increasing productivity gain from such learning. This is simply because the other two options are specific to local team production associated with learning. Second, there is a non-monotonic relationship between improvements in policy targets and the allocation of

[2] Actually, functional specialization of cities can be interpreted from the view point of team production. The high-skilled in larger cities as managers and the low-skilled in smaller cities as workers are consolidated in a line of production system (through either trade within firms or outsourcing).

[3] This implies that the low-skilled who are less knowledgeable than the high-skilled cannot become managers since there is no benefit from forming a team the manager of which is an l agent, a person with the lowest knowledge.

creativity. More specifically, the fraction of the high-skilled who choose to work in the local sector has an inverted-U-shaped relationship with improvements in the policy targets. Thus, continuing improvements in the targets first succeed to attract creativity, and then their effects become limited. This is because as the relative price of local good decreases due to the increasing relative supply of local good due to improvements in the targets, the low-skilled exit from the local sector, resulting in an equilibrium where local good is produced by team production only. This then implies that improvements in the targets, which are specific to team production, propagate across all economic activities in the local sector, resulting an increase in the relative supply increasing faster than the relative demand.

This chapter is related to three lines of research. The first is models of team production. There is a strand of literature of team production stressing economic activity in the modern knowledge economy including a seminal paper by Garicano and Rossi-Hansberg (2006a). The model in this chapter is an extension of their simplest form described in Garicano and Rossi-Hansberg (2006b) to multiple sectors and additional form of team production. Unlike this line of research, the focus is on the allocation of creativity or high-skilled agents not in income inequalities.

The second is the literature of urban specialization, especially that with the focus on functional specialization such as Duranton and Puga (2005) and Michaels et al. (2013) among others. In this chapter, functional specialization of cities is interpreted as a form of team production, and the model is designed for its relation to local advantages, which provides a new perspective in this literature.

The third is the strand of the literature of knowledge creation including Berliant and Fujita (2012) among others. Unlike their research stressing collaborations between creative people, the current focus is on those between high-skilled and low-skilled agents.

The remaining structure of this chapter is organized as follows: In Section 3.2, I describe the structure of the model including the forms of team production considered in this chapter. In Section 3.3, I then derive possible scenarios of the allocation of creativity when policy-related parameters consisting of communication cost in the local sector, cost of within-team learning about local advantages, and productivity gain from such learning are improved. Likelihoods of those scenarios are quantified by resorting simulations in Section 3.4. In the final section, Section 3.5, I conclude this chapter.

3.2 The Model

3.2.1 Environment

I consider a closed two-sector economy consisting of two types of economic agents. The sectors consist of a local ℓ sector and a global g sector, and homogeneous goods, ℓ good and g good, are produced in both sectors. The difference between the two sectors is discussed later. The economic agents consist of low-skilled l and high-skilled h, each is endowed with one unit of time. The differences between the two are productivities. Specifically, letting ℓ_i and g_i denote i-agent labor productivity per unit of time in sector ℓ and g, respectively, I assume that h agents have absolute advantages in both sectors and a comparative advantage in g sector: $\ell_l < \ell_h$; $g_l < g_h$; and $g_l/\ell_l < g_h/\ell_h$. The agents are assumed to be competitive, and thus i agents can earn her income of g_i ($p\ell_i$) if she worked in sector g (ℓ), where p denotes the relative price of ℓ good. For simplicity, I assume that the number of each type of agents is fixed, and specifically the relative supply of l agents is given by a parameter $\rho > 1$.

In the following, I introduce team production, production activity by one manager and workers, into the above Ricardian model. More specifically, there are two types of team production. One of the two is allowed in both ℓ and g sectors, while the other is allowed only in ℓ sector, which is the difference between the two sectors.

3.2.2 Team Production

In order to introduce team production, I resort to the framework developed by Garicano and Rossi-Hansberg (2006b). In this framework, producing goods is equivalent to solving problems arising in production activity, and agents' productivities, ℓ_i and g_i, are interpreted as their knowledge allowing them to solve a fraction of problems they are faced with, i.e., $\ell_i, g_i \in (0,1)$. More specifically, each agent draws one problem per unit of time, and each problem is characterized by some level of knowledge required to solve it. If the agent is sufficiently knowledgeable, then she can solve the problem. Therefore, given one unit time endowment and productivities, i agents can solve g_i (ℓ_i) problems in g (ℓ) sector, by the law of large numbers, resulting in (expected) self-employment income of g_i ($p\ell_i$) if they worked in g (ℓ) sector.[4]

Common Form

Instead of self-employment, agents are allowed to form a team consisting of one manager and n workers. For expositional purpose, assume for a moment that the economy consists of a single sector k, where agents' productivities are given by k_l and k_h, and $0 < k_l < k_h < 1$. Then, suppose that an h agent and n_k l workers are to form a team in the sector. Team production then allows agents to conduct the following production processes: First, l workers draw and try to solve problems by themselves. In total, $k_l n_k$ problems are solved by n_k l workers, implying $(1-k_l)n_k$ problems are left unsolved. Second, workers pass $(1-k_l)n_k$ unsolved problems to their manager with communication cost c_k per unit of problems. Finally, the manager with knowledge of k_h suggests how to fix those problems to her workers if she knows, i.e., if k_h is higher than the required level, and the suggested workers solve the related problems. Therefore, in total, the team as a whole can solve $k_h n_k$ problems.

Note that in the above example, team size n_k is determined by the time constraint of the manager: time for communications is limited to unity, i.e., $c_k(1-k_l)n_k = 1$. This implies the total output of the team is given by $k_h n_k = k_h/(c_k(1-k_l))$. Note also that an agent cannot become a manger of

[4] There is no complementarity between sectors, and thus agents will specialization in one of two sectors.

a team with workers who are more knowledgeable than him or those who have the same level of knowledge. In the former case, the total output $k_l n_k$ must be distributed among one l manger and n_k h workers. However, this output is smaller than the sum of outputs associated with self-employment of all members, i.e., $k_l + n_k k_h$, implying someone will be unwilling to participate in this team. A similar argument is also applied to the latter case.

The remaining question is then how this revenue is distributed among members in the team. In order to think about this problem, I assume that the relative supply ρ of l agents is sufficiently large, resulting in excess supply of l agents compared with the number of workers demanded by h managers. Specifically, even if all h agents become managers, the supply of l workers is less than the demand which is equal to n_k times the number of h agents: $n_k = 1/(c_k(1-k_l)) < \rho$.[5]

Given the above assumption, the wage rate w_l of l workers is determined by their outside option value, i.e., income k_l of self-employment since l workers cannot become managers. The zero-profit condition then determines the income w_h of h managers: $k_h n_k = w_h + k_l n_k$, or $w_h = (k_h - k_l)n_k = (k_h - k_l)/(c_k(1-k_l))$. In order to focus on a meaningful case, I assume that this income level is greater than that associated with self-employment, i.e., $(k_h - k_l)/(c_k(1-k_l)) > k_h$.

In sum, communication cost c_k must be some intermediate level. In the rest of this chapter, I assume that this condition is satisfied for both g and ℓ sectors:

$$\frac{1}{\rho(1-g_l)} < c_g < \frac{g_h - g_l}{g_h(1-g_l)}; \quad \frac{1}{\rho(1-\ell_l)} < c_\ell < \frac{\ell_h - \ell_l}{\ell_h(1-\ell_l)}. \tag{3.1}$$

ℓ-specific Form: Learning about Local Advantages

In addition to the common form of team production, I also introduce another type of team production in ℓ sector. The difference from the common one is that within this type of team, an h agent can learn about local advantages, e.g., scenery, culture, history, weather and etc., from l workers. This learning is then reflected in a productivity gain $a \in (1, \bar{a})$ of h manager's productivity, where $\bar{a} \equiv \ell_h^{-1}$, i.e., h manager's productivity increases from ℓ_h to $a\ell_h$. An interpretation is as follows: h manager can apply their knowledge well-suitable to activities in g sector to those in ℓ sector and can still earn income more than l agents do. However, with learning about local advantages, the quality of output increases further. Rather than simply designing a conventional building in a beautiful scenery, designing a building in accord with such nature makes the place more valuable. Such activities involve some arts and thus require more knowledge and creativity. It is assumed that it is prohibitively costly for l agents to conduct this kind of learning.

Learning described above of course costs to some extent. Here, I assume that an h manager incurs iceberg-type cost $\tau > 1$ when learning. Specifically, with learning, passing $(1-\ell_l)n_\ell$ unsolved

[5] Note that this inequality provides some lower bound for communication cost c_k. Although communication cost less than this lower bound generates theoretically interesting equilibrium outcomes as shown by Garicano and Rossi-Hansberg (2006b), I do not consider such situation since it seems the most natural case that the high-skilled or creative people have bargaining power in the wage determination.

problems in a team requires the manager $\tau c_\ell(1-\ell_l)n_\ell$ units of time (not $c_\ell(1-\ell_l)n_\ell$ units).

Given the above cost and benefit of learning, team size and output are given by $1/[\tau c_\ell(1-\ell_l)]$ and $a\ell_h/[\tau c_\ell(1-\ell_l)]$, respectively.

3.2.3 l-agent Choice

Since l agents cannot become a manager, each of them works for h managers either in g or ℓ sector or for herself, i.e., self-employment. In addition, noting that l agents have no bargaining power, wage rates are equalized across the above options. Therefore, l agents simply choose which sector they work for.

This environment is exactly the same as in a simple Ricardian model. That is, letting $p_l^* \equiv g_l/\ell_l$, l agents choose g sector if $p < p_l^*$ and ℓ sector otherwise. This optimal choice implies that $w_l = g_l$ if $p < p_l^*$, and $w_l = p\ell_l$ otherwise.

3.2.4 h-agent Choice

In addition to self-employment, which generates income of g_h or $p\ell_h$ depending on sectors, h agents can become the manager of a team in either g or ℓ sector. In addition, if an h agent chose to form a team in ℓ sector, then she must also choose which type of team she forms. For notational convenience, let g_s and ℓ_s denote self-employment in g sector and that in ℓ sector, respectively. Also let g, $\ell_{w/}$, and $\ell_{w/o}$ denote teams in g sector, those with learning in ℓ sector, and those without learning in ℓ sector, respectively. Therefore, h agents choose one of $\{g_s, \ell_s, g, \ell_{w/}, \ell_{w/o}\}$. Note that for a chosen form $f \in \{g, \ell_{w/}, \ell_{w/o}\}$ of a team, productivity z_f and team size n_f are determined, implying that the wage rate $w_{h,f}$of h manager is given by $w_{h,f} = (z_f - w_l)n_f$.

Since the wage rate $w_{h,f}$ of managers depends on the wage rate w_l of l workers which in turn depends on the relative price p, h-agent choice should be discussed conditional on the level of the relative price p.

Indifference Curves: $p < p_l^*$

If $p < p_l^*$, then the wage rate w_l of l workers is given by $w_l = g_l$. Due to comparative advantage, h agents also choose g_s if they chose self-employment. However, note that g_s is never chosen by h agents when $p < p_l^*$. This is simply because $w_{h,g} > g_s$, which holds under (3.1).

Therefore, h agents are effectively faced with three options: g, $\ell_{w/}$, or $\ell_{w/o}$. It is convenient to provide equations associated with indifference curves (or income equalization):

$$I_{\ell_{w/}\sim\ell_{w/o}} \quad : \quad \frac{pa\ell_h - g_l}{\tau c_\ell(1-\ell_l)} = \frac{p\ell_h - g_l}{c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = \frac{g_l}{\ell_h}\frac{\tau-1}{\tau-a}(\tau \neq a) \tag{3.2}$$

$$I_{g\sim\ell_{w/}} \quad : \quad \frac{g_h - g_l}{c_g(1-g_l)} = \frac{pa\ell_h - g_l}{\tau c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = \frac{g_l}{a\ell_h}\left[1+\tau\frac{c_\ell(1-\ell_l)}{c_g(1-g_l)}\frac{g_h-g_l}{g_l}\right] \tag{3.3}$$

$$I_{g\sim\ell_{w/o}} \quad : \quad \frac{g_h - g_l}{c_g(1-g_l)} = \frac{p\ell_h - g_l}{c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = \frac{g_l}{\ell_h}\left[1+\frac{c_\ell(1-\ell_l)}{c_g(1-g_l)}\frac{g_h-g_l}{g_l}\right]. \tag{3.4}$$

For notational convenience, let $p_{\ell_{w/}\sim\ell_{w/o}}$ denote the relative price corresponding to the indifference curve associated with $\ell_{w/}\sim\ell_{w/o}$. In a similar manner, I use similar notations for the other cases. When emphasizing that the relative price is a function of some parameter θ, I use the expression like $p_{\ell_{w/}\sim\ell_{w/o}}(\theta)$.

Indifference Curves: $p > p_l^*$

If $p > p_l^*$, then the wage rate w_l of l workers is given by $w_l = p\ell_l$. Due to comparative advantage, h agents also choose g_s if they chose self-employment. In this case, preferring g to g_s is not necessarily the case since the choice is dependent on the relative price p.

Therefore, h agents are effectively faced with four options: g_s, g, $\ell_{w/}$, or $\ell_{w/o}$. It is convenient to provide equations associated with indifference curves (or income equalization):[6]

$$I_{\ell_{w/}\sim\ell_{w/o}} \quad : \quad \frac{pa\ell_h - p\ell_l}{\tau c_\ell(1-\ell_l)} = \frac{p\ell_h - p\ell_l}{c_\ell(1-\ell_l)} \quad \Longrightarrow \quad \tau = \frac{a\ell_h-\ell_l}{\ell_h-\ell_l}, \tag{3.5}$$

$$I_{g\sim\ell_{w/}} \quad : \quad \frac{g_h - p\ell_l}{c_g(1-g_l)} = \frac{pa\ell_h - p\ell_l}{\tau c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = \frac{g_h}{\ell_l + (a\ell_h-\ell_l)\frac{c_g(1-g_l)}{\tau c_\ell(1-\ell_l)}}, \tag{3.6}$$

$$I_{g\sim\ell_{w/o}} \quad : \quad \frac{g_h - p\ell_l}{c_g(1-g_l)} = \frac{p\ell_h - p\ell_l}{c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = \frac{g_h}{\ell_l + (\ell_h-\ell_l)\frac{c_g(1-g_l)}{c_\ell(1-\ell_l)}}, \tag{3.7}$$

$$I_{g_s\sim\ell_{w/}} \quad : \quad g_h = \frac{pa\ell_h - p\ell_l}{\tau c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = g_h\frac{\tau c_\ell(1-\ell_l)}{a\ell_h-\ell_l}, \tag{3.8}$$

$$I_{g_s\sim\ell_{w/o}} \quad : \quad g_h = \frac{p\ell_h - p\ell_l}{c_\ell(1-\ell_l)} \quad \Longrightarrow \quad p = g_h\frac{c_\ell(1-\ell_l)}{\ell_h-\ell_l}, \tag{3.9}$$

$$I_{g_s\sim g} \quad : \quad g_h = \frac{g_h - p\ell_l}{c_g(1-g_l)} \quad \Longrightarrow \quad p = \frac{g_h}{\ell_l}[1-c_g(1-g_l)] \tag{3.10}$$

Since parameters of interest are learning cost τ, productivity gain a, communication cost c_ℓ, I summarize h-agent choice in (θ, p) coordinate for each $\theta \in \{\tau, a, c_\ell\}$ below, which is used when conducting comparative statics with some initial parameter set. In the following, I assume that the

[6] For example, first, compare $\ell_{w/}$ and $\ell_{w/o}$. Second, compare the winner of the first step with g. Finally, compare the winner of the second step with g_s.

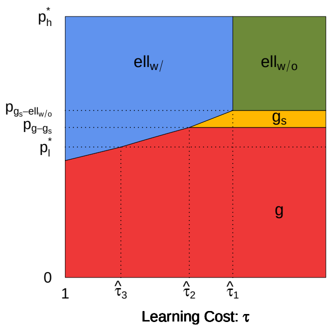

Figure 3.1: h-agent Choice in (τ, p) Coordinate, Pattern 1, $\frac{c_\ell}{c_g} < c_1^*$

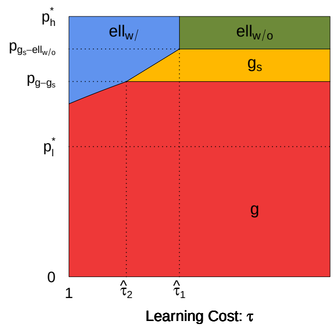

Figure 3.2: h-agent Choice in (τ, p) Coordinate, Pattern 2, $\frac{c_\ell}{c_g} > c_1^*$, $\hat{\tau}_2 > 1$

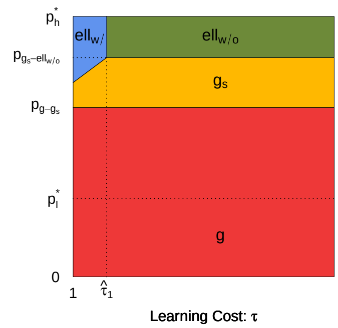

Figure 3.3: h-agent Choice in (τ, p) Coordinate, Pattern 3, $\frac{c_\ell}{c_g} > c_1^*$, $\hat{\tau}_2 < 1$

initial parameters satisfy the following:

$$1 < c_g(1 - g_l) + c_\ell(1 - \ell_l)\frac{\ell_l}{\ell_h - \ell_l}. \tag{3.11}$$

That is, communication costs (c_g, c_ℓ) are not too small. Given this assumption, there exists a range of g_s of some positive measure in (θ, p) coordinate, which becomes clear soon. Stated differently, this condition is expressed as $p_{g\sim g_s} < p_{g_s\sim\ell_{w/o}}$, where $p_{g\sim g_s}$ and $p_{g_s\sim\ell_{w/o}}$ are given by (3.10) and (3.9), respectively.

h-agent Choice in (τ, p) Coordinate

It is found that h-choice in (τ, p) coordinate exhibits one of three patterns depicted in Figure 3.1-3.3. As mentioned in the last paragraph, there exists an area of g_s due to $p_{g\sim g_s} < p_{g_s\sim\ell_{w/o}}$. The thresholds $\hat{\tau}_1$, $\hat{\tau}_2$, and $\hat{\tau}_3$ are given by (3.5), substituting $p_{g\sim g_s}$ into (3.6), and substituting p_l^* into (3.6), respectively:

$$\begin{aligned}
\hat{\tau}_1 &\equiv \frac{a\ell_h - \ell_l}{\ell_h - \ell_l}, \\
\hat{\tau}_2 &\equiv \frac{a\ell_h - \ell_l}{\ell_l}\frac{1 - c_g(1 - g_l)}{c_\ell(1 - \ell_l)}, \\
\hat{\tau}_3 &\equiv \frac{g_l}{\ell_l}\frac{a\ell_h - \ell_l}{g_h - g_l}\frac{1 - g_l}{1 - \ell_l}\frac{c_g}{c_\ell}.
\end{aligned}$$

Given (3.11), $\hat{\tau}_3 < \hat{\tau}_2 < \hat{\tau}_1$.

There are two relationships which are used to distinguish these three patterns. The first is about the level of the ratio c_ℓ/c_g of communication costs. If $c_\ell/c_g < c_1^*$, where the threshold c_1^* is given by

$$c_1^* \equiv \frac{g_l}{\ell_l}\frac{a\ell_h - \ell_l}{g_h - g_l}\frac{1-g_l}{1-\ell_l},$$

then the case corresponds to Pattern 1. Otherwise, Pattern 2 or 3 is relevant. The crucial difference between Pattern 1 and Pattern 2 and 3 is that $p_{g\sim\ell_{w/}}(1) < p_l^*$ in the former case,[7] while the contrary holds in the latter cases.[8] The intuition is that, other things being equal, a lower communication cost c_ℓ in ℓ sector makes h agents more likely to choose $\ell_{w/}$ or $\ell_{w/o}$, and for c_ℓ/c_g less than one, a threshold price $p_{g\sim\ell_{w/}}(1)$ become lower than p_l^*. This distinction becomes important when deriving the relative supply curve. In Pattern 1, $\hat{\tau}_3 > 1$, implying $\hat{\tau}_2 > 1$. In other cases, whether $\hat{\tau}_2 > 1$ or not depdens on cases, which are discussed in the second relationship below.

The second relationship is about the level of $\hat{\tau}_2$ which is the intersection of three areas, g_s, g, and $\ell_{w/}$. Pattern 2 and 3 correspond to cases with $\hat{\tau}_2 > 1$ and $\hat{\tau}_2 < 1$, respectively. This distinction also becomes important when deriving the relative supply curve.

h-agent Choice in (a,p) Coordinate

It is found that without the upper bound $\bar{a}$ for productivity gain a, h-choice in (a,p) coordinate exhibits only one pattern depicted in Figure 3.4.[9] As mentioned, there exists an area of g_s due to $p_{g\sim g_s} < p_{g_s\sim\ell_{w/o}}$. The thresholds $\hat{a}_1$, $\hat{a}_2$, and $\hat{a}_3$ are given by solving (3.5) for a, substituting $p_{g\sim g_s}$ into (3.6), and substituting p_l^* into (3.6), respectively:

$$\begin{aligned}
\hat{a}_1 &\equiv \left(1-\frac{\ell_l}{\ell_h}\right)\tau + \frac{\ell_l}{\ell_h},\\
\hat{a}_2 &\equiv \frac{\ell_l}{\ell_h}\left[1+\tau\frac{c_\ell(1-\ell_l)}{1-c_g(1-g_l)}\right],\\
\hat{a}_3 &\equiv \frac{\ell_l}{\ell_h}\left[1+\tau\frac{c_\ell(1-\ell_l)}{c_g(1-g_l)}\frac{g_h-g_l}{g_l}\right],
\end{aligned}$$

and the ranking $1 < \hat{a}_1 < \hat{a}_2 < \hat{a}_3$ holds.

h-agent Choice in (c_ℓ,p) Coordinate

It is found that without the lower bound for ℓ-sector communication cost c_ℓ, h-choice in (c_ℓ,p) coordinate exhibits one of two patterns depicted in Figure 3.5-3.6.[10] As mentioned, there exists an area of g_s in both patterns due to $p_{g\sim g_s} < p_{g_s\sim\ell_{w/o}}$. The thresholds $\hat{c}_{\ell,1}$, $\hat{c}_{\ell,2}$, and $\hat{c}_{\ell,3}$ in Pattern 1 are given by

[7] Note that $p_{g\sim\ell_{w/}}(1)$ is the value of the relative price corresponding to the indifference curve $I_{g\sim\ell_{w/}}$ evaluated at $\tau = 1$.

[8] I simply omit the knife-edge case, i.e., $c_\ell/c_g = c_1^*$.

[9] The upper bound of a is taken into account when conducting comparative statics. For the summary of h-choice in (a,p) coordinate, it is convenient to omit the upper bound for a for a moment.

[10] As in the previous case, for the summary of h-choice in (c_ℓ,p) coordinate, it is convenient to omit the lower bound for c_ℓ for a moment.

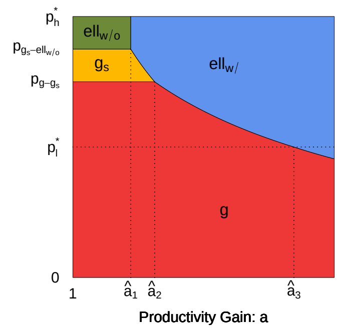

Figure 3.4: h-agent Choice in (a, p) Coordinate

substituting $p_{g\sim g_s}$ into (3.7), substituting p_l^* into (3.7), and substituting $p_{\ell_{w/}\sim\ell_{w/o}}$ given by (3.5) into (3.7), respectively:

$$
\begin{aligned}
\hat{c}_{\ell,1} &\equiv \frac{\ell_h-\ell_l}{\ell_l(1-\ell_l)}[1-c_g(1-g_l)],\\
\hat{c}_{\ell,2} &\equiv \frac{g_l}{\ell_l}\frac{\ell_h-\ell_l}{g_h-g_l}\frac{1-g_l}{1-\ell_l}c_g,\\
\hat{c}_{\ell,3} &\equiv \frac{a-1}{\tau-a}\frac{g_l}{g_h-g_l}\frac{1-g_l}{1-\ell_l}c_g,
\end{aligned}
$$

and the ranking $0<\hat{c}_{\ell,3}<\hat{c}_{\ell,2}<\hat{c}_{\ell,1}$ holds. The thresholds $\tilde{c}_{\ell,1}$ and $\tilde{c}_{\ell,2}$ in Pattern 2 are given by substituting $p_{g\sim g_s}$ into (3.6) and substituting p_l^* into (3.6), respectively:

$$
\begin{aligned}
\tilde{c}_{\ell,1} &\equiv \frac{a\ell_h-\ell_l}{\ell_l}\frac{1-c_g(1-g_l)}{\tau(1-\ell_l)},\\
\tilde{c}_{\ell,2} &\equiv \frac{g_l}{\ell_l}\frac{a\ell_h-\ell_l}{g_h-g_l}\frac{1-g_l}{1-\ell_l}\frac{c_g}{\tau},
\end{aligned}
$$

and the ranking $0<\tilde{c}_{\ell,2}<\tilde{c}_{\ell,1}$ holds.

Pattern 1 and 2 correspond to the case where $\hat{\tau}_1<\tau$ and the one where $\hat{\tau}_1>\tau$, respectively. Given the indifference curve $I_{\ell_{w/}\sim\ell_{w/o}}$ in (3.5), the former case is associated with a relative higher learning cost, making h agent to prefer $\ell_{w/o}$ to $\ell_{w/}$ if $p>p_l^*$, while the contrary holds in Pattern 2. In the case of $p<p_l^*$, the dominance of $\ell_{w/}$ over $\ell_{w/o}$ again applies to Pattern 2, while some mixed result holds in Pattern 1 because $p_{\ell_{w/}\sim\ell_{w/o}}<p_l^*$, where $p_{\ell_{w/}\sim\ell_{w/o}}$ is given by (3.2).

3.3 Equilibrium

Given the agents' optimal choices sumamrized in the previous section, a general equilibrium is defined as a set of the relative price p, quantities, and shares of $\{g_s, \ell_s, g, \ell_{w/}, \ell_{w/o}\}$ in h agents including

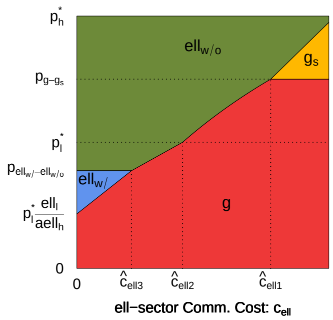

Figure 3.5: h-agent Choice in (c_ℓ, p) Coordinate, Pattern 1, $\hat{\tau}_1 < \tau$

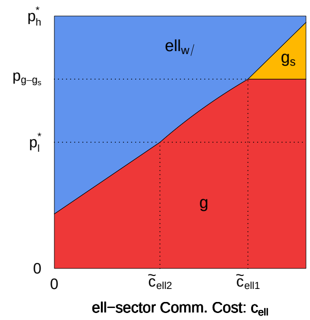

Figure 3.6: h-agent Choice in (c_ℓ, p) Coordinate, Pattern 2, $\hat{\tau}_1 > \tau$

the ℓ-sector share $\lambda_h \in [0,1]$, the sum of shares of $\ell_{w/}$ and $\ell_{w/o}$, in h agents such that both g-good and ℓ-good markets clear; and agents behave optimally.

For this purpose, in the following, I derive the relative supply of ℓ good in Subsection 3.3.1. In Subsection 3.3.2, conducting comparative statics, I identity possible scenarios associated with a decrease in learning cost τ or ℓ-sector communication cost or an increase in productivity gain a, all of which seem to be related to some policy instruments and thus provide some implications for policy targets.

3.3.1 Relative Supply Curve

In the following, I derive the relative supply curve of ℓ good for various values of three parameters: τ, a, and c_ℓ. In doing so, I omit the lower and upper bounds imposed on these parameters for ease of proposition. These restrictions will be considered when conducting comparative statics of the equilibrium ℓ-sector share λ_h in h agents. In all cases, the behavior of l agents has common properties, i.e., they choose g-sector self-employment g_s or workers if $p < p_l^*$; otherwise, they choose ℓ-sector self-employment ℓ_s or workers.

Decreasing Learning Cost τ

In order to derive the relative supply curve, it is sufficient to consider four phases: $\hat{\tau}_1 < \tau$; $\hat{\tau}_2 < \tau < \hat{\tau}_1$; $\hat{\tau}_3 < \tau < \hat{\tau}_2$; and $1 < \tau < \hat{\tau}_3$. Therefore, I derive the relative supply curve for each phase in the following.

Phase: $\hat{\tau}_1 < \tau$

As the relative price p increases from zero to p_h^*, a typical h-agent choice changes as follows: they choose g, g_s, and $\ell_{w/o}$ for $p \in (0, p_{g\sim g_s})$, for $p \in (p_{g\sim g_s}, p_{g_s\sim \ell_{w/o}})$, and for $p \in (p_{g_s\sim \ell_{w/o}}, p_h^*]$, respectively. For $p = p_{g\sim g_s}$ and for $p = p_{g_s\sim \ell_{w/o}}$, they are indifferent between g and g_s and between g_s and $\ell_{w/o}$, respectively.

Noting that the thresholds above are independent of τ, it holds that for all $\tau > \hat{\tau}_1$, the relative supply curve is the same as depicted in Figure 3.7. s_g denotes the the relative supply of ℓ good when all h are g managers, and l agents are either ℓ_s or g worker, and s_{g_s} is defined by the relative supply of ℓ good when all agents are self-employed and they specialize in the sectors for which they have comparative advantages:

$$s_g \equiv \frac{\ell_l}{g_h}[\rho c_g(1-g_l)-1], \quad s_{g_s} \equiv \frac{\ell_l}{g_h}\rho.$$

Phase: $\hat{\tau}_2 < \tau < \hat{\tau}_1$

As the relative price p increases from zero to p_h^*, a typical h-agent choice changes as follows: they choose g, g_s, and $\ell_{w/}$ for $p \in (0, p_{g\sim g_s})$, for $p \in (p_{g\sim g_s}, p_{g_s\sim \ell_{w/}}(\tau))$, and for $p \in (p_{g_s\sim \ell_{w/}}(\tau), p_h^*]$, respectively. For $p = p_{g\sim g_s}$ and for $p = p_{g_s\sim \ell_{w/}}(\tau)$, they are indifferent between g and g_s and between g_s and $\ell_{w/}$, respectively.

Therefore, the relative supply curve has three steps as in the previous case. However, the crucial difference is that now the third step of the relative supply curve shifts downward as τ decreases from $\hat{\tau}_1$ to $\hat{\tau}_2$ (Figure 3.8). More specifically, starting from the same one as in the previous phase, the relative supply curve converges to a two-step functions, the height of the second step of which is equal to $p_{g\sim g_s}$. This is clear if we note that the height of the third step is given by $p_{g_s\sim \ell_{w/}}(\tau)$; and that $p_{g_s\sim \ell_{w/}}(\tau)$ converges to $p_{g\sim g_s}$ as suggested by Figure 3.1 or 3.2.

Phase: $\hat{\tau}_3 < \tau < \hat{\tau}_2$

As the relative price p increases from zero to p_h^*, a typical h-agent choice changes as follows: they choose g and $\ell_{w/}$ for $p \in (0, p_{g\sim \ell_{w/}}(\tau))$ and for $p \in (p_{g\sim \ell_{w/}}(\tau), p_h^*]$, respectively. For $p = p_{g\sim \ell_{w/}}(\tau)$, they are indifferent between g and $\ell_{w/}$.

Therefore, the relative supply curve has two steps as depicted in Figure 3.9. In addition, since $p_{g\sim \ell_{w/}}(\tau)$ decreases from $p_{g\sim g_s}$ to p_l^* as suggested by Figure 3.1, the second step of the relative supply curve shifts downward, starting from the same height as in the previous phase at $\tau = \hat{\tau}_2$ and converging to the height of p_l^*. In the latter limit, the relative supply curve becomes a one-step function.

Phase: $1 < \tau < \hat{\tau}_3$

Although h-agent choice has the same pattern as in the previous phase, note an important difference, that is, in the current case, the threshold price $p_{g\sim \ell_{w/}}(\tau) < p_l^*$. This implies a relative supply curve, Figure 3.10, with two steps where the quantity corresponding to the second jump is now given by $s_{\ell_{w/}}$ (not s_g), the relative supply of ℓ good when all h agents are $\ell_{w/}$ managers (not g managers), and l

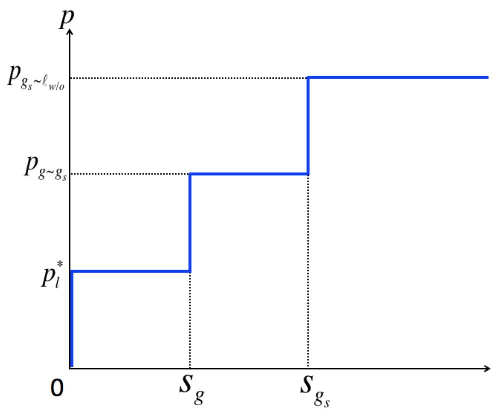

Figure 3.7: $\hat{\tau}_1 < \tau$

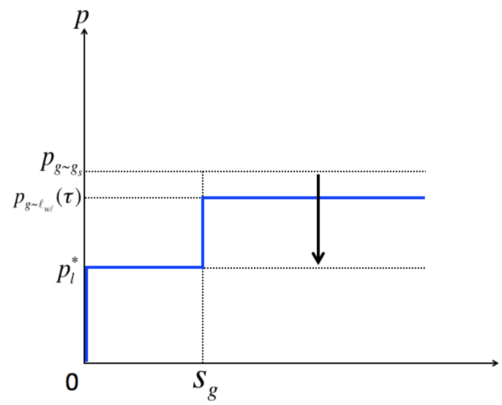

Figure 3.9: $\hat{\tau}_3 < \tau < \hat{\tau}_2$

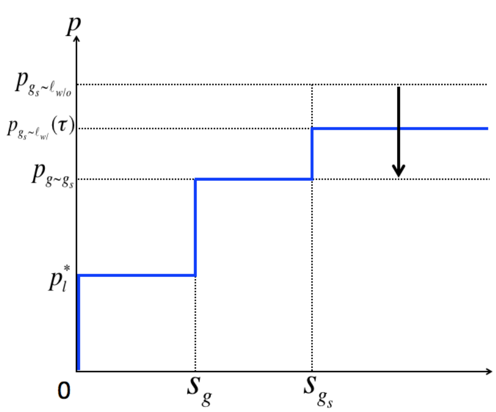

Figure 3.8: $\hat{\tau}_2 < \tau < \hat{\tau}_1$

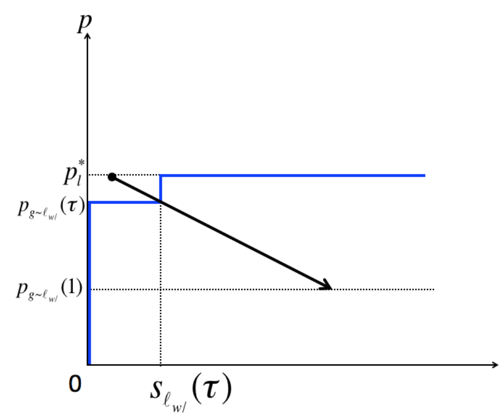

Figure 3.10: $1 < \tau < \hat{\tau}_3$

agents are either g-sector self-employment or $\ell_{w/}$ workers such that g-manager demand for workers is satisfied:

$$s_{\ell_{w/}} \equiv \frac{a\ell_h}{g_l} \frac{1}{\rho\tau c_\ell(1-\ell_l)-1}.$$

In Figure 3.10, the notation $s_{\ell_{w/}}(\tau)$ stresses that the quantity $s_{\ell_{w/}}$ depends on τ.

In addition, since $s_{\ell_{w/}}(\tau)$ as well as $p_{g\sim\ell_{w/}}(\tau)$ depends on τ, where the former is decreasing in τ, the corner between the first and second steps shifts southeast as τ decreases.

Increasing Productivity Gain a

There are four phases in productivity gain a for the derivation of the relative supply curve: $1 < a < \hat{a}_1$; $\hat{a}_1 < a < \hat{a}_2$; $\hat{a}_2 < a < \hat{a}_3$; and $\hat{a}_3 < a$. Given that a decrease in learning cost τ and an increase in

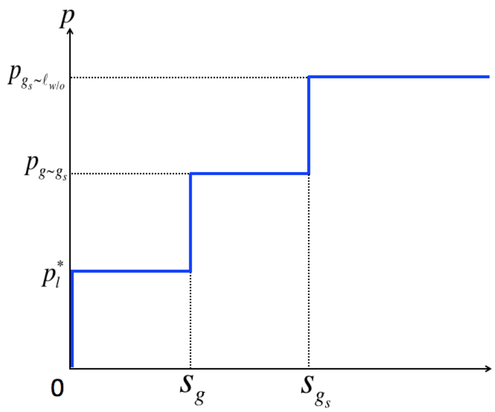

Figure 3.11: $1 < a < \hat{a}_1$

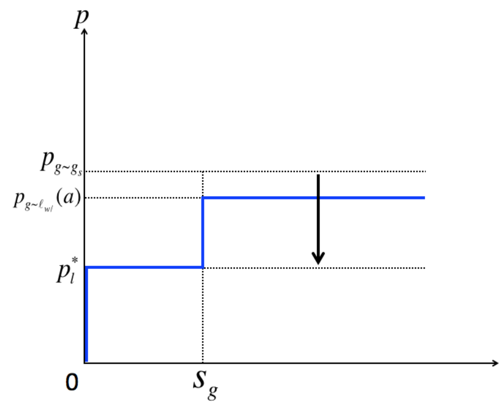

Figure 3.13: $\hat{a}_2 < a < \hat{a}_3$

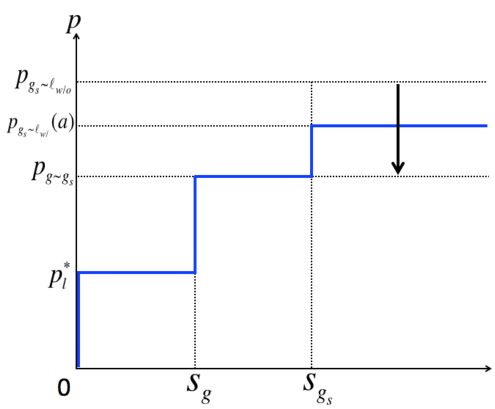

Figure 3.12: $\hat{a}_1 < a < \hat{a}_2$

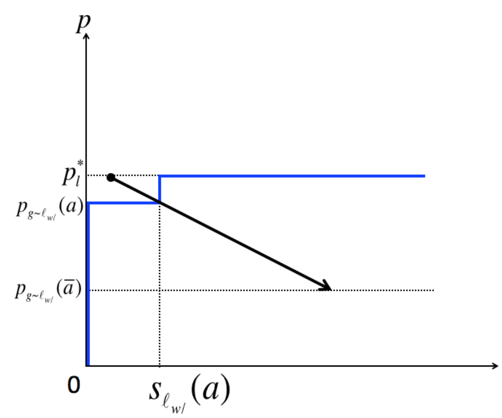

Figure 3.14: $\hat{a}_3 < a$

productivity gain a are close to a mirror-image relationship, it is guessed that the (shifts in) supply curves corresponding to these phases have the same patterns. Actually, this is the case, and thus I only report the resulting relative supply curves in Figure 3.11-3.14.

Decreasing Communication Cost c_ℓ: Pattern 1

Unlike the other two cases, the derivation of the relative supply curve for decreasing ℓ-sector communication cost c_ℓ requires case-by-case analysis: Pattern 1, where $\hat{\tau}_1 < \tau$, and Pattern 2, where $\tau < \hat{\tau}_1$, which are discussed in order below.

In order to derive the relative supply curve for Pattern 1, it is necessary to consider four phases: $\hat{c}_{\ell,1} < c_\ell$; $\hat{c}_{\ell,2} < c_\ell < \hat{c}_{\ell,1}$; $\hat{c}_{\ell,3} < c_\ell < \hat{c}_{\ell,2}$; and $\underline{c}_\ell < c_\ell < \hat{c}_{\ell,3}$, where $\underline{c}_\ell$ is given by $\underline{c}_\ell \equiv 1/[\rho(1-\ell_l)]$, and it is assumed that $\underline{c}_\ell < \hat{c}_{\ell,3}$. Therefore, I derive the relative supply curve for each phase in the following.

Phase: $\hat{c}_{\ell,1} < c_\ell$

As the relative price p increases from zero to p_h^*, h-agent choice changes as follows: they choose g, g_s, and $\ell_{w/o}$ for $p \in (0, p_{g\sim g_s})$, for $p \in (p_{g\sim g_s}, p_{g_s\sim\ell_{w/o}}(c_\ell))$, and for $p \in (p_{g_s\sim\ell_{w/o}}(c_\ell), p_h^*]$, respectively. For $p = p_{g\sim g_s}$ and for $p = p_{g_s\sim\ell_{w/o}}(c_\ell)$, they are indifferent between g and g_s and between g_s and $\ell_{w/o}$, respectively.

Figure 3.15 shows the resulting relative supply curve. It has three steps, the third of which shifts downward as c_ℓ decreases from a height of $p_{g_s\sim\ell_{w/o}}(c_\ell^0))$ to that of $p_{g\sim g_s}$, where c_ℓ^0 denotes some starting level of c_ℓ.

Phase: $\hat{c}_{\ell,2} < c_\ell < \hat{c}_{\ell,1}$

As the relative price p increases from zero to p_h^*, h-agent choice changes as follows: they choose g and $\ell_{w/o}$ for $p \in (0, p_{g\sim\ell_{w/o}}(c_\ell))$ and for $p \in (p_{g\sim\ell_{w/o}}(c_\ell), p_h^*]$, respectively. For $p = p_{g\sim\ell_{w/o}}(c_\ell)$, they are indifferent between g and $\ell_{w/o}$. Here, $p_{g\sim\ell_{w/o}}(c_\ell)$ is increasing in c_ℓ, and, starting from $p_{g\sim g_s}$ at $c_\ell = \hat{c}_{\ell,1}$, it converges to p_l^* as c_ℓ converges to $\hat{c}_{\ell,2}$.

Then, it is shown that the relative supply curve has two steps as depicted in Figure 3.16, the second step of which shifts downward as c_ℓ decreases.

Phase: $\hat{c}_{\ell,3} < c_\ell < \hat{c}_{\ell,2}$

The h-agent choice is same as in the previous phase except that $p_{g\sim\ell_{w/o}}(c_\ell) < p_l^*$, and $p_{g\sim\ell_{w/o}}(c_\ell)$ changes monotonically from p_l^* to $p_{\ell_{w/}\sim\ell_{w/o}}$ as c_ℓ converges from $\hat{c}_{\ell,2}$ to $\hat{c}_{\ell,3}$.

The associated relative supply curve is depicted in Figure 3.17, where $s_{\ell_{w/o}}$ denotes the relative supply of ℓ good when all h agents are $\ell_{w/o}$ managers, and l agents engage in either g-sector self-employment or $\ell_{w/o}$ team such that g-manager demand for workers is satisfied:

$$s_{\ell_{w/o}} \equiv \frac{\ell_h}{g_l} \frac{1}{\rho\tau c_\ell(1-\ell_l)-1} > s_{\ell_{w/}},$$

where the inequality holds under $a < \tau$ which is implied by the current assumption that $\hat{\tau}_1 < \tau$. The figure shows that the corner between the first and second steps shifts southeast since the corresponding price level $p_{g\sim\ell_{w/o}}(c_\ell)$ and the relative quantity $s_{\ell_{w/o}}(c_\ell)$ are increasing and decreasing functions of c_ℓ.

Phase: $\underline{c}_\ell < c_\ell < \hat{c}_{\ell,3}$

As the relative price p increases from zero to p_h^*, h-agent choice changes as follows: they choose g, $\ell_{w/}$, and $\ell_{w/o}$ for $p \in (0, p_{g\sim\ell_{w/}}(c_\ell))$, for $p \in (p_{g\sim\ell_{w/}}(c_\ell), p_{\ell_{w/}\sim\ell_{w/o}})$, and for $p \in (p_{\ell_{w/}\sim\ell_{w/o}}, p_h^*]$, respectively. For $p = p_{g\sim\ell_{w/}}(c_\ell)$ and for $p = p_{\ell_{w/}\sim\ell_{w/o}}$, they are indifferent between g and $\ell_{w/}$ and between $\ell_{w/}$ and $\ell_{w/o}$, respectively. Here, $p_{g\sim\ell_{w/}}(c_\ell)$ is increasing in c_ℓ, and, starting from $p_{\ell_{w/}\sim\ell_{w/o}}$ at $c_\ell = \hat{c}_{\ell,3}$, it converges to $p_l^*\ell_l(a\ell_h)$ as c_ℓ converges to zero.

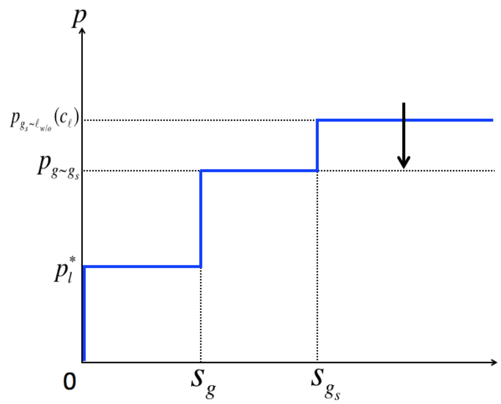

Figure 3.15: $\hat{c}_{\ell,1} < c_\ell$

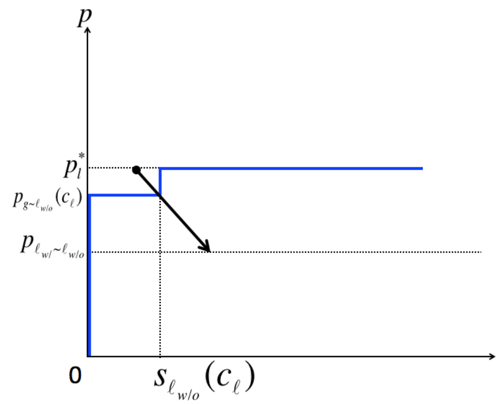

Figure 3.17: $\hat{c}_{\ell,3} < c_\ell < \hat{c}_{\ell,2}$

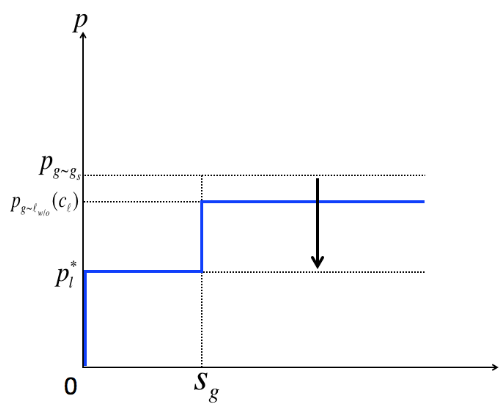

Figure 3.16: $\hat{c}_{\ell,2} < c_\ell < \hat{c}_{\ell,1}$

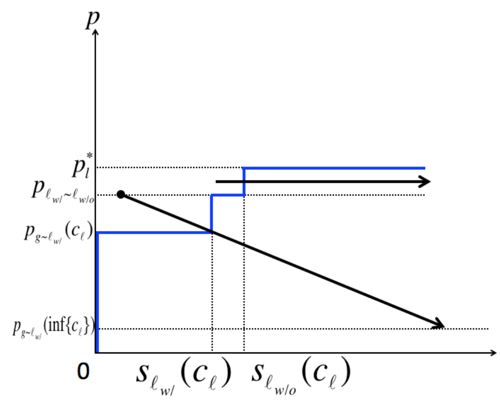

Figure 3.18: $(\inf\{c_\ell\} =)\underline{c}_\ell < c_\ell < \hat{c}_{\ell,3}$

The associated relative supply curve has therefore three steps (Figure 3.18). However, unlike those discussed before, it is most complicated. Specifically, the corner between the first and second steps and that between the second and third ones shift southeast and eastward, respectively, since $p_{\ell_{w/} \sim \ell_{w/o}}$, $p_{g \sim \ell_{w/}}(c_\ell)$, $s_{\ell_{w/o}}(c_\ell)$, and $s_{\ell_{w/}}(c_\ell)$ are constant, increasing, decreasing, and decreasing in c_ℓ, respectively. In addition, $s_{\ell_{w/o}}(c_\ell) \to \infty$ as $c_\ell \downarrow \underline{c}_\ell$.

Decreasing Communication Cost c_ℓ: Pattern 2

In the derivation of the relative supply curve for Pattern 2, there are three phases to consider: $\tilde{c}_{\ell,1} < c_\ell$; $\tilde{c}_{\ell,2} < c_\ell < \tilde{c}_{\ell,1}$; and $\underline{c}_\ell < c_\ell < \tilde{c}_{\ell,2}$, where the observed properties are similar to the last three phases in the case of decreasing τ and increasing a as suggested by comparison between Figure 3.1, 3.4, and 3.6. Thus, I omit step-by-step derivations of the relative supply curves, which are depicted in Figure 3.19-3.21.

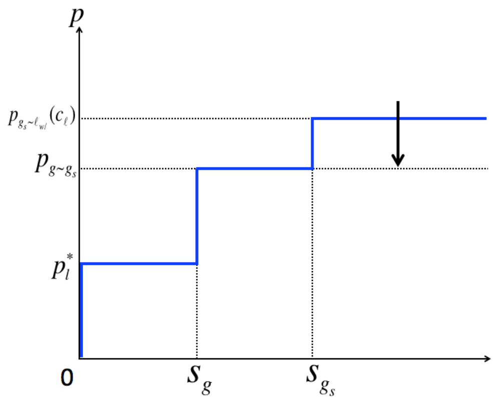

Figure 3.19: $\tilde{c}_{\ell,1} < c_\ell$

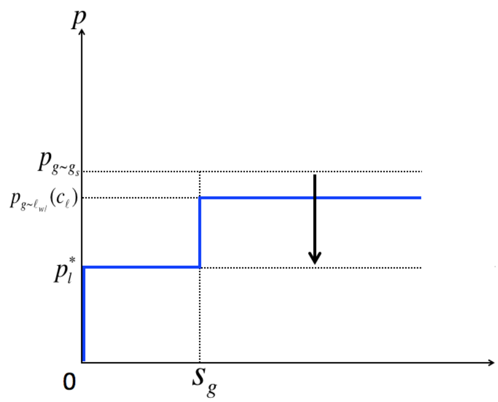

Figure 3.20: $\tilde{c}_{\ell,2} < c_\ell < \tilde{c}_{\ell,1}$

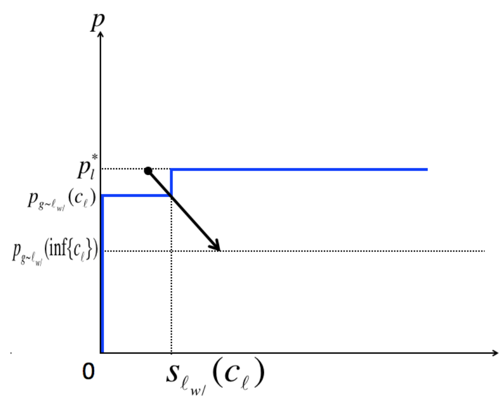

Figure 3.21: $(\inf\{c_\ell\} =)\underline{c}_\ell < c_\ell < \tilde{c}_{\ell,2}$

3.3.2 Possible Scenarios

In this chapter, I focus on the simplest case where agents' preference is specified by a Cobb-Douglas function with expenditure shares, α_g and α_ℓ, of g and ℓ goods, i.e., $\alpha_g + \alpha_\ell = 1$, implying that the relative demand for ℓ good is given by αp^{-1}, where $\alpha \equiv \alpha_\ell/\alpha_g > 0$. Depending on the ratio α of expenditure shares, various equilibria are possible, and thus there are many "dynamics" of ℓ-sector share λ_h in h agents when τ or c_ℓ decreases or when a increases.

Therefore focusing on equilibria with a certain property to avoid exhaustive analyses, I summarize possible scenarios under the Cobb-Douglas demand structure. In the following subsection, I describe the equilibria of interest, and then the results of comparative statics of λ_h with respect to τ, a, and c_ℓ are summarized in order.

Equilibrium of Interest

The equilibria of interest are those where the equilibrium relative price p is included in a range $(p_l^*, p_{g\sim g_s})$, and all h agents engage in team production in g sector. Stated differently, I impose a set of restrictions to initial parameters: First, α must satisfy

$$\frac{g_l}{g_h}[\rho c_g(1-g_l)-1] < \alpha < [\rho c_g(1-g_l)-1][1-c_g(1-g_l)]. \tag{3.12}$$

Second, it must hold that $\max\{\hat{\tau}_2, 1\} < \tau$, which is equivalent to $1 < a < \hat{a}_2$ given that $\tau > 1$; or $\hat{c}_{\ell,1} < c_\ell$ if $\hat{\tau}_1 < \tau$ and $\tilde{c}_{\ell,1} < c_\ell$ otherwise.

The reason why I focus on this situation is that it to some extent resembles functional specialization of cities reported by the literature (Duranton and Puga, 2005). Relatively skilled agents tend to participate in creative activity such as research and development, output of which are then, in a world of functional specialization, distributed across cites as blue prints used for repetitive, routine production activity such as line production. The preferred interpretation here is that this system of production corresponds to team production in g sector.

In the following, instead of imposing the above restrictions rigorously, I first identify possible scenarios under the assumption that the relative demand and supply curves intersect at an equilibrium relative price p less than $p_{g\sim g_s}$ for changes in τ, a, and c_ℓ.[11]

τ-λ_h Dynamics

In order to identify possible scenarios of the dynamics of the ℓ-sector share λ_h in h agents when learning cost τ decreases (hereafter, τ-λ_h dynamics), it is necessary to consider each of four phases discussed in Subsection 3.3.1. In the following, I list the possible τ-λ_h dynamics for each phase (under the assumption that $1 < \hat{\tau}_3 < \hat{\tau}_2 < \hat{\tau}_1$) and then summarize the possible scenarios in a whole range $(1, \tau^0]$ of τ by considering possibility that $\hat{\tau}_2 < 1$; or $\hat{\tau}_3 < 1$, where τ^0 is an initial value of τ satisfying $\max\{\hat{\tau}_2, 1\} < \tau^0$.

Phase: $\hat{\tau}_1 < \tau$

Since the relative supply curve does not shift in this phase, the equilibrium does not change, implying that $\lambda_h = 0$ for all $\tau > \hat{\tau}_1$ (Figure 3.22).

Phase: $\hat{\tau}_2 < \tau < \hat{\tau}_1$

Since the relative supply curve shifts only above $p_{g\sim g_s}$, the equilibrium is not affected, implying that $\lambda_h = 0$ for all $\tau \in (\hat{\tau}_2, \hat{\tau}_1)$ (Figure 3.22).

[11] For a more rigorous analysis, I resort to simulations in Section 3.4.

Phase: $\hat{\tau}_3 < \tau < \hat{\tau}_2$

In this phase, there exists some threshold of τ, above which the equilibrium does not change and below which it is affected, resulting in a continuous increase in λ_h (Figure 3.23).

When $\lambda_h > 0$, the relative demand curve intersects the horizontal line of the second step, i.e., $p = p_{g\sim\ell_{w/}(\tau)}$, where h agents are indifferent between g and $\ell_{w/}$, and λ_h and $1-\lambda_h$ fractions of h agents engage in $\ell_{w/}$ and g teams, respectively. Therefore, the demand for l workers are the sum of that $\lambda_h N_h/[\tau c_\ell(1-\ell_l)]$ from $\ell_{w/}$-team managers and that $(1-\lambda_h)N_h/[c_g(1-g_l)]$ from g-team managers. l agents not working for h managers (either in g or $\ell_{w/}$) engage in self-production ℓ_s in ℓ sector since $p > p_l^*$.

The associated market clearing condition is thus expressed as follows:

$$s = \frac{\frac{a\ell_h\lambda_h}{\tau c_\ell(1-\ell_l)} + \ell_l\left[\rho - \frac{\lambda_h}{\tau c_\ell(1-\ell_l)} - \frac{1-\lambda_h}{c_g(1-g_l)}\right]}{\frac{g_h(1-\lambda_h)}{c_g(1-g_l)}} = d(p_{g\sim\ell_{w/}}(\tau)) \quad \Longrightarrow \quad \lambda_h = \frac{d(p_{g\sim\ell_{w/}}(\tau)) - s_g}{d(p_{g\sim\ell_{w/}}(\tau)) + \frac{1}{\tau}\frac{a\ell_h-\ell_l}{g_h}\frac{c_g(1-g_l)}{c_\ell(1-\ell_l)} + \frac{\ell_l}{g_h}},$$

where s and $d(p_{g\sim\ell_{w/}}(\tau))$ denote the relative supply and demand of ℓ good, respectively. Since $\partial\lambda_h/\partial\tau \propto -(1+\alpha) < 0$, a decrease in τ results in an increase in λ_h. In addition, as $\tau \nearrow \hat{\tau}_2$, $\lambda_h \searrow 0$ since $p \nearrow p_{g\sim g_s}$, and thus $d(p_{g\sim\ell_{w/}}(\tau)) \searrow s_g$.

Phase: $1 < \tau < \hat{\tau}_3$

Focus on a case where $\tau \uparrow \hat{\tau}_3$. Then, depending on α, the relative demand at $p = p_l^*$ might exceed the relative supply $s_{\ell_{w/}}$. Specifically, if $s_{\ell_{w/}}(\hat{\tau}_3) < d(p_l^*)$, then the equilibrium relative price $p = p_l^*$, making l agents are indifferent between g_s and ℓ_s. The excess demand $d(p_l^*) - s_{\ell_{w/}}(\hat{\tau}_3)$ is then satisfied by l-agent self-production. As described, $\lambda_h = 1$ in this case.

On the other hand, if $d(p_l^*) < s_{\ell_{w/}}(\hat{\tau}_3)$, then $p = p_{g\sim\ell_{w/}}(\hat{\tau}_3)(= p_l^*)$, and h agents are indifferent between g and $\ell_{w/}$. Given the excess supply, some fraction of h agents are allocated to g, implying $\lambda_h \in (0,1)$.

As for τ decreasing to one, there are two and one τ-λ_h dynamics for $s_{\ell_{w/}}(\hat{\tau}_3) < d(p_l^*)$ and for $s_{\ell_{w/}}(\hat{\tau}_3) > d(p_l^*)$, respectively, given that

$$\frac{\partial}{\partial\tau}\left[\frac{s_{\ell_{w/}}(\tau)}{d(p_{g\sim\ell_{w/}}(\tau))}\right] \propto -[(g_h - g_l) + \rho c_g(1-g_l)g_l] < 0.$$

If $s_{\ell_{w/}}(\hat{\tau}_3) < d(p_l^*)$, then there are two possible τ-λ_h dynamics as depicted in Figure 3.24. One (solid line) is a case where the relative supply $s_{\ell_{w/}}(1)$ at $\tau = 1$ is smaller than the relative demand $d(p_{g\sim\ell_{w/}}(1))$ at $\tau = 1$, and the other (dashed line) is a case where the contrary holds. In the former case, $\lambda_h = 1$ for all $\tau \in (1, \hat{\tau}_3)$, while λ_h starts to decrease at some threshold of τ in the last case. The latter fact follows from the result that the partial derivative of λ_h with respect to τ is positive, where

λ_h is given by the relevant market clearing condition and specified as

$$\lambda_h = \frac{(g_h - g_l)\frac{c_\ell(1-\ell_l)}{c_g(1-g_l)} + g_l \rho c_\ell (1-\ell_l)}{(g_h - g_l)\frac{c_\ell(1-\ell_l)}{c_g(1-g_l)} + \frac{1}{\tau}(g_l + \frac{a\ell_h}{d(p_{g\sim\ell_{w/}}(\tau))})}.$$

If $d(p_l^*) < s_{\ell_{w/}}(\hat{\tau}_3)$, then λ_h starts from some positive share in $(0,1)$ as mentioned above and continuously decreases to some lower level given that $\partial\lambda_h/\partial\tau > 0$.

Summary: Possible Scenarios

Given the above results, I summarize the possible scenarios in the following proposition in which I now consider possibility that $\hat{\tau}_2 < 1$; or $\hat{\tau}_3 < 1$, the latter of which is equivalent to $c_\ell/c_g > c_1^*$.

Proposition 3.1 (Possible Scenarios in τ-λ_h Dynamics)**.** *Suppose that productivities of agents satisfy the conditions on absolute and relative advantages; and that initial parameters satisfy (3.1) and (3.11). In addition, focus on a case where the benchmark equilibrium is characterized as a relative price p less than $p_{g\sim g_s}$ and h agents all engaging in team production in g sector. Then, possible scenarios are as follows:*

1. *If $c_\ell/c_g < c_1^*$, then there are three scenarios: Scenario 2, 3, and 4.*

   (a) *Suppose that the relative demand for ℓ good at a price level of p_l^* is less than the relative output of ℓ good with learning cost of $\hat{\tau}_3$ when all h agents are $\ell_{w/}$ managers, and l agents are either g-sector self-employed or $\ell_{w/}$ workers, i.e., $d(p_l^*) < s_{\ell_{w/}}(\hat{\tau}_3)$ or equivalently*

   $$\alpha < \frac{\ell_h}{\ell_l}\frac{a}{\rho\hat{\tau}_3 c_\ell(1-\ell_l)-1}.$$

   Then, as learning cost τ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase, then jumps to some interior level, and finally starts to decrease (Scenario 2).

   (b) *Suppose that $s_{\ell_{w/}}(\hat{\tau}_3) < d(p_l^*)$ and that the relative demand for ℓ good at a price level of $p_{g\sim\ell_{w/}}(1)$ is less than the relative output of ℓ good with learning cost of unity when all h agents are $\ell_{w/}$ managers, and l agents are either g-sector self-employed or $\ell_{w/}$ workers, i.e., $d(p_{g\sim\ell_{w/}}(1)) < s_{\ell_{w/}}(1)$ or equivalently*

   $$\alpha < \frac{1}{\rho c_\ell(1-\ell_l)-1}\left[1+\frac{c_\ell(1-\ell_l)}{c_g(1-g_l)}\frac{g_h-g_l}{g_l}\right].$$

   Then, as learning cost τ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase, then jumps to unity, and finally starts to decrease (Scenario 3).

(c) *Suppose that* $s_{\ell_{w/}}(\hat{\tau}_3) < d(p_l^*)$ *and that* $s_{\ell_{w/}}(1) < d(p_{g\sim\ell_{w/}}(1))$. *Then, as learning cost* τ *decreases, the* ℓ*-sector share* λ_h *in h agents eventually starts to increase and jumps to unity (Scenario 4).*

2. *If* $c_\ell/c_g > c_1^*$ *and if* $\hat{\tau}_2 > 1$, *then there are two scenarios: Scenario 0 and 1.*

(a) *Suppose that the relative demand for* ℓ *good at a price level of* $p_{g\sim\ell_{w/}}(1)$ *is less than the relative output of* ℓ *good when all h (l) agents choose self-production in g (*ℓ*) sector, i.e.,* $d(p_{g\sim\ell_{w/}}(1)) < s_g$, *or equivalently,*

$$\alpha < \frac{\rho c_g(1-g_l)-1}{1+\frac{a\ell_h-\ell_l}{\ell_l}\frac{c_g(1-g_l)}{c_\ell(1-\ell_l)}}.$$

Then, the ℓ*-sector share* λ_h *in h agents is constant at* $\lambda_h = 0$ *for all possible learning costs* τ *(Scenario 0).*

(b) *Suppose in contrast that* $d(p_{g\sim\ell_{w/}}(1)) > s_g$. *Then, as learning cost* τ *decreases, the* ℓ*-sector share* λ_h *in h agents eventually starts and continues to increase (Scenario 1).*

3. *If* $c_\ell/c_g > c_1^*$ *and if* $\hat{\tau}_2 < 1$, *then there is one scenario: The* ℓ*-sector share* λ_h *in h agents is constant at* $\lambda_h = 0$ *for all possible learning costs* τ *(Scenario 0).*

Scenario 0-4 are illustrated in Figure 3.26-3.30.

a-λ_h and c_ℓ-λ_h Dynamics

Similar arguments are also applied to the other two dynamics, a-λ_h and c_ℓ-λ_h dynamics. Here, I report the results only in the following propositions. As for a-λ_h dynamics, the upper bound $\bar{a} \equiv \ell_h^{-1}$ for productivity gain a is considered, while the lower bound $\underline{c}_\ell$ for ℓ-sector communication cost c_ℓ, in c_ℓ-λ_h dynamics, where $\underline{c}_\ell \equiv 1/[\rho(1-\ell_l)]$.

Proposition 3.2 (Possible Scenarios in a-λ_h Dynamics)**.** *Suppose that productivities of agents satisfy the conditions on absolute and relative advantages; and that initial parameters satisfy (3.1) and (3.11). In addition, focus on a case where the benchmark equilibrium is characterized as a relative price p less than* $p_{g\sim g_s}$ *and h agents all engaging in team production in g sector. Then, possible scenarios are as follows:*

1. *If* $\bar{a} < \hat{a}_2$, *then there are one scenario: The* ℓ*-sector share* λ_h *in h agents is constant at* $\lambda_h = 0$ *for all possible productivity gains a (Scenario 0).*

2. *If* $\hat{a}_2 < \bar{a} < \hat{a}_3$, *then there are two scenarios: Scenario 0 and 1.*

(a) *Suppose that the relative demand for ℓ good at a price level of $p_{g\sim\ell_{w/}}(\bar{a})$ is less than the relative output of ℓ good when all h (l) agents choose self-production in g (ℓ) sector, i.e., $d(p_{g\sim\ell_{w/}}(\bar{a})) < s_g$, or equivalently,*

$$\alpha < \frac{\rho c_g(1-g_l)-1}{1+\frac{1}{\ell_l}\frac{c_g(1-g_l)}{\tau c_\ell}}.$$

Then, the ℓ-sector share λ_h in h agents is constant at $\lambda_h = 0$ for all possible productivity gains a (Scenario 0).

(b) *Suppose in contrast that $d(p_{g\sim\ell_{w/}}(\bar{a})) > s_g$. Then, as productivity gain a increases, the ℓ-sector share λ_h in h agents eventually starts and continues to increase (Scenario 1).*

3. *If $\hat{a}_3 < \bar{a}$, then there are two scenarios: Scenario A1 and A2.*

(a) *Suppose that the relative demand for ℓ good at a price level of p_l^* is less than the relative output of ℓ good with productivity gain of $\hat{a}_3$ when all h agents are $\ell_{w/}$ managers, and l agents are either g-sector self-employed or $\ell_{w/}$ workers, i.e., $d(p_l^*) < s_{\ell_{w/}}(\hat{a}_3)$, or equivalently,*

$$\alpha < \frac{\ell_h}{\ell_l}\frac{\hat{a}_3}{\rho\tau c_\ell(1-\ell_l)-1}.$$

Then, as productivity gain a increases, the ℓ-sector share λ_h in h agents eventually starts to increase, then jumps to and continues to be some interior level (Scenario A1).

(b) *Suppose in contrast that $d(p_l^*) > s_{\ell_{w/}}(\hat{a}_3)$. Then, as productivity gain a increases, the ℓ-sector share λ_h in h agents eventually starts to increase, then jumps to and continues to be unity (Scenario A2).*

Proposition 3.3 (Possible Scenarios in c_ℓ-λ_h Dynamics)**.** *Suppose that productivities of agents satisfy the conditions on absolute and relative advantages; and that initial parameters satisfies (3.1) and (3.11). In addition, focus on a case where the benchmark equilibrium is characterized as a relative price p less than $p_{g\sim g_s}$ and h agents all engaging in team production in g sector. Then, possible scenarios are as follows:*

1. *If $\hat{\tau}_1 < \tau$ and if $\hat{c}_{\ell,1} < \underline{c}_\ell$, then there is one scenario: The ℓ-sector share λ_h in h agents is constant at $\lambda_h = 0$ for all possible communication cost c_ℓ (Scenario 0).*

2. *If $\hat{\tau}_1 < \tau$ and if $\hat{c}_{\ell,2} < \underline{c}_\ell < \hat{c}_{\ell,1}$, then there are two scenarios: Scenario 0 and 1.*

(a) *Suppose that $d(p_{g\sim\ell_{w/o}}(\underline{c}_\ell)) < s_g$, then the ℓ-sector share λ_h in h agents is constant at $\lambda_h = 0$ for all possible communication cost c_ℓ (Scenario 0).*

(b) *Suppose in contrast that* $d(p_{g\sim\ell_{w/o}}(\underline{c}_\ell)) > s_g$*. Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts and continues to increase (Scenario 1).*

3. *If* $\hat{\tau}_1 < \tau$ *and if* $\hat{c}_{\ell,3} < \underline{c}_\ell < \hat{c}_{\ell,2}$*, then there are two scenarios: Scenario 2 and 3.*

   (a) *Suppose that* $d(p_l^*) < s_{\ell_{w/o}}(\hat{c}_{\ell,2})$*. Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase, then jumps to some interior level, and finally starts to decrease (Scenario 2).*

   (b) *Suppose in contrast that* $d(p_l^*) < s_{\ell_{w/o}}(\hat{c}_{\ell,2})$*. Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase, then jumps to unity, and finally starts to decrease (Scenario 3).*

4. $\hat{\tau}_1 < \tau$ *and if* $\underline{c}_\ell < \hat{c}_{\ell,3}$*, then there are eight scenarios: Scenario 3-4 and C1-C6.*

   (a) *Suppose that* $d(p_l^*) < s_{\ell_{w/o}}(\hat{c}_{\ell,2})$ *and that* $d(p_{\ell_{w/}\sim\ell_{w/o}}) < s_{\ell_{w/}}(\hat{c}_{\ell,3})$*. Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase and then repeats jumping to some level and decreasing twice (Scenario C1).*

   (b) *Suppose that* $d(p_l^*) < s_{\ell_{w/o}}(\hat{c}_{\ell,2})$*, that* $s_{\ell_{w/}}(\hat{c}_{\ell,3}) < d(p_{\ell_{w/}\sim\ell_{w/o}})$*, and that* $d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) < s_{\ell_{w/}}(\underline{c}_\ell)$*. Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase, jumps to some interior level and then decreases, jumps to unity, and finally starts to decrease (Scenario C2).*

   (c) *If* $d(p_l^*) < s_{\ell_{w/o}}(\hat{c}_{\ell,2})$*, if* $s_{\ell_{w/}}(\hat{c}_{\ell,3}) < d(p_{\ell_{w/}\sim\ell_{w/o}})$*, and if* $s_{\ell_{w/}}(\underline{c}_\ell) < d(p_{g\sim\ell_{w/}}(\underline{c}_\ell))$*. Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase, then jumps to some interior level and decreases continuously, and finally becomes constant (Scenario C3).*

   (d) *If* $s_{\ell_{w/o}}(\hat{c}_{\ell,2}) < d(p_l^*)$*, if* $s_{\ell_{w/o}}(\hat{c}_{\ell,3}) < d(p_{\ell_{w/}\sim\ell_{w/o}})$*, and if* $d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) < s_{\ell_{w/}}(\underline{c}_\ell)$*, Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase, then jumps to unity, and finally starts to decrease (Scenario 3).*

   (e) *If* $s_{\ell_{w/o}}(\hat{c}_{\ell,2}) < d(p_l^*)$*, if* $s_{\ell_{w/o}}(\hat{c}_{\ell,3}) < d(p_{\ell_{w/}\sim\ell_{w/o}})$*, and if* $s_{\ell_{w/}}(\underline{c}_\ell) < d(p_{g\sim\ell_{w/}}(\underline{c}_\ell))$*, Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase and jumps to unity (Scenario 4).*

   (f) *If* $s_{\ell_{w/o}}(\hat{c}_{\ell,2}) < d(p_l^*)$*, if* $d(p_{\ell_{w/}\sim\ell_{w/o}}) < s_{\ell_{w/o}}(\hat{c}_{\ell,3})$*, and if* $d(p_{\ell_{w/}\sim\ell_{w/o}}) < s_{\ell_{w/}}(\hat{c}_{\ell,3})$*, Then, as communication cost* c_ℓ *decreases, the* ℓ*-sector share* λ_h *in* h *agents eventually starts to increase, then jumps to unity, and finally starts to decrease with a downward jump at some level of communication cost (Scenario C4).*

   (g) *If* $s_{\ell_{w/o}}(\hat{c}_{\ell,2}) < d(p_l^*)$*, if* $d(p_{\ell_{w/}\sim\ell_{w/o}}) < s_{\ell_{w/o}}(\hat{c}_{\ell,3})$*, if* $s_{\ell_{w/}}(\hat{c}_{\ell,3}) < d(p_{\ell_{w/}\sim\ell_{w/o}})$*, and if* $d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) < s_{\ell_{w/}}(\underline{c}_\ell)$*,* *Then, as communication cost c_ℓ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase and then repeats jumping to unity, being constant for a while and then decreasing twice (Scenario C5).*

(h) *If $s_{\ell_{w/o}}(\hat{c}_{\ell,2}) < d(p_l^*)$, if $d(p_{\ell_{w/}\sim\ell_{w/o}}) < s_{\ell_{w/o}}(\hat{c}_{\ell,3})$, if $s_{\ell_{w/}}(\hat{c}_{\ell,3}) < d(p_{\ell_{w/}\sim\ell_{w/o}})$, and if $s_{\ell_{w/}}(\underline{c}_\ell) < d(p_{g\sim\ell_{w/}}(\underline{c}_\ell))$,*
Then, as communication cost c_ℓ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase, jumps to and stay at unity, starts to decrease, and finally jumps to unity (Scenario C6).

5. *$\tau < \hat{\tau}_1$ and if $\tilde{c}_{\ell,1} < \underline{c}_\ell$, then there is one scenario: Scenario 0.*

6. *$\tau < \hat{\tau}_1$ and if $\tilde{c}_{\ell,2} < \underline{c}_\ell < \tilde{c}_{\ell,1}$, then there are two scenarios: Scenario 0 and 1.*

   (a) *Suppose that $d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) < s_g$. Then, the ℓ-sector share λ_h in h agents is constant at $\lambda_h = 0$ for all possible communication cost c_ℓ (Scenario 0).*

   (b) *Suppose in contrast that $d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) > s_g$. Then, as communication cost c_ℓ decreases, the ℓ-sector share λ_h in h agents eventually starts and continues to increase (Scenario 1).*

7. *$\tau < \hat{\tau}_1$ and if $\underline{c}_\ell < \tilde{c}_{\ell,2}$, then there are three scenarios: Scenario 2, 3 and 4.*

   (a) *Suppose that $d(p_l^*) < s_{\ell_{w/}}(\tilde{c}_{\ell,2})$. Then, as communication cost c_ℓ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase, then jumps to some interior level, and finally starts to decrease (Scenario 2).*

   (b) *Suppose that $s_{\ell_{w/}}(\tilde{c}_{\ell,2}) < d(p_l^*)$ and that $d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) < s_{\ell_{w/}}(\underline{c}_\ell)$. Then, as communication cost c_ℓ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase, then jumps to unity, and finally starts to decrease (Scenario 3).*

   (c) *Suppose that $s_{\ell_{w/}}(\tilde{c}_{\ell,2}) < d(p_l^*)$ and that $s_{\ell_{w/}}(\underline{c}_\ell) < d(p_{g\sim\ell_{w/}}(\underline{c}_\ell))$. Then, as communication cost c_ℓ decreases, the ℓ-sector share λ_h in h agents eventually starts to increase and jumps to unity (Scenario 4).*

In Proposition 3.3, the following equivalences hold:

$$
\begin{aligned}
d(p_{g\sim\ell_{w/o}}(\underline{c}_\ell)) \lesseqgtr s_g &\iff \alpha \lesseqgtr \frac{\rho c_g(1-g_l)-1}{1+\frac{\ell_h-\ell_l}{\ell_l}\rho c_g(1-g_l)} \\
d(p_l^*) \lesseqgtr s_{\ell_{w/o}}(\hat{c}_{\ell,2}) &\iff \alpha \lesseqgtr \frac{\ell_h}{\ell_l}\frac{1}{\rho\hat{c}_{\ell,2}(1-\ell_l)-1} \\
d(p_{\ell_{w/}\sim\ell_{w/o}}) \lesseqgtr s_{\ell_{w/o}}(\hat{c}_{\ell,3}) &\iff \alpha \lesseqgtr \frac{\tau-1}{\tau-a}\frac{1}{\rho\hat{c}_{\ell,3}(1-\ell_l)-1} \\
d(p_{\ell_{w/}\sim\ell_{w/o}}) \lesseqgtr s_{\ell_{w/}}(\hat{c}_{\ell,3}) &\iff \alpha \lesseqgtr \frac{\tau-1}{\tau-a}\frac{a}{\rho\tau\hat{c}_{\ell,3}(1-\ell_l)-1} \\
d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) \lesseqgtr s_{\ell_{w/}}(\underline{c}_\ell) &\iff \alpha \lesseqgtr \frac{1}{\tau-1}\left[1+\frac{\tau}{\rho}\frac{1}{c_g(1-g_l)}\frac{g_h-g_l}{g_l}\right] \\
d(p_{g\sim\ell_{w/}}(\underline{c}_\ell)) \lesseqgtr s_g &\iff \alpha \lesseqgtr \frac{\rho c_g(1-g_l)-1}{1+\frac{a\ell_h-\ell_l}{\ell_l}\frac{\rho}{\tau}c_g(1-g_l)} \\
d(p_l^*) \lesseqgtr s_{\ell_{w/}}(\tilde{c}_{\ell,2}) &\iff \alpha \lesseqgtr \frac{\ell_h}{\ell_l}\frac{a}{\rho\tau\tilde{c}_{\ell,2}(1-\ell_l)-1}.
\end{aligned}
$$

Scenario A1-A2 and Scenario C1-C6 are illustrated in Figure 3.31-3.32 and Figure 3.33-3.38, respectively.

3.4 Numerical Experiment

As mentioned in the previous section, I assume that initial parameters satisfy (3.1), (3.11), (3.12), and $\max\{\hat{\tau}_2, 1\} < \tau$ when conducting comparative statics. Then I identify scenarios that actually realize under these restrictions when learning cost τ decreases, when productivity gain a from learning increases, and when ℓ-sector communication cost c_ℓ decreases.

However, since there are many possible scenarios in those dynamics, I resort to Monte Carlo simulation where a large number of random samples satisfying the above set of restrictions are generated, and comparative statics is conducted for each sample. Given the set of restrictions, it is suffice to set upper bounds for only two parameters: the relative supply $\bar{\rho}$ of l agents and learning cost $\bar{\tau}$.

3.4.1 Algorithm

For each fixed set of upper bounds, $\bar{\rho}$ and $\bar{\tau}$, the algorithm below is used to generate random samples and conduct comparative statics.

Step 1: Generate samples of parameters $(\ell_l, \ell_h, g_l, g_h, \rho, c_g, c_\ell, a, \tau)$ of size of one million from the uniform distribution over the subset of parameters satisfying the stated conditions:

(a) Generate ℓ_l at random such that $0 < \ell_l < (1-\bar{\rho}^{-1})$.

(b) Generate ℓ_h at random such that $[\bar{\rho}/(\bar{\rho}-1)]\ell_l < \ell_h < 1$.

(c) Generate g_h at random such that $0 < g_h < 1$.

(d) Generate g_l at random such that $0 < g_l < (\ell_l/\ell_h)g_h$.

(e) Generate ρ at random such that $\ell_h/(\ell_h - \ell_l) < \rho \leq \bar{\rho}$.

(f) Generate a at random such that

$$1 < a < \min\left\{\ell_h^{-1}, \frac{\ell_l}{\ell_h}\left(1 + \bar{\tau}\frac{g_h}{g_l}\frac{\ell_h - \ell_l}{\ell_h}\right)\right\}.$$

(g) Generate c_g at random such that

$$\max\left\{\frac{1}{\rho(1-g_l)}, \frac{\ell_h - \ell_l}{\ell_h(1-g_l)}, \frac{1}{1-g_l} - \frac{\bar{\tau}}{1-g_l}\frac{\ell_l}{\ell_h}\frac{\ell_h - \ell_l}{a\ell_h - \ell_l}\right\} < c_g < \frac{g_h - g_l}{g_h(1-g_l)}.$$

(h) Generate c_ℓ at random such that

$$\max\left\{\frac{1}{\rho(1-\ell_l)}, \frac{\ell_h - \ell_l}{\ell_l(1-\ell_l)}[1 - c_g(1-g_l)], \frac{a\ell_h - \ell_l}{\bar{\tau}\ell_l(1-\ell_l)}[1 - c_g(1-g_l)]\right\} < c_\ell < \frac{\ell_h - \ell_l}{\ell_h(1-\ell_l)}.$$

(i) Generate τ at random such that $\max\{\hat{\tau}_2, 1\} < \tau \leq \bar{\tau}$.

Step 2: For each sample, construct equidistant grid points on the following closed interval of α under which an equilibrium relative price p satisfies $p_l^* \leq p \leq p_{g\sim g_s}$:

$$\frac{g_l}{g_h}[\rho c_g(1-g_l) - 1] \leq \alpha \leq [\rho c_g(1-g_l) - 1][1 - c_g(1-g_l)],$$

and compute the share of each scenario in the grid points using Proposition 3.1-3.3.

Step 3: Compute the sample average of the share of each scenario.

Step 1 in the algorithm ensures that parameters satisfy the required conditions. For $\bar{\tau}$, I consider three values: low (1.1), middle (1.5), and high (2.0).[12] $\bar{\rho}$ ranges from 2 to 20.

3.4.2 Results

Compared with the other two parameters, a decrease in ℓ-sector communication cost c_ℓ is most effective in shifting h agents from global to local sectors in the sense that the measure of Scenario 0, the share of Scenario 0 in the samples, is lowest for most of the pair $(\bar{\rho}, \bar{\tau})$ of the upper bounds for the relative supply of l agents and learning cost. Figure 3.39 shows the measure of Scenario 0 in each dynamics for each $(\bar{\rho}, \bar{\tau})$. It is clear that for most of $(\bar{\rho}, \bar{\tau})$, the measure of Scenario 0 in c_ℓ-λ_h dynamics is lower than those in the other two dynamics. This is mainly because unlike the other two dynamics,

[12] Learning within ℓ team increases time cost of passing unsolved problems by $100 \times (\tau - 1)\%$. $\bar{\tau}$ gives the upper bound for this increase, and "low," "middle," and "high" correspond to the maximal increase of 10%, 50%, and 100%, respectively.

the measure of Scenario 0 decreases faster as $\bar{\rho}$ increases. Although the measure of Scenario 0 in τ-λ_h dynamics and that in a-λ_h dynamics decreases as $\bar{\tau}$ increases and decreases, respectively, the measure of Scenario 0 in c_ℓ-λ_h dynamics is far smaller than the other dynamics.[13]

When focusing on $\ell_{w/}$-specific parameters, i.e., τ and a, there is no clear ranking in the sense that the measure of Scenario 0 tends to be higher in τ-λ_h dynamics when $\bar{\tau}$ is low, while the contrary holds when $\bar{\tau}$ is high as shown in Figure 3.39. This suggests that an effective policy target related to local team production depends on cases. When learning cost τ is likely to be high (the case of high $\bar{\tau}$), team production with learning is relatively costly, making the effect of increasing productivity gain a limited. Figure 3.39 also shows that there exists some threshold of $\bar{\tau}$ (conditional on the value of $\bar{\rho}$) below which increasing productivity gain a is more effective than lowering learning cost τ.

In addition to the measure of Scenario 0, I also report the measure of each scenario other than Scenario 0 for each dynamics in Figure 3.40. The left, the center, and the right panels show measures of scenarios in τ-λ_h, a-λ_h, and c_ℓ-λ_h dynamics, respectively, while the lower, the middle, and the upper panels show measures of scenarios for low, middle, and high $\bar{\tau}$, respectively. Depending on parameter values, any scenario can emerge, i.e., measures of all scenarios become positive in some parameter range.

There are at least two properties common across all dynamics. The first is simply that as the measure of Scenario 0 decreases due to changes in $\bar{\tau}$, then measures of other scenarios shift upward.

The second is that for lower $\bar{\rho}$, intermediate λ_h, i.e., $\lambda_h \in [0,1)$, is likely to happen, while the corner equilibrium, i.e., $\lambda_h = 1$, does for higher $\bar{\rho}$. More specifically, for low $\bar{\rho}$, Scenario 2 (or Scenario A1 in the case of a-λ_h dynamics) and Scenario 1 are likely to happen, while scenarios with $\lambda_h = 1$ happen with zero probability for sufficiently low $\bar{\rho}$. However, for high $\bar{\rho}$, scenarios with $\lambda_h = 1$ become more likely to happen, and among those scenarios, Scenario 4 in τ-λ_h dynamics, Scenario A2 in a-λ_h dynamics, and Scenario 3 in c_ℓ-λ_h dynamics are of measures comparable with those of scenarios with $\lambda_h \in [0,1)$.

The mechanism behind the second property is the result of general equilibrium effects of changing relative supply ρ of l agents since individual choices do not directly depend on ρ. Further, the property is simply due to increasing supply capacities of ℓ good by ℓ teams, i.e., $s_{\ell_{w/}}$ and $s_{\ell_{w/o}}$, relative to the relative demand for ℓ good. Given the definitions, $s_{\ell_{w/}}$ and $s_{\ell_{w/o}}$ are decreasing functions of ρ since the supply for g good including that from self-employment of l agents, those not employed by h agents, increases as the relative supply ρ of l agents increases. Thus, when ρ is relative large, the relative demand for ℓ good is more likely to be higher than the supply, resulting in h agents all engaging in team production in ℓ sector.

An important policy implication observed in Figure 3.40 is that when encouraging team production in ℓ sector, i.e., reducing communication cost c_ℓ or learning cost τ or increasing productivity gain a, effects on λ_h are likely to be non-monotonic, and thus a careful policy management is necessary.

[13] More precisely, as $\bar{\tau}$ increases, the measure of Scenario 0 in c_ℓ-λ_h dynamics increases. However, the magnitude of this change is small compared with those in the other two dynamics.

More specifically, Scenario 2 (or Scenario A1) has the highest measure than those except for Scenario 0.[14] For example, reducing communication cost c_ℓ first succeeds at shifting h agents from g to ℓ sectors and then fails. The mechanism behind this non-monotonicity is as follows: When λ_h is increasing, the equilibrium relative price p of ℓ good is higher than p_l^*, implying that l agents not employed by h managers in g sectors all engage in self-employment in ℓ sector, i.e., ℓ_s. Then, given that l agents engaging in ℓ_s do not benefit from reduced communication cost c_ℓ, a lower communication cost c_ℓ, by increasing the relative demand for ℓ good due to a decrease in the price p, requires some h agents to fix the excess demand for ℓ good. In a phase where λ_h is decreasing (or constant in a-λ_h dynamics), $p < p_l^*$, and thus l agents not employed by h managers all engage in self-employment in g sector (not ℓ sector), making a decrease in communication cost c_ℓ beneficial for all agents engaging in ℓ-good production.

3.5 Conclusion

In this chapter, I introduce team production á la Garicano and Rossi-Hansberg (2006b) into a two-sector Ricardian comparative advantage model in order to derive implications for policies encouraging team production activity in local sectors and aiming at a shift of creativity from global to local sectors. The policy-related parameters in the model consist of cost of communications within team production in local sectors, cost of within-team learning about local advantages, and productivity gain from such learning. The first implication is that team production changes the nature of comparative advantage, or stated differently, team production could be a tool of shifting creativity from global to local sectors. However, policy targets should be selected carefully since the three options above are different in the likelyhood of success in shifting creativity, and the most likely case of success is associated with non-monotone dynamics of the allocation of creativity, the second implication.

Given not only the stylized specification of the model but also non-monotonicity, several extensions are worth pursuing in order to investigate the robustness of the implications and derive further policy implications. Examples are to introduce a more general utility function such as CES, explicit space (and competition between local governments), urban externalities, and continuous heterogeneity of agents.

[14]Note that Scenario A1 has no decreasing phase in a-λ_h because the relative demand and supply have the same elasticity with respect to a. However, this phase can be interpreted as an ineffectiveness of policies aiming at an increase in productivity gain a.

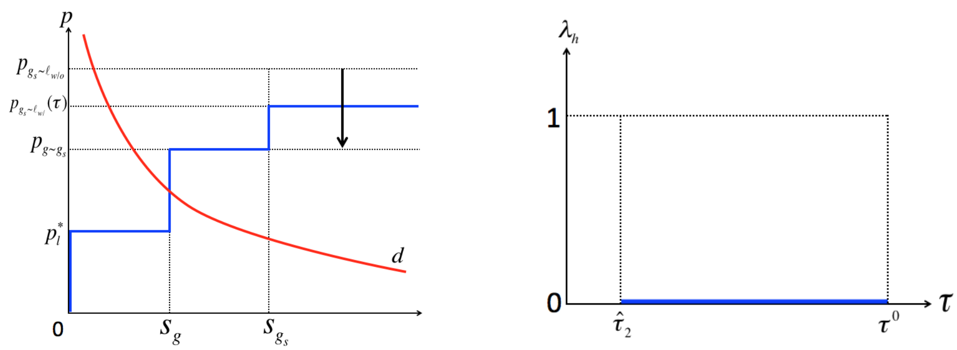

Figure 3.22: Relative Quantities and τ-λ_h Dynamics, $\hat{\tau}_2 < \tau$

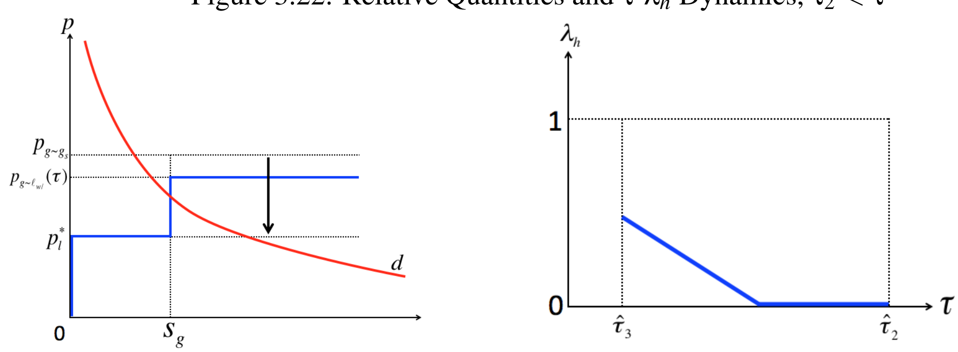

Figure 3.23: Relative Quantities and τ-λ_h Dynamics, $\hat{\tau}_3 < \tau < \hat{\tau}_2$

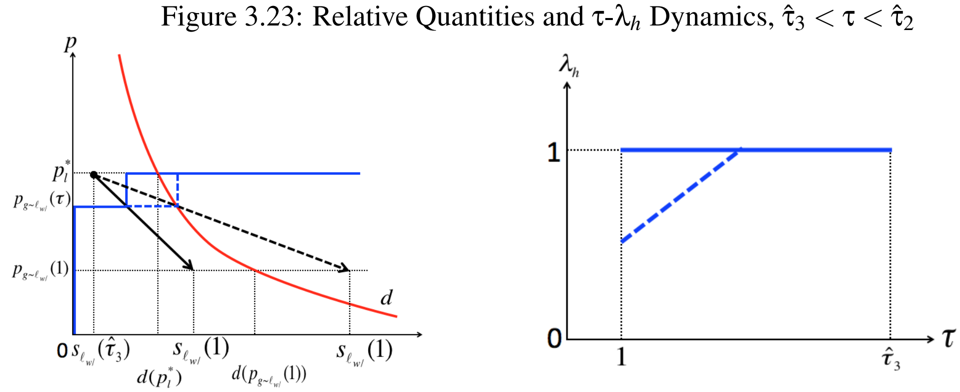

Figure 3.24: Relative Quantities and τ-λ_h Dynamics, $1 < \tau < \hat{\tau}_3$, $s_{\ell_{w/}}(\hat{\tau}_3) < d(p_l^*)$

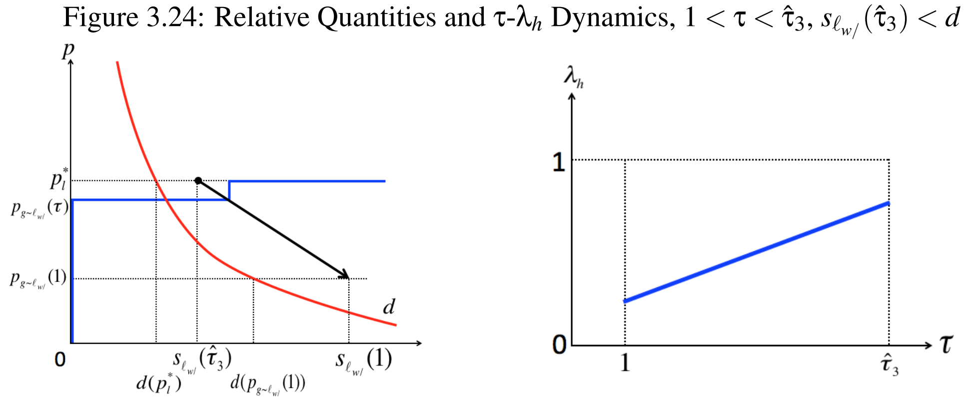

Figure 3.25: Relative Quantities and τ-λ_h Dynamics, $1 < \tau < \hat{\tau}_3$, $d(p_l^*) < s_{\ell_{w/}}(\hat{\tau}_3)$

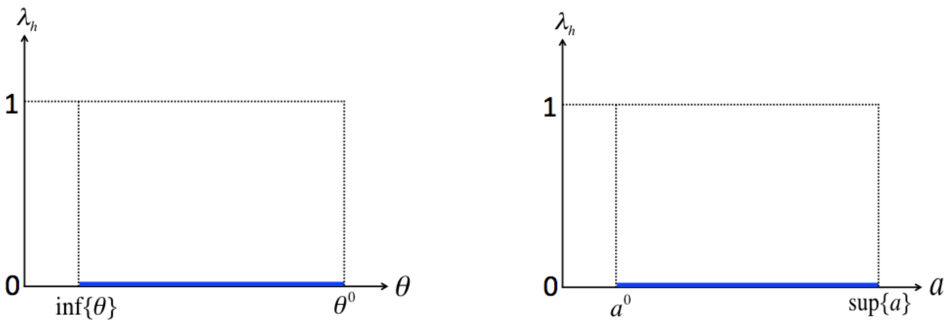

Figure 3.26: Scenario 0

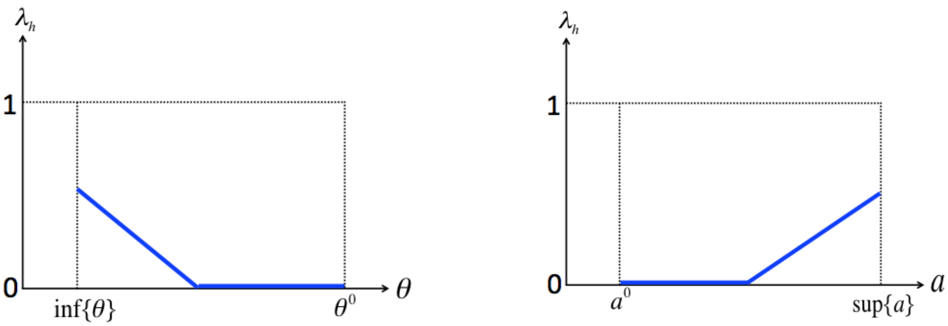

Figure 3.27: Scenario 1

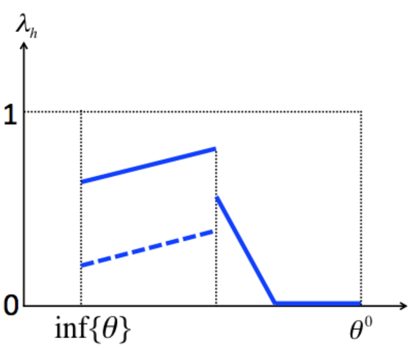

Figure 3.28: Scenario 2

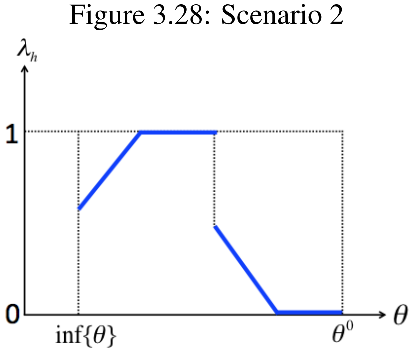

Figure 3.29: Scenario 3

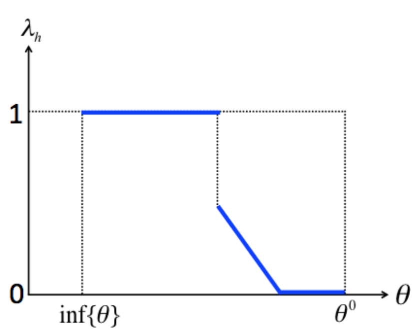

Figure 3.30: Scenario 4

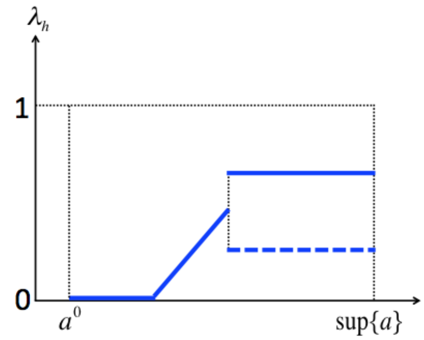

Figure 3.31: Scenario A1

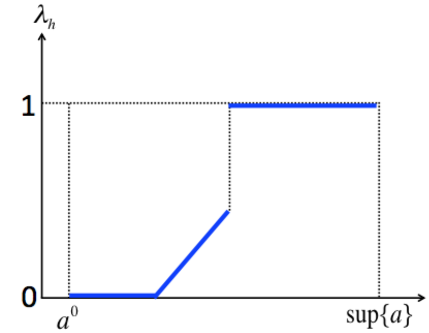

Figure 3.32: Scenario A2

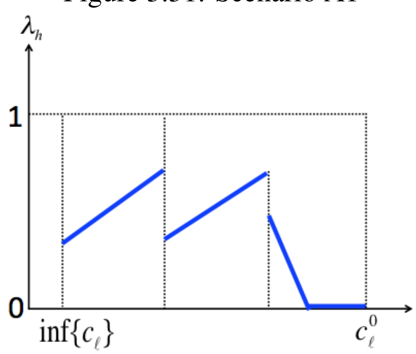

Figure 3.33: Scenario C1

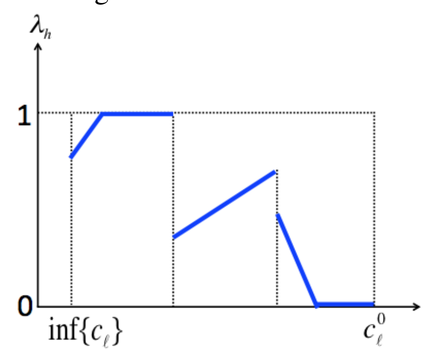

Figure 3.34: Scenario C2

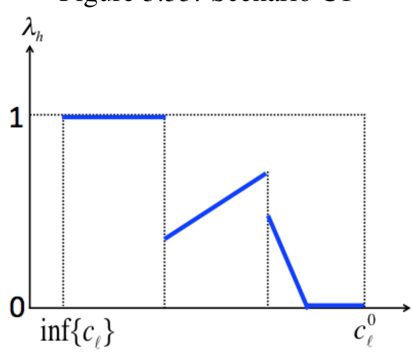

Figure 3.35: Scenario C3

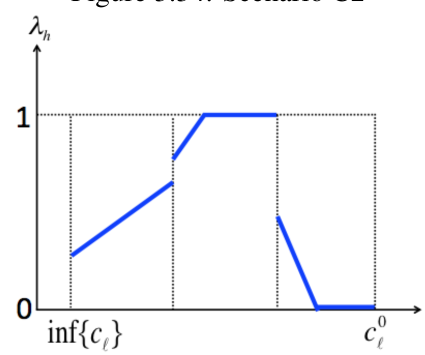

Figure 3.36: Scenario C4

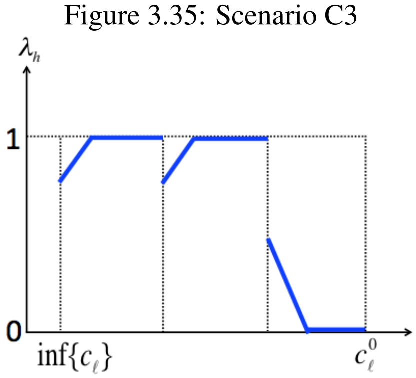

Figure 3.37: Scenario C5

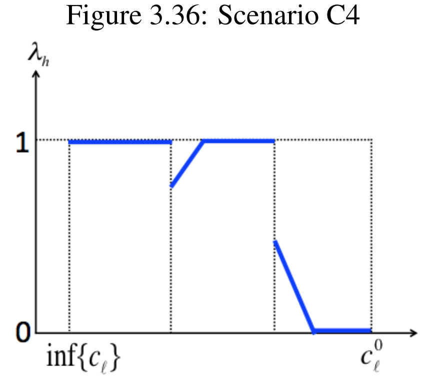

Figure 3.38: Scenario C6

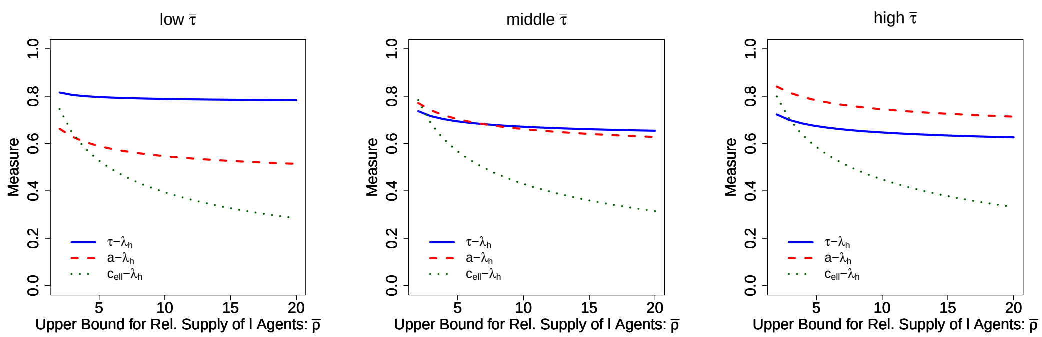

Figure 3.39: Measure of Scenario 0

Note: Low, middle and high $\bar{\tau}$ correspond to values of 1.1, 1.5, and 2.0, respectively.

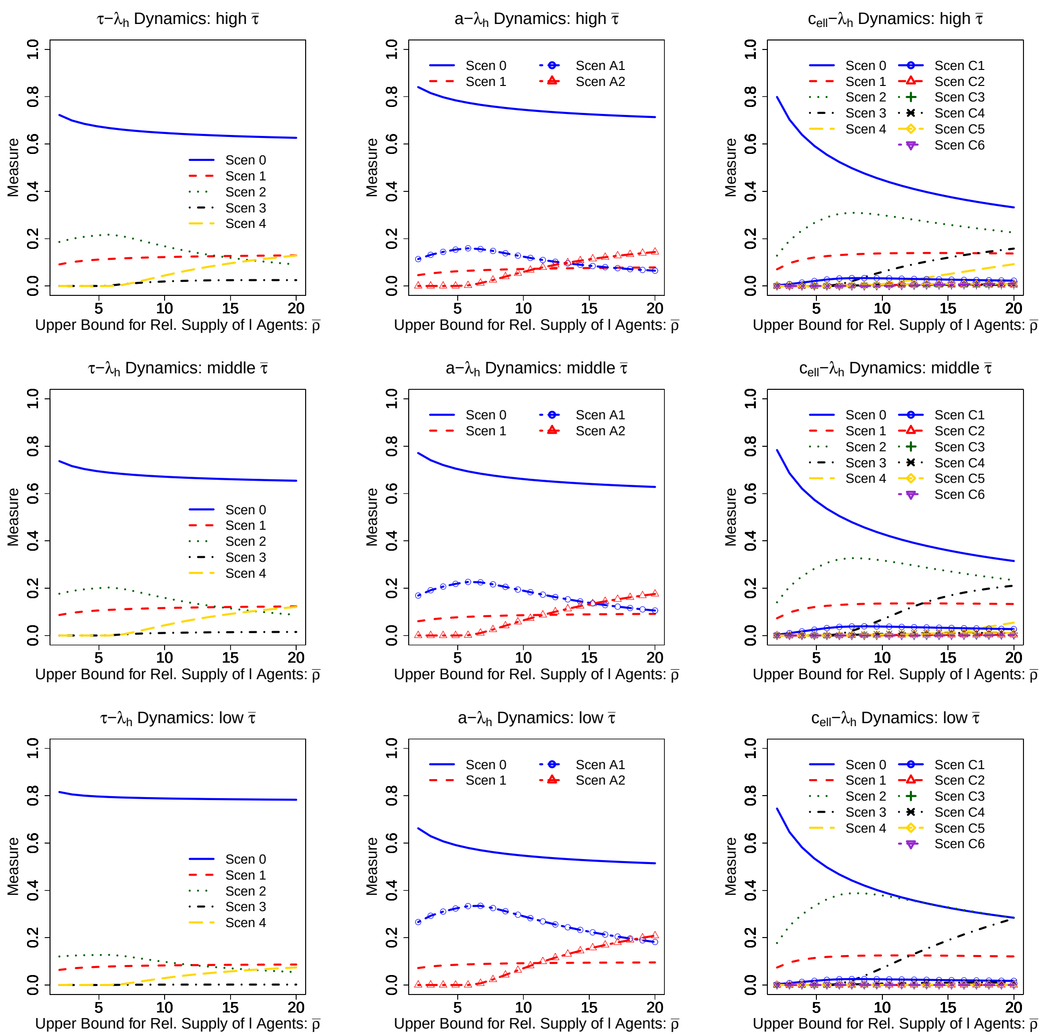

Figure 3.40: Measures of Scenarios in Dynamics

Note: Low, middle and high $\bar{\tau}$ correspond to values of 1.1, 1.5, and 2.0, respectively.

Conclusion

As reported by previous researches, urban specialization has been transforming from traditional sectoral to functional specialization. The latter form of specialization is a system of cities, in which new technologies and varieties of goods and services are developed in larger cities with the help of greater urban diversity and positive externalities, while in smaller cities production based on technologies developed in the former cities are conducted, often characterized as less skill-intensive, more routine economic activity.

Given this motivating fact, the dissertation provides a simple model of functional specialization of cities in Chapter 1 which formalizes mechanisms behind the specialization and also investigates implications of such specialization for welfare and optimal income redistribution across cities. The specialization is a result of the interplay between two counteracting forces, agglomeration and dispersion forces, the former and the latter of which correspond to urban diversity and market competition, respectively. Chapter 2 then complements the analysis in Chapter 1 in a way that heterogeneous agents not considered in Chapter 1 are introduced to show that a similar specialization emerges, being associated with income inequality in the form of skill premium. Finally, in Chapter 3, the role of team production in attracting creativity, especially from the view point of smaller cities, is analyzed in light of the increasing importance of knowledge and creativity in overall economic activity and widening income inequalities suggested by previous researches and chapters. Although such organization changes the nature of comparative advantages of agents and could thus act as a device to attract creativity, the associated effects on the allocation of creativity are likely to be very limited, and even in a case of success in attracting creativity, a careful policy management is necessary due to the non-monotonic effects of policies.

Bibliography

Albouy, David and Gabriel Ehrlich, "Metropolitan Land Values and Housing Productivity," NBER Working Papers 18110, National Bureau of Economic Research, Inc May 2012.

Anderson, James E. and Eric van Wincoop, "Trade Costs," *Journal of Economic Literature*, September 2004, *42* (3), 691–751.

Behrens, Kristian, Gilles Duranton, and Frederic Robert-Nicoud, "Productive Cities: Sorting, Selection, and Agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507 – 553.

_ , Giordano Mion, Yasusada Murata, and Jens Suedekum, "Spatial frictions," Technical Report 2014.

Berliant, Marcus and Masahisa Fujita, "Culture and diversity in knowledge creation," *Regional Science and Urban Economics*, 2012, *42* (4), 648–662.

Black, Duncan and Vernon Henderson, "A Theory of Urban Growth," *Journal of Political Economy*, April 1999, *107* (2), 252–284.

Davis, Donald R. and Jonathan I. Dingel, "A spatial knowledge economy," NBER Working Papers 18188, National Bureau of Economic Research, Inc June 2012.

Davis, Morris A. and Francois Ortalo-Magne, "Household Expenditures, Wages, Rents," *Review of Economic Dynamics*, April 2011, *14* (2), 248–261.

Desmet, Klaus and Esteban Rossi-Hansberg, "Urban Accounting and Welfare," *American Economic Review*, 2013. forthcoming.

Dhingra, Swati and John Morrow, "Monopolistic competition and optimum product diversity under firm heterogeneity," *London School of Economics, mimeograph*, 2012.

Dixit, Avinash K and Gene M Grossman, "Trade and Protection with Multistage Production," *Review of Economic Studies*, October 1982, *49* (4), 583–94.

_ and Joseph E Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, June 1977, *67* (3), 297–308.

Dornbusch, Rudiger, Stanley Fischer, and Paul A Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, December 1977, *67* (5), 823–839.

Duranton, Gilles, "Urban Evolutions: The Fast, the Slow, and the Still," *American Economic Review*, March 2007, *97* (1), 197–221.

_ and Diego Puga, "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," *American Economic Review*, December 2001, *91* (5), 1454–1477.

_ and _ , "From sectoral to functional urban specialisation," *Journal of Urban Economics*, March 2005, *57* (2), 343–370.

Eeckhout, Jan, "Gibrat's Law for (All) Cities," *American Economic Review*, December 2004, *94* (5), 1429–1451.

Feenstra, Robert C. and Gordon H. Hanson, "Foreign Investment, Outsourcing, and Relative Wages," in Robert C. Feenstra, Gene M. Grossman, and Douglas A. Irwin, eds., *The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati*, Cambridge, MA: MIT Press, 1996, pp. 89–127.

Forslid, Rikard and Gianmarco I.P. Ottaviano, "An analytically solvable core-periphery model," *Journal of Economic Geography*, July 2003, *3* (3), 229–240.

Gabaix, Xavier, "Power Laws in Economics and Finance," *Annual Review of Economics*, 05 2009, *1* (1), 255–294.

_ and Yannis M. Ioannides, "The evolution of city size distributions," in J. V. Henderson and J. F. Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4 of *Handbook of Regional and Urban Economics*, Elsevier, April 2004, chapter 53, pp. 2341–2378.

Garicano, Luis and Esteban Rossi-Hansberg, "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 2006, *121* (4), 1383–1435.

_ and _ , "The Knowledge Economy at the Turn of the Twentieth Century: The Emergence of Hierarchies," *Journal of the European Economic Association*, 04-05 2006, *4* (2-3), 396–403.

Grossman, Gene M. and Esteban Rossi-Hansberg, "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, December 2008, *98* (5), 1978–97.

Henderson, J. V., "The Sizes and Types of Cities," *American Economic Review*, September 1974, *64* (4), 640–56.

Hsu, Wen - Tai, "Central Place Theory and City Size Distribution," *Economic Journal*, 09 2012, *122* (563), 903–932.

Kohler, Wilhelm, "International outsourcing and factor prices with multistage production," *Economic Journal*, 03 2004, *114* (494), C166–C185.

Matsuyama, Kiminori, "Complementarities and Cumulative Processes in Models of Monopolistic Competition," *Journal of Economic Literature*, June 1995, *33* (2), 701–729.

_ , "Endogenous Ranking and Equilibrium Lorenz Curve Across (ex ante) Identical Countries," *Econometrica*, 09 2013, *81* (5), 2009–2031.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding, "Task Specialization in U.S. Cities from 1880-2000," CEP Discussion Papers dp1186, Centre for Economic Performance, LSE February 2013.

Moretti, E., *The New Geography of Jobs*, Houghton Mifflin Harcourt, 2012.

Nagamachi, Kohei, "Comparative Advantage and Skill Premium of Regions," *Economics Bulletin*, 2013, *33* (3), 1681–1694.

Pflüger, Michael and Takatoshi Tabuchi, "The size of regions with land use for production," *Regional Science and Urban Economics*, November 2010, *40* (6), 481–489.

Rossi-Hansberg, Esteban and Mark L. J. Wright, "Urban Structure and Growth," *Review of Economic Studies*, 2007, *74* (2), 597–624.

Valentinyi, Akos and Berthold Herrendorf, "Measuring Factor Income Shares at the Sector Level," *Review of Economic Dynamics*, October 2008, *11* (4), 820–835.

Yi, Kei-Mu, "Can Vertical Specialization Explain the Growth of World Trade?," *Journal of Political Economy*, February 2003, *111* (1), 52–102.