CIRJE-F-886

Task Trade and the Size Distribution of Cities

Kohei Nagamachi
University of Tokyo

April 2013; Revised in May 2013

Task Trade and the Size Distribution of Cities

Kohei Nagamachi[1]

Graduate School of Economics, The University of Tokyo

May 31, 2013

Abstract

Taking account of the increasing importance of task trade in urban contexts, this paper provides a model of a system of cities in which ex ante identical locations specialize in tasks that differ in their skill intensity, resulting in a unique size distribution of cities. The necessary and sufficient condition for a power law including Zipf's law is derived, and a quantitative analysis shows that the model is consistent with the size distribution of U.S. cities. A welfare analysis is also conducted, suggesting excess agglomeration with sizable welfare loss under laissez-faire.

Keywords: size distribution of cities, task trade, spatial equilibrium, symmetry breaking
JEL classification: F12, R12, R13

[1]We would like to thank Kiminori Matsuyama, Takatoshi Tabuchi, and Daichi Shirai for their helpful and insightful comments and suggestions. All remaining errors are ours. We also acknowledge the financial support by the JSPS Grant-in-Aid for Research Activity Start-up No.24830025.
E-mail: kohei.nagamachi@gmail.com

1 Introduction

Urban economics is now undergoing what trade theory had experienced — a refinement of traditional thinking, recognizing the increasing importance of tasks as an essential unit of modern economic activities (Duranton and Puga, 2005). However, what distinguishes this refinement in urban economics from that in the trade literature is the focus of economists on the relation of tasks with urban diversity. In their seminal work, Duranton and Puga (2001) discuss the task trade between diversified and specialized cities. The former specialize in the development of new products and their appropriate production technologies, whereas the latter specialize in stylized production based on the methods developed in the former cities.[2] Urban diversity, which generally involves a network of people with different ideas, helps to provide solutions to particular problems or develop new ideas, acting as an agglomeration force or so-called urbanization economies (Jacobs, 1969). Due to this agglomeration force, cities can exist even if they face substantial congestion diseconomies.[3] The specialization of cities by task then results in regional disparities reflecting variations in the relative magnitudes of these counteracting forces.

Therefore, we have good reason to examine the implications of task trade for a system of cities, especially with respect to the size distribution of cities. However, to our knowledge, no single paper has investigated the implications of task trade across cities for the size distribution of cities using a rigorous general equilibrium framework.[4] Instead, the literature of the size distribution either does not focus on specialization of cities itself or focuses on the specialization of cities in different

[2] More precisely, the model provided by Duranton and Puga (2001) is the one of process innovation. However, the same mechanism seems to work in the case of product innovation. Empirical studies such as Feldman and Audretsch (1999) suggest this type of mechanism.

[3] Empirical studies such as Feldman and Audretsch (1999) and Davis and Henderson (2008) support this line of thinking by verifying the importance of urban diversity and differentiated local service supplies in promoting innovative activities and enhancing the productivity of firms. Furthermore, the increasing importance of interactive tasks empirically shown by Michaels et al. (2013) indirectly suggests urban diversity as a clue to understanding urban agglomeration.

[4] Although the hierarchy model of Beckmann (1958) is interpreted as a model with task trade, it is not a general equilibrium model. Models of the central place theory could be another alternative. However, Hsu (2012) focuses on industry-level factors rather than tasks. More importantly, our focus is the interactions between economic agents and product varieties within each city through either market or non-market activities, which are not stressed in models of the central place theory.

industries. However, as stressed by Duranton and Puga (2005), the nature of the specialization of cities has been changing: from sectoral to functional specialization.

The purpose of this paper is to fill this gap between theory and reality. To this end, we extend Matsuyama's (2013) international trade model to a spatial equilibrium model of a system of cities with task trade in which ex ante identical locations specialize ex post in different sets of tasks. Tasks differ in their skill intensity, and those with higher skill intensity require complex and differentiated services, characterized by monopolistic competition á la Dixit and Stiglitz (1977), to a greater extent in production. Product varieties here act as an agglomeration force, that ensures that concentration into a particular location, despite being associated with higher congestion costs, is sustainable. Thus, the current model employs the traditional mechanism for the existence of cities and their disparities while adding a new mechanism of specialization through task trades, which affects the balance between agglomeration and dispersion forces and thus the size distribution of cities.

We then show that an equilibrium with specialization of cities through task trade 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 which we interpret as urban diversity. We also 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.

A contribution of this paper with respect to these analytical results is that it develops an interpretation that allows us to ensure that the 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 crucial step in the interpretation is the introduction of competitive developers, which appear in studies from the urban literature such as Henderson (1974) and Rossi-Hansberg and Wright (2007). The model accommodates the entry and exit of developers and thus is consistent with the active creation and destruction of cities that is reported by Henderson and Wang (2007). Yet these developers are not perfect in the sense that their tool for competition is limited to subsidies to workers, which implies that the market outcome is inefficient due to the unresolved distortion arising from monopolistic competition.

Our another and main contribution is that we test the theory using data not just deriving the condition for a power law. More specifically, we calibrate the model using an occupational dataset for the United States and compare the prediction of the model for the size distribution of cities to the actual size distribution. Interestingly, the model can reproduce the observed size distribution fairly well under the hypothesis that the spatial allocation of tasks is well captured by the theory.

Therefore, using the calibrated model, we also conduct a welfare analysis in order to investigate whether laissez-faire is associated with excess agglomeration. More specifically, we extend the model by introducing a government implementing an income redistribution policy with an income tax and cross-city lump-sum transfer. Welfare is measured in terms of the equilibrium utility of identical workers who are freely mobile, thus have a same utility level and are the only agents who derive utility from consumption in the economy. The results show that laissez-faire is associated with excess agglomeration, which is Pareto-dominated by not only *autarky* but also equilibrium with *any* non-zero income tax rate.[5] More specifically, the relationship between the tax rate and equilibrium utility is hump-shaped with a single peak: utility under laissez-faire is 6.3% smaller than that under autarky; then the utility increases to its highest level, 32.0% larger than that under autarky or 38.3% larger than that under laissez-faire as the tax rate increases to 5.77%; and a further increase in the tax rate reduces utility, which gradually converges to the level achieved under autarky as the tax rate converges to 100%. That is, the welfare loss due to the excess agglomeration under laissez-faire is substantial.

This paper is related to three strands of literature. The first strand is the literature on the international trade theory of Ricardian comparative advantage. The current model is an application of Matsuyama's (2013) model to the regional con-

[5] Given an externality, it is not surprising that an equilibrium with some positive income tax rate is more efficient than in the laissez-faire case. However, what is surprising is that even compared to an equilibrium with a sufficiently high tax rate, which results in a dispersed spatial distribution and thus reduces the benefits from agglomeration, laissez-faire is still inefficient.

text. Matsuyama 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 extends the analysis to multiple and arbitrarily large number of countries by developing a new method in which an equilibrium reduces to a second-order difference or differential equations. The crucial differences between his and ours are as follows: First, we interpret the sectors in his analysis as tasks in order to take account of the importance of specialization of cities by task as argued by Duranton and Puga (2005). Second, we introduce developers, as mentioned above. Third, we consider spatial equilibrium by introducing migration of workers and land, the latter of which is used for consumption as well as production. Finally and importantly, we apply the model to empirical data.

The second strand of literature is the research on the size distribution of cities. To our knowledge, this is the first general equilibrium model that discusses both theoretical and quantitative implications of the specialization of cities through task trade for the size distribution of cities. 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. Although the recent study by Behrens et al. (2010) takes account of the findings of Hendricks (2011) and consequently focuses on within-industry aspects as this paper does, our study differs from theirs, which focuses on heterogeneous entrepreneurs within a framework á la Melitz (2003). Model-based quantitative studies such as Desmet and Rossi-Hansberg (2013) and Behrens et al. (2013) are also related to this research. However, our model focuses on task trade and abstracts heterogeneous preference and exogenous productivity differentials.

The third is trade models which specify the production of goods as a continuum of fragments, intermediate goods, or tasks. The current model does not include the trade costs of tasks and thus heterogeneous trade costs, introduced by Grossman and Rossi-Hansberg (2008) in order to validate the interpretation of each element in the continuum as a task. Instead, we allow such an interpretation by calibrating the model with data closely related to the tasks. Therefore, although the model specification is somewhat similar to previous models with a continuum of fragments or intermediate inputs, we have a clear-cut rationale for interpreting each element in the continuum as a task.[6] Even without heterogeneous trade costs, the fraction of traded tasks is determined endogenously, reflecting the comparative advantages of cities that are in turn determined endogenously.

The remainder of this paper is organized as follows. We first provide the model in Section 2. In Section 3, we discuss the equilibrium properties and the theoretical implications of task trade for the size distribution of cities. We then calibrate the model and investigate its quantitative implications in Section 4. Using the calibrated model, we also conduct a welfare analysis by introducing an income redistribution policy using an income tax and lump-sum transfers across cities in Section 5. Finally, we conclude this paper in Section 6.

2 The Model

In this section, we provide a simple spatial equilibrium model with a continuum of tasks that are traded across cities and within firms. The economic agents consist of mobile workers, developers, final good firms, and local firms within each city, the last of which include monopolistically competitive and perfectly competitive firms. The model is essentially an application of Matsuyama's (2013) framework to the urban context.

In the following, we explain the optimization problems of all 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, each of which is endowed with one unit of land.[7] We

[6] 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.

[7] In this paper, we do not distinguish cities, regions, and locations and use these words inter-

subsequently modify this assumption by making J diverge to infinity but making the mass of each location converges to zero in such a way that the total mass of locations is equal to unity, which allows us to consider the distribution of developers and accommodate their free entry and exit.

2.1 Workers

There is 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 $\bar{R}_j$, the latter of which is received from the developer who managing location j as discussed in the next subsection. She uses these sources of income to consume homogeneous tradeable goods c_j and housing h_j, the prices of which are P and R_j, respectively. The constancy of the expenditure share α of housing consumption is consistent with the prior studies such as Davis and Ortalo-Magne (2011).

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

$$(1) \qquad \frac{W_j + \bar{R}_j}{W_{j'} + \bar{R}_{j'}} = \left(\frac{R_j}{R_{j'}} \right)^{\alpha} \qquad \forall j \neq j',$$

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

changeably.

2.2 Developers

As in urban models such as Henderson (1974) and Rossi-Hansberg and Wright (2007), among others, we introduce competitive developers competing for workers using subsidies $\bar{R}_j$ whose revenues consist of their rents from the cities that they manage in order to take account of the active creation and destruction of cities reported by Henderson and Wang (2007).

The associated zero-profit condition then implies that[8]

$$\bar{R}_j = \frac{R_j}{N_j} \quad \forall j, \tag{2}$$

where N_j denotes the population of location j, and R_j is interpreted as the total land rents given one unit of land. Note that this is the relationship that applies to location j, at which some particular developer succeeded in attracting workers and has thus revealed its existence in the economy. Stated differently, J is the number of developers revealed in this manner, and although its finiteness is thus arbitrary, it is relaxed when J becomes arbitrarily large and we consider the *distribution* of developers that is presumed to be observed under free entry and exit.

2.3 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 requires a continuum of tasks $\{t : 0 \leq t \leq 1\}$:[9]

$$Y = \exp\left[\int_0^1 \ln(y(t))dt\right], \tag{3}$$

[8] We abstract the endogenous determination of the physical area of each city because our focus is not on how the urban structure is determined.

[9] In Matsuyama (2013), what we consider to be a continuum of tasks here is interpreted as a continuum of sectors, and the technology specified by (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).

where Y and $y(t)$ denote outputs of the final good and task t, respectively. The equal weights and unit mass of the tasks 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 task t.

Firms decide where each of these tasks is performed. For each fixed task $t \in [0,1]$, a typical firm decides the quantity $y_j(t)$ of production of task t at an existing location $j \in \{1,2,\cdots,J\}$. Once the task has been performed at each location, outputs $\{y_j(t)\}_{j=1}^{J}$ are aggregated, and the result is used as a production input:

$$y(t) = \sum_{j=1}^{J} y_j(t) \quad \forall t. \tag{4}$$

Furthermore, task t performed at location j is a combination of skill-intensive services $X_{S,j}(t)$ and labor-intensive services $X_{L,j}(t)$ by local suppliers. Its output $y_j(t)$ is given by

$$y_j(t) = X_{S,j}(t)^{\gamma(t)} X_{L,j}(t)^{1-\gamma(t)} \quad \forall j,t, \tag{5}$$

where $\gamma(t) \in [0,1]$ represents the skill-intensity of task 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 $\gamma(1) = 1$.[10]

Therefore, letting $P_{S,j}$ and $P_{L,j}$ denote the prices of the skill-intensive and labor-intensive services, respectively, we can write the profit maximization of the final good firm as follows:

$$\max PY - \sum_{j=1}^{J} \int_{\mathbb{T}_j} \left[P_{S,j} X_{S,j}(t) + P_{L,j} X_{L,j}(t)\right] dt \quad s.t. \quad (3)-(5), \text{ and } \mathbb{T}_j \subseteq [0,1],$$

where the control variables consist of the *measurable set* $\mathbb{T}_j$ of tasks performed at location j as well as quantities $(Y, \{y(t)\}_t, \{y_j(t), X_{S,j}(t), X_{L,j}(t)\}_{j,t})$. 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,

[10] We generalize this assumption in Section 4.

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 task t are distributed to the skill-intensive and labor-intensive sectors, respectively. Thus, the skill-intensive 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 skill-intensive and labor-intensive sectors are monopolistically and perfectly competitive, respectively.[11] More specifically, $X_{S,j}(t)$ denotes the composite of a continuum of varieties $\{x_{S,j}(v,t)\}_{v\in[0,D_j]}$. The quantity of each variety is also a control variable, with the following technology:

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

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 task t performed at location j:

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

where $P_{S,j}$ is the price index of the skill-intensive services at location j given by

$$P_{S,j} = \left[\int_0^{D_j} p_{S,j}(v)^{-\frac{1}{\theta}}dv\right]^{-\theta} \quad \forall j. \tag{6}$$

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

[11] This stylized specification reflects our view of the nature of skill-intensive services such as management, research and development, and legal services as well as labor-intensive services such as line production based on previously developed blueprints.

2.4 Skill-intensive Local Service Sector

As mentioned in the previous subsection, the skill-intensive 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

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

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 $p_{S,j}(v) = (1+\theta) m R_j^{\beta} W_j^{1-\beta}$, and substituting this into (6) and taking the ratio across two different locations, j and j', we obtain

$$\frac{P_{S,j}}{P_{S,j'}} = \left(\frac{R_j}{R_{j'}} \right)^{\beta} \left(\frac{W_j}{W_{j'}} \right)^{1-\beta} \left(\frac{D_j}{D_{j'}} \right)^{-\theta} \quad \forall j \neq j'. \tag{7}$$

2.5 Labor-intensive Local Service Sector

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

$$\max_{H_{L,j},\ L_{L,j}} P_{L,j} H_{L,j}^{\beta} L_{L,j}^{1-\beta} - R_j H_{L,j} - W_j L_{L,j}.$$

Note that the production of labor-intensive services also requires both labor and land. For simplicity, we assume the same cost share parameter β as in the skill-intensive sector.

The profit maximization then implies

$$\frac{P_{L,j}}{P_{L,j'}} = \left(\frac{R_j}{R_{j'}}\right)^{\beta} \left(\frac{W_j}{W_{j'}}\right)^{1-\beta} \quad \forall j \neq j'. \tag{8}$$

2.6 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 skill-intensive sector. That is, if cities are different, we should observe a perfect sorting of tasks, and the higher the location index j is, the higher the average skill intensity Γ_j should be. Note that this configuration of the equilibrium is consistent with the argument advanced by 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 skill-intensive and labor-intensive 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. *A sorting equilibrium is a set of prices* $(P, \{P_{S,j}, P_{L,j}, R_j, W_j\}_j, \{p_{S,j}(v)\}_{v,j})$, *quantities* $(Y, \{c_j, h_j, H_{L,j}, L_{L,j}\}_j, \{y(t)\}_t \{y_j(t), X_{S,j}(t), X_{L,j}(t)\}_{j,t}, \{x_{S,j}(v,t)\}_{v,t,j}, \{q_j(v)\}_v)$, *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:*

$$R_j = (1-\alpha)\beta|\mathbb{T}_j|E + \alpha(W_j N_j + \bar{R}_j N_j), \quad \text{(Land)} \tag{9}$$

$$W_j N_j = (1-\alpha)(1-\beta)|\mathbb{T}_j|E; \quad \text{(Labor)} \tag{10}$$

4. *there is free entry into developer and skill-intensive local service sectors; and*
5. *thresholds* $\{T_j\}_{j=0}^J$ *are consistent with comparative advantage:*

$$\frac{P_{j+1}(T_j)}{P_j(T_j)} = \left(\frac{P_{S,j+1}}{P_{S,j}}\right)^{\gamma(T_j)} \left(\frac{P_{L,j+1}}{P_{L,j}}\right)^{1-\gamma(T_j)} = 1, \tag{11}$$

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

Here, the market clearing conditions, i.e., (9) and (10), are evident from the specifications of the utility and production technologies presented in the previous subsections. Those of goods markets are omitted for ease of exposition. The conditions that $P_{S,j+1}/P_{S,j} < 1$ and $P_{L,j+1}/P_{L,j} > 1$ imply that location $j+1$, compared to location j, has a comparative advantage in the skill-intensive service sector and thus has a comparative advantage in tasks with higher skill intensity. The threshold T_j here is the task for which locations j and $j+1$ have the same comparative advantage.

3 Equilibrium Properties and the Size Distribution of Cities

In this section, we first consolidate the equilibrium condition presented in the previous section to obtain the fundamental equation governing the equilibrium of the economy in Subsection 3.1. We then prove the existence and uniqueness of a sorting equilibrium in Subsection 3.2, which are consistent with the various observed facts. The theoretical implications of our model for the size distribution of cities are also derived in Subsection 3.3.

3.1 Fundamental Equation

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

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

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 (9) and (10) along with the free-entry condition of the developer sector, (2), yields

$$\text{(12)} \qquad \frac{R_{j+1}}{R_j} = \frac{W_{j+1}N_{j+1}+R_{j+1}}{W_jN_j+R_j} = \frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|} \quad \forall j.$$

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 are. In addition, it is also implied that differentials of land rent R_j and market share $|\mathbb{T}j$ are equal.

This result is then combined with the free-migration condition (1) to obtain

$$\text{(13)} \qquad \frac{N_{j+1}}{N_j} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{1-\alpha} \quad \forall j,$$

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 (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 (10) implies that the differential of labor compensation $W_j N_j$ is equal to that of the market share $|\mathbb{T}_j|$:

$$(14) \qquad \frac{W_{j+1}}{W_j} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{\alpha} \quad \forall j.$$

These results immediately imply that a higher market share $|\mathbb{T}_j|$ is associated with higher congestion costs, which is a *dispersion force*, in the sense that the unit production cost and thus the price $P_{L,j}$ of the labor-intensive local services is higher in that location. More specifically, substituting (12) and (14) into (8), we obtain

$$(15) \qquad \frac{P_{L,j+1}}{P_{L,j}} = \left(\frac{|\mathbb{T}_{j+1}|}{|\mathbb{T}_j|}\right)^{\alpha(1-\beta)+\beta} \quad \forall j.$$

Another important implication of this result is that a location with a higher market share $|\mathbb{T}_j|$ is likely to exhibit comparative advantage in the production of skill-intensive services. This can be seen by discussing the determination of urban diversity, a factor that generates comparative advantage in skill-intensive services. Substituting (7), (12), (14) and (15) into (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, higher congestion costs in a location with a higher market share $|\mathbb{T}_j|$ should be associated with greater urban diversity D_j, leading the location to exhibit comparative advantage in the production of skill-intensive 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 task trade within firms and across locations. In a sorting equilibrium, if it were to exist, regional disparities would emerge as a balance between the agglomeration and dispersion forces, and this balance would be affected by the task trade reflecting the function $\gamma(t)$ and, thus, the distribution of the skill intensities in the economy.

Finally, by utilizing the free-entry condition into the skill-intensive 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 skill-intensive sector yields $D_j p_{S,j} q = (1-\alpha)\Gamma_j |\mathbb{T}_j| E$, or taking the ratio of this equation, we obtain

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

As mentioned in the previous section, the finiteness of J is arbitrary in the sense that it does not accommodate free entry into the developer sector. To avoid this arbitrariness, we investigate the distribution of the existing developers or cities in a sorting equilibrium. We accomplish this by making J diverge to infinity while limiting the total mass of cities to unity. Then, applying the Matsuyama's (2013) method to the fundamental equation,[12] we obtain the following boundary value problem for a second-order ordinary differential equation (ODE):

$$\text{(16)} \qquad \frac{\Phi''}{\Phi'} = \frac{\theta\gamma'(\Phi)\Phi'}{G(\Phi)} \quad \text{with } \Phi(0) = 0 \text{ and } \Phi(1) = 1,$$

where

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

Each city is characterized by the task 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 existing 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 task 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 task, we call the city that hosts task t

[12] 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 task in the limit and thus is characterized by t.

city t. We assume that $G(1) > 0$ in order to focus on a meaningful case.[13]

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 (16). Importantly, we can obtain a unique solution to the fundamental equation analytically, which 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 task each city specializes in ex post.[14]

More specifically, we obtain the inverse function of the Lorenz curve (let $H : z \to t$ denote the function, i.e., $H(z) \equiv \Phi^{-1}(z)$):

$$H(z) = \frac{\int_0^z G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du}{\int_0^1 G(u)^{\frac{1}{(1-\alpha)(1-\beta)}} du} \quad \forall z \in [0,1]. \tag{17}$$

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. *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*

[13] Intuitively, this assumption implies that the magnitude of congestion costs, i.e., the power $\alpha(1-\beta)+\beta$ appearing in (15), 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.

[14] 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 quantitative analysis discussed in Section 4. If we have a cross-sectional dataset of cities, we can calibrate and test the model.

addition, this equilibrium has the following properties: The market share 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 skill-intensive sector) in city t are given by

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

for all $t \in [0,1]$. Therefore, as t increases, i.e., as a location specializes in a more skill-intensive task, 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 (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 skill-intensive services, i.e., $D_j p_{S,j} q =$

$(1-\alpha)\Gamma_j|\mathbb{T}_j| = (1-\alpha)\int_{T_{j-1}}^{T_j}\gamma(t)dt$, 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 $\gamma'(t), \Phi'(t), \Phi''(t) > 0$. □

The comovement across these variables seems fit the reality. As argued in the previous subsection, the result that $R'(t) > 0$ derives from urban congestion. The comovement between $W(t)$ and $N(t)$ is an implication of the spatial equilibrium, and the 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 skill-intensive sector is less than that of the market size of the sector.[15]

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. (2010).

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 condi-

[15] 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

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

tion under which the associated sorting equilibrium exhibits a power law including Zipf's law as a special case. As will be discussed below, this is equivalent to identifying the functional form of $\gamma(t)$ guaranteeing that the size distribution of cities obeys Zipf's law. 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, especially random growth models. However, this is not necessarily the case, at least if the approach is consistent with a power law. This is because a static model, regardless of whether it is a symmetry-breaking one, requires us to assume a non-degenerate distribution of some exogenous variable *and* because assuming such a distribution itself is not inconsistent with random growth factors given the result shown clearly by Gabaix (2009): the random growth of some variable with a lower bound results in a Pareto distribution. In addition, it is noteworthy that we calibrate $\gamma(t)$ and then test the model in Section 4 instead of assuming a particular function $\gamma(t)$ that results in Zipf's law.

The next proposition states the implications of our model for the size distribution of cities:

Proposition 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*

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

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, 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 (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

$$\gamma'\left[\gamma^{-1}(B(\lambda))\right] \propto \lambda^{\tilde{\eta}}, \quad B(\lambda) \equiv \frac{\alpha(1-\beta)+\beta-\lambda^{-1}}{(1-\alpha)(1-\beta)}. \tag{18}$$

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 task 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 (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. □

4 Quantitative Analysis

In this section, we apply the model to empirical data. In Subsection 4.1, we describe our methodology, our procedure and the results of the calibration. Then, in Subsection 4.2, we report the implications of the calibrated model for the size distribution of cities.

4.1 Calibration

In order to compute the equilibrium numerically, we need to specify the values of the parameters (α, β, σ) and the functional form of $\gamma(t)$. However, before proceeding to the calibration of these primitives, we should note that our model does not concern how the distribution of the skill intensity $\gamma(t)$ is determined; rather, we examine how the distribution of income, the market share of skill-intensive and labor-intensive sectors, or population is determined for a given distribution of skill intensity $\gamma(t)$. In addition, we also note that the population distribution is only indirectly related to those of the other variables through general equilibrium effects. Therefore, testing the model by examining the size distribution of cities allows us to use the other distribution as a target of the calibration. Here, as a specific target, we employ the distribution of the market size $\gamma(\Phi(t))\Phi'(t)$ of the skill-intensive sector.[16]

Given this consideration, we first generalize $\gamma(t)$ in such a way that $0 \leq \gamma(0) = \gamma_0 < \gamma_1 = \gamma(1) \leq 1$ and then specify the functional form of $\gamma(t)$ as follows by guessing that the upper tail of the market size is well approximated by a power law:

$$\gamma(t) = \begin{cases} \tilde{a} - \{(\tilde{a}-\gamma_0)^\eta - [(\tilde{a}-\gamma_0)^\eta - (\tilde{a}-\gamma_1)^\eta]\, t\}^{\frac{1}{\eta}} & \text{if } \eta \neq 0, \\ \tilde{a} - \exp\{\ln(\tilde{a}-\gamma_0) - [\ln(\tilde{a}-\gamma_0) - \ln(\tilde{a}-\gamma_1)]\, t\} & \text{if } \eta = 0, \end{cases}$$

where $\tilde{a} > 1$ is a parameter that is different from a in Proposition 2. Thus, instead of assuming an exact power law directly, here we consider a slightly more general $\gamma(t)$ that includes an exact power law as a special case.

It is then suggested that we need to add two restrictions to the parameters (γ_0, γ_1) to bound the max-min ratio of the market size of the skill-intensive sector as observed in the data and to address the identification problem. More specifically, we assume that $\gamma_0 > 0$ and $\gamma_1 = 1$. The former restriction follows immediately if we

[16] An analogous example from the macroeconomics context is that macro economists using Bewley-type dynamic general equilibrium models with heterogeneous agents, e.g., individuals with exogenous heterogeneity in labor efficiency, calibrate the stochastic process of labor efficiency in such a way that the model replicates the observed income distribution (Huggett (1993) and Aiyagari (1994) are classical examples). However, they evaluate their performance by comparing the model with the data in terms of the distribution of an asset or that of another endogenous variable that is at the center of the discussion.

notice that we need to use the market size of the skill-intensive sector normalized by the minimal one, i.e., $\gamma(\Phi(0))\Phi'(0) = \gamma_0\Phi'(0)$. The latter restriction derives from the formula for determining the normalized market size $m_s(t)$ of the skill-intensive sector in city t:

$$
\begin{aligned}
m_s(t) &\equiv \ln\left[\frac{\gamma(\Phi(t))\Phi'(t)}{\gamma(0)\Phi'(0)}\right] = \ln\left[\frac{\gamma(\Phi(t))}{\gamma_0\Phi'(0)}\right] + \ln\left[\frac{\Phi'(t)}{\Phi'(0)}\right] \\
&= \ln\left[\frac{\gamma(\Phi(t))}{\gamma_0}\right] + \frac{1}{(1-\alpha)(1-\beta)}\ln\left[\frac{\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma_0}{\alpha(1-\beta)+\beta-(1-\alpha)(1-\beta)\theta\gamma(\Phi(t))}\right]
\end{aligned}
$$

for all $t \in [0,1]$, where the second line follows from the expression for $\Phi'(t)$, obtained by differentiating $t = H[\Phi(t)]$ with respect to t and using the inverse function theorem. This clearly suggests that if the values of the other parameters are taken as given, there are likely multiple pairs of (γ_0, γ_1) that achieve the same $m_s(t)$.

Regarding the other parameters, we calibrate α and β independently in order to approximate the observed housing expenditure share, reported by Davis and Ortalo-Magne (2011), and the cost share of non-labor production factors, reported by Valentinyi and Herrendorf (2008). σ is calibrated in such a way that the model can match the natural logarithm of the observed max-min ratio of the market size of all sectors, which is given by

$$
\ln\left[\frac{\Phi'(1)}{\Phi'(0)}\right] = \frac{1}{(1-\alpha)(1-\beta)}\ln\left[\frac{G(0)}{G(1)}\right].
$$

Formally, we solve the following constrained minimization problem in a brute force manner with a discretized parameter space:

$$
\begin{aligned}
&\min_{\tilde{a},\eta,\gamma_0} \max_{i\in\{1,2,\cdots,\hat{N}\}} \left|\hat{F}_{m_s}(m_i) - F_{m_s}(m_i;\tilde{a},\eta,\gamma_0)\right| \\
&s.t. \\
&\tilde{a} > 1, \qquad 0 < \gamma_0 < 1, \\
&m_{\text{data}} = \frac{1}{(1-\alpha)(1-\beta)}\ln\left[\frac{G(0)}{G(1)}\right],
\end{aligned}
$$

where $\hat{F}_{m_s}$ and F_{m_s} are the distribution functions of the natural logarithm of the mar-

Parameter	Meaning	Restriction	Target	Source
α	Housing expenditure share	$(0,1)$	Housing expenditure share	Davis and Ortalo-Magne (2011)
β	Cost share of housing	$(0,1)$	Cost share of non-labor factors	Valentinyi and Herrendorf (2008)
σ	Elasticity of substitution	$(1,+\infty)$	Max-Min ratio of labor compensation	Occupational Employment Statistics
$\tilde{a}$	Parameter of $\gamma(t)$	$(1,+\infty)$	Distribution of the skill share	Occupational Employment Statistics
η	Parameter of $\gamma(t)$	$(-\infty,+\infty)$		
γ_0	$\min\gamma(t)$	$(0,1)$		
γ_1	$\max\gamma(t)$	$\gamma_1 = 1$		

Table 1: Restrictions on and Targets of the Calibrated Parameters

α	β	σ	$\tilde{a}$	η	γ_0	γ_1
0.240	0.300	2.163	1.011	-0.477	0.480	1.000

Table 2: Calibrated Parameters

ket size of the skill-intensive sector for the data and the model, respectively. m_{data} is the natural logarithm of the observed max-min ratio of the market size of all sectors. We discretize the function with $\hat{N}$ equidistant grid points and evaluate the distance between two distributions using the maximal deviation. Table 1 summarizes the restrictions on and targets of the calibrated parameters.

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 the skill-intensive sector and the normalized labor compensation of that sector 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 of the skill-intensive sector from the data on labor compensation. We assume that the skill-intensive and labor-intensive occupations are those listed in Table 3. These occupations cover approximately 50.9% of the total occupations in the U.S. (17.1% for the skill-intensive sector and 33.8% for the labor-intensive sector), reflecting the exclusion of occupations that do not seem to be well captured by the model, such as "Farming, Fishing, and Forestry Occupation (45-0000)" and "Transportation and Material Moving Occupations (53-0000)". We should also note that not all MSAs are included in the sample, partly because of a lack of data for some of the major

Skill-intensive Occupations	Labor-intensive Occupations
Management Occupations (11-0000)	Sales and Related Occupations (41-0000)
Business and Financial Operations Occupations (13-0000)	Office and Administrative Support Occupations (43-0000)
Computer and Mathematical Occupations (15-0000)	Production Occupations (51-0000)
Architecture and Engineering Occupations (17-0000)	
Life, Physical, and Social Science Occupations (19-0000)	
Legal Occupations (23-0000)	
Arts, Design, Entertainment, Sports, and Media Occupations (27-0000)	

Table 3: Definition of Skill-intensive and Labor-intensive Occupations

occupations listed in Table 3. However, more importantly, we restrict our attention to the upper tail of the distribution in the sense that, after excluding MSAs that are too small, the model can replicate the distribution of the natural logarithm of the market size of the skill-intensive sector fairly well. In effect, 335 MSAs are covered after sample selection. This figure is not particularly different from those in studies such as Rossi-Hansberg and Wright (2007). It is implied that $m_{\text{data}} = 5.96$.

The results of the calibration are reported in Table 2. As shown in Figure 1, the model replicates the observed distribution of the natural logarithm of the market size of the skill-intensive sector with the maximal deviation of 2.30%.

4.2 Quantitative Implications

Interestingly, our model successfully reproduces the observed city size distribution. As shown in Figure 2, the rank-size plot of the model (the blue line) appropriately approximates that of the data (the red dashed line) where the latter corresponds to the sample that contains the largest 335 MSAs in the "Annual Estimates of the Population of Metropolitan Statistical Areas" compiled by the Census Bureau. In order to derive the prediction of the model, we first discretize the space of t, i.e., $[0,1]$, with 335 equidistant grid points and then compute the natural logarithm of population for each point using

$$\ln\left[\frac{N(t)}{N(0)}\right] = \frac{1}{1-\beta}\ln\left[\frac{G(0)}{G(\Phi(t))}\right].$$

Does this result imply that the model successfully explain the observed size

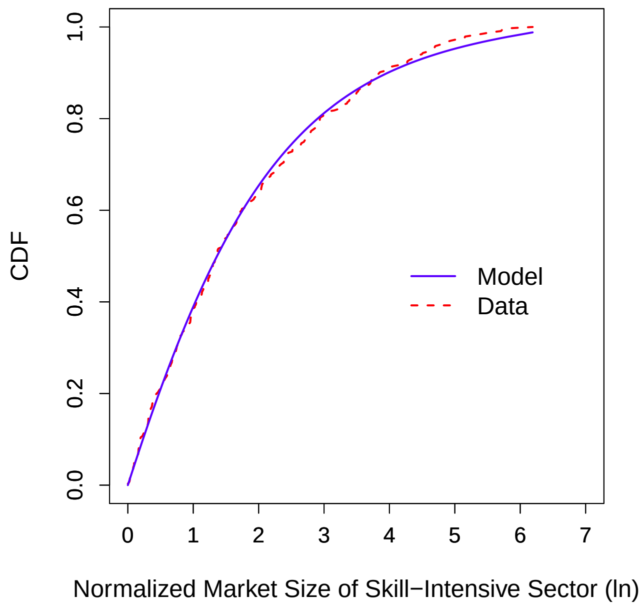

Figure 1: Cumulative Distribution Function of the Normalized Market Size of the Skill-intensive Sector

Note: The cumulative distribution of the data is constructed using the labor compensation of the skill-intensive occupations as defined in Table 3.

Source: May 2011 Occupational Employment Statistics.

distribution of cities? We note that the result is conditional on the hypothesis that the observed cross-city variations in the market size of the skill-intensive sector are well captured by the model. Thus, one direction for future research is to investigate the determinants of the economy-wide distribution of skill intensity. Important examples include studies on so-called directed technical change such as Acemoglu (2002). Moreover, because the size distribution of cities is stable, there should not be significant change in the distribution of skill intensity for the normalized market size of the skill-intensive sector to be stable if the model is promising. This point raises another interesting question. Is it possible for the distribution of skill intensity to be stable even if the economy has experienced skill-biased technical changes as the U.S. has since the late 1970s? Unfortunately, because there is no consistent

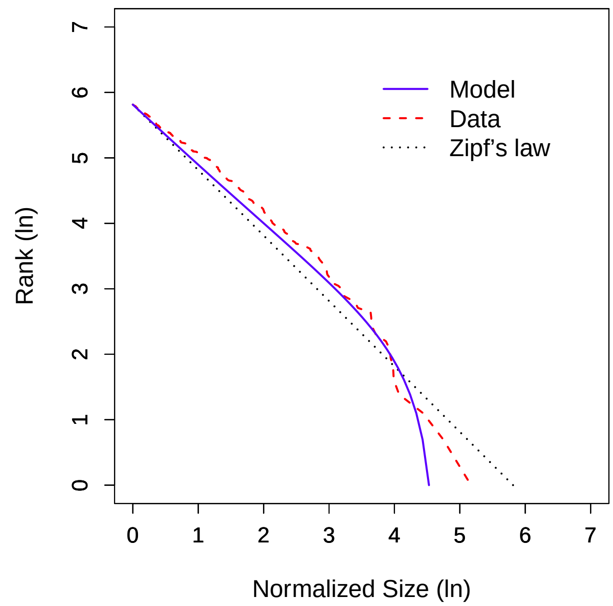

Figure 2: Rank-Size Plot

Note: The rank-size plot for the data includes the top 335 U.S. MSAs.
Source: U.S. Census Bureau.

panel dataset (at least for the U.S.), this type of time series analysis is limited to some extent.

5 Welfare Analysis of Income Redistribution Policy

Finally, we extend the model by including a government that implements an income redistribution policy in order to examine whether the market equilibrium is characterized by excess agglomeration.[17] In addition, we compute the optimal income tax rate under the income redistribution policy specified in Subsection 5.1.

[17] To our knowledge, there are no studies in the literature on the size distribution of cities that discusses the welfare effect of an income redistribution policy.

5.1 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:

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

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

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

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

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

for all j.

5.2 Modified Fundamental Equation

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

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

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

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

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 we must use 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.[18] However, an advantageous property of this ODE is that it is very easy

[18] 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 labor-intensive good and thus the congestion costs are to check the 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 a boundary condition $\Phi(0) = 0$. 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.

5.3 Optimal Tax Rate

Since we have already obtained empirically relevant values of the parameters in the previous section, we use the parameter values listed in Table 2 as a benchmark. Then, 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 for each fixed tax rate $\tau \in [0,1]$, the equilibrium utility is calculated as[19]

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

where

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

increasing in t if ($\Phi'(t) > 0$ and) $\Phi''(t) > 0$ for all $t \in [0,1]$. Then, the argument in Subsection 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 skill-intensive 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.

[19] Here, we omit terms that are independent of τ.

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 the 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. In addition, $E_a(t)/(N(t)R(t)^\alpha)$ is equal to the total land-rent-adjusted after-tax income divided by the total population. To see this, let $\tilde{e}$ denote the constant such that

$$\frac{E_a(t)}{N(t)R(t)^\alpha} = \tilde{e} \quad \forall t.$$

Then, multiplying both sides by $N(t)$ and integrating the result over $t \in [0,1]$, we obtain

$$\tilde{e} = \frac{\int_0^1 E_a(t)/R(t)^\alpha dt}{\int_0^1 N(t)dt} = \int_0^1 \frac{E_a(t)}{R(t)^\alpha} dt,$$

where the second equality follows from the normalized population. Furthermore, using the above results for $E_a(t)$ and $R(t)$, we obtain

$$\tilde{e} = [1 - \alpha(1-\tau)]^{-(1-\alpha)} e^{-c_0},$$

where the first term represents the *income multiplier effect* under a non-zero tax rate τ, and the second seems to represent the effect of the economy-wide or average land rent. As a result, the equilibrium utility is calculated by

$$\ln \bar{U} = -(1-\alpha)\ln[1-\alpha(1-\tau)] - c_0 - (1-\alpha)\ln P. \tag{19}$$

Figure 3 then depicts the consequent relationship between tax rate τ and the natural logarithm of the equilibrium utility $\ln \bar{U}$. This figure clearly shows (i) that the laissez-faire outcome is Pareto-dominated by any non-zero tax rate; and (ii) that $\ln \bar{U}$ has a unique peak at a tax rate of approximately 5.77% with an utility level that is 1.53% higher than in the laissez-faire case. The former result is primarily due to excess concentration in the top cities as suggested by the population profile $N(t)$ depicted in the upper-right panel in Figure 5.[20] $N(t)/N(0)$ under laissez-faire

[20] This finding agrees with Henderson (1974) in that the market economy (without competitive

and that under the optimal tax are only significantly different at the upper tail of the distribution. The decomposition of utility, depicted in the left panel of Figure 4, does supports this interpretation. That is, as the tax rate τ converges to zero or as individuals concentrate into super-star cities, as shown in the upper-right panel of Figure 5, the term $-c_0$ decreases rapidly compared to the other two counteracting effects, i.e., $-(1-\alpha)\ln[1-\alpha(1-\tau)]$ and $-(1-\alpha)\ln P$ in (19).

The latter is due to the result that the positive welfare effect $-c_0$ of a more equalized population distribution through a decrease in the average land rent is significant at lower levels of τ and then becomes weaker for higher levels of τ. The other two effects are both decreasing in τ, and the negative slope of the income multiplier effect $-(1-\alpha)\ln[1-\alpha(1-\tau)]$ dominates the positive slope of $-c_0$ even after the slope of the price index effect $-(1-\alpha)\ln P$ becomes nearly flat, resulting in a single peak. Here, the result that $-(1-\alpha)\ln P$ is decreasing in τ implies that the price index $\ln P$ is increasing in τ. As suggested in the right panel of Figure 4 and Figure 5, this is because as τ increases and thus as individuals out-migrate from super-star cities, the numbers of varieties in those cities decreases significantly. As a result, even with the associated easing of congestion costs, i.e., decreases in land rents $R(t)$ and wage rates $W(t)$ in those cities as depicted in the middle panels in Figure 5, the price index $\ln P$ increases. Interestingly, given an increase in the tax rate τ from zero to the optimal level, the average establishment profile $\zeta(t)$ adjusts in a way that it weakens the positive effect of out-migration from top cities on $\ln P$, helping the overall welfare effect of an increase in τ to be positive. That is, a decreased establishment size in top cities weakens the negative effect of out-migration on the number of varieties, and a stable establishment size in small- and medium-sized cities enhances the increase in the numbers of varieties associated with in-migration, as depicted in the lower panels of Figure 5.

It is noteworthy that the perfect redistribution, i.e., $\tau = 1$, achieves a higher utility than we find under the laissez-faire case. The utility that is achieved under laissez-faire is 0.25% lower than that of the equilibrium with $\tau = 1$ (Figure 3).

developers) is characterized by excess agglomeration. In addition, this result is not inconsistent with that reported by Pflüger and Tabuchi (2010), who argue that with land use in both consumption and production, the spatial configuration is characterized by either efficient or excess agglomeration.

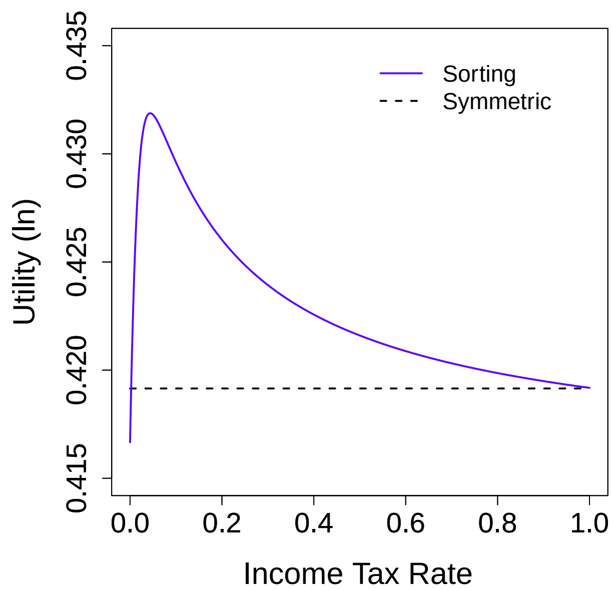

Figure 3: Income Tax τ and the Equilibrium Utility (ln)

Importantly, the equilibrium with $\tau = 1$ is *not* equivalent to the symmetric equilibrium. That is, it is a sorting equilibrium, as is shown in Figure 5. The result that population $N(t)$ and the average establishment size $\zeta(t)$ are decreasing t is a crucial qualitative difference from the other two cases: the optimal tax and laissez-faire cases. It is also noteworthy that the equilibrium with $\tau = 1$, although it is a sorting equilibrium, does not differ quantitatively from the symmetric equilibrium. The Lorenz curve $\Phi(t)$ (the green dot-dashed line) is fairly close to the 45-degree line. In addition, except for cities with smaller market size, i.e., cities with smaller t, the profiles of all variables are fairly flat, resulting in a good approximation of the symmetric equilibrium in which all variables are the same across cities. Thus, consistent with this observation, the two equilibria achieve the same utility level as shown in Figure 3.

Does this result imply that free migration (and trade) harms welfare?[21] Clearly,

[21] Given that the symmetric equilibrium is equivalent to autarky, this question can be replaced with the question, "Do free migration and trade Pareto-improve welfare?", which is the regional version of the question that is frequently asked in the trade literature.

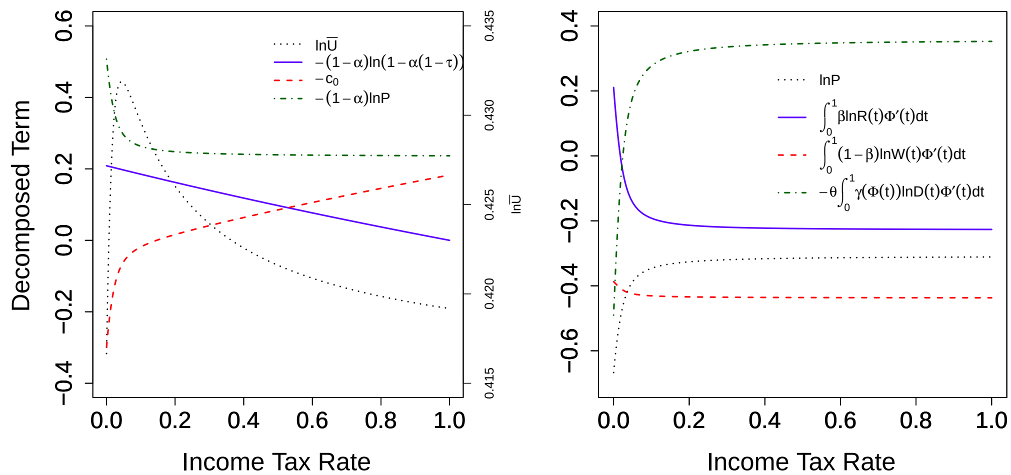

Figure 4: Decomposition of Utility $\ln \bar{U}$ (left panel) and Final Good Price $\ln P$ (right panel)

this is not the case. The policy implication of the above results is that only a slight difference in the tax rate, i.e., 0.00% vs. 5.77%, can result in completely different welfare consequences. If the government wants the laissez-faire economy to achieve greater utility, it suffices to employ any non-zero tax rate. In contrast, if we live in an economy that already features an income redistribution policy rule, we need to be careful about the policy in order to achieve greater utility. Importantly, the quantitative analysis validates the necessity of such a careful management of the policy because if we measure the welfare in terms of the lifetime utility, the utility level under the optimal tax rate is 32.0% and 38.3% larger than those under autarky and laissez-faire, respectively.[22]

[22] Here, the lifetime utility is computed as the utility of an infinitely-lived agent with discounted factor of 0.96. The resulting substantial change in welfare is contrasting with the result reported by Desmet and Rossi-Hansberg (2013), who use a neo-classical model and argue that the welfare impact of a change in the spatial distribution of economic activity is small for the United States.

6 Conclusion

What are the implications of task trade across cities for the size distribution of cities? In order to answer this question, this paper develops a spatial equilibrium model of a system of cities with task trade in which ex ante identical locations specialize in different sets of tasks in a symmetry-breaking manner. The city size is determined as a balance between agglomeration and dispersion forces, the former and latter of which arise from market interactions between varieties and from scarce land, respectively. The specialization of cities then results in disparities between cities, and the higher the skill intensity of a city is, the larger the city size, which favors urban diversity and thus agglomeration. We have shown that a sorting equilibrium exists and is unique, exhibiting comovement between income, population, the wage rate, the land rate, urban diversity, and the average establishment size in the skill-intensive sector. The necessary and sufficient condition for the size distribution to obey a power law including Zipf's law as a special case is also derived, and a quantitative analysis confirms that the model is consistent with the size distribution of U.S. cities. A welfare analysis is also conducted, suggesting that although the laissez-faire is characterized by excess agglomeration with a substantial welfare loss.

References

ACEMOGLU, D., "Directed Technical Change," *Review of Economic Studies* 69 (October 2002), 781–809.

AIYAGARI, S. R., "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics* 109 (August 1994), 659–84.

BECKMANN, M. J., "City Hierarchies and the Distribution of City Size," *Economic Development and Cultural Change* 6 (1958), 243–248.

BEHRENS, K., G. DURANTON AND F. ROBERT-NICOUD, "Productive cities: Sorting, selection and agglomeration," CEPR Discussion Papers 7922, C.E.P.R. Discussion Papers, July 2010.

Behrens, K., G. Mion, Y. Murata and J. Südekum, "Spatial Frictions," Working papers, IZA, January 2013.

Davis, J. C. and J. V. Henderson, "The agglomeration of headquarters," *Regional Science and Urban Economics* 38 (September 2008), 445–460.

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

Desmet, K. and E. Rossi-Hansberg, "Urban Accounting and Welfare," *American Economic Review* (2013), forthcoming.

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

Dixit, A. K. and J. E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67 (June 1977), 297–308.

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

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

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

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

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

Feenstra, R. C. and G. H. Hanson, "Foreign Investment, Outsourcing, and Relative Wages," in R. C. Feenstra, G. M. Grossman and D. A. Irwin, eds., *The*

Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati (Cambridge, MA: MIT Press, 1996), 89–127.

FELDMAN, M. P. AND D. B. AUDRETSCH, "Innovation in cities: Science-based diversity, specialization and localized competition," *European Economic Review* 43 (February 1999), 409–429.

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

GABAIX, X. AND Y. M. IOANNIDES, "The evolution of city size distributions," in J. V. Henderson and J. F. Thisse, eds., *Handbook of Regional and Urban Economics*volume 4 of *Handbook of Regional and Urban Economics*, chapter 53 (Elsevier, 2004), 2341–2378.

GROSSMAN, G. M. AND E. ROSSI-HANSBERG, "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review* 98 (December 2008), 1978–97.

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

HENDERSON, J. V. AND H. G. WANG, "Urbanization and city growth: The role of institutions," *Regional Science and Urban Economics* 37 (May 2007), 283–313.

HENDRICKS, L., "The Skill Composition Of U.S. Cities," *International Economic Review* 52 (02 2011), 1–32.

HSU, W., "Central Place Theory and City Size Distribution," *Economic Journal* 122 (09 2012), 903–932.

HUGGETT, M., "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control* 17 (1993), 953–969.

JACOBS, J., *The economy of cities*, Vintage international (Random House, 1969).

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

MATSUYAMA, K., “Endogenous Ranking and Equilibrium Lorenz Curve Across (ex ante) Identical Countries,” *Econometrica* (2013), forthcoming.

MELITZ, M. J., “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica* 71 (November 2003), 1695–1725.

MICHAELS, G., F. RAUCH AND S. J. REDDING, “Task Specialization in U.S. Cities from 1880-2000,” CEP Discussion Papers dp1186, Centre for Economic Performance, LSE, February 2013.

PFLÜGER, M. AND T. TABUCHI, “The size of regions with land use for production,” *Regional Science and Urban Economics* 40 (November 2010), 481–489.

ROSSI-HANSBERG, E. AND M. L. J. WRIGHT, “Urban Structure and Growth,” *Review of Economic Studies* 74 (04 2007), 597–624.

VALENTINYI, A. AND B. HERRENDORF, “Measuring Factor Income Shares at the Sector Level,” *Review of Economic Dynamics* 11 (October 2008), 820–835.

YI, K.-M., “Can Vertical Specialization Explain the Growth of World Trade?,” *Journal of Political Economy* 111 (February 2003), 52–102.

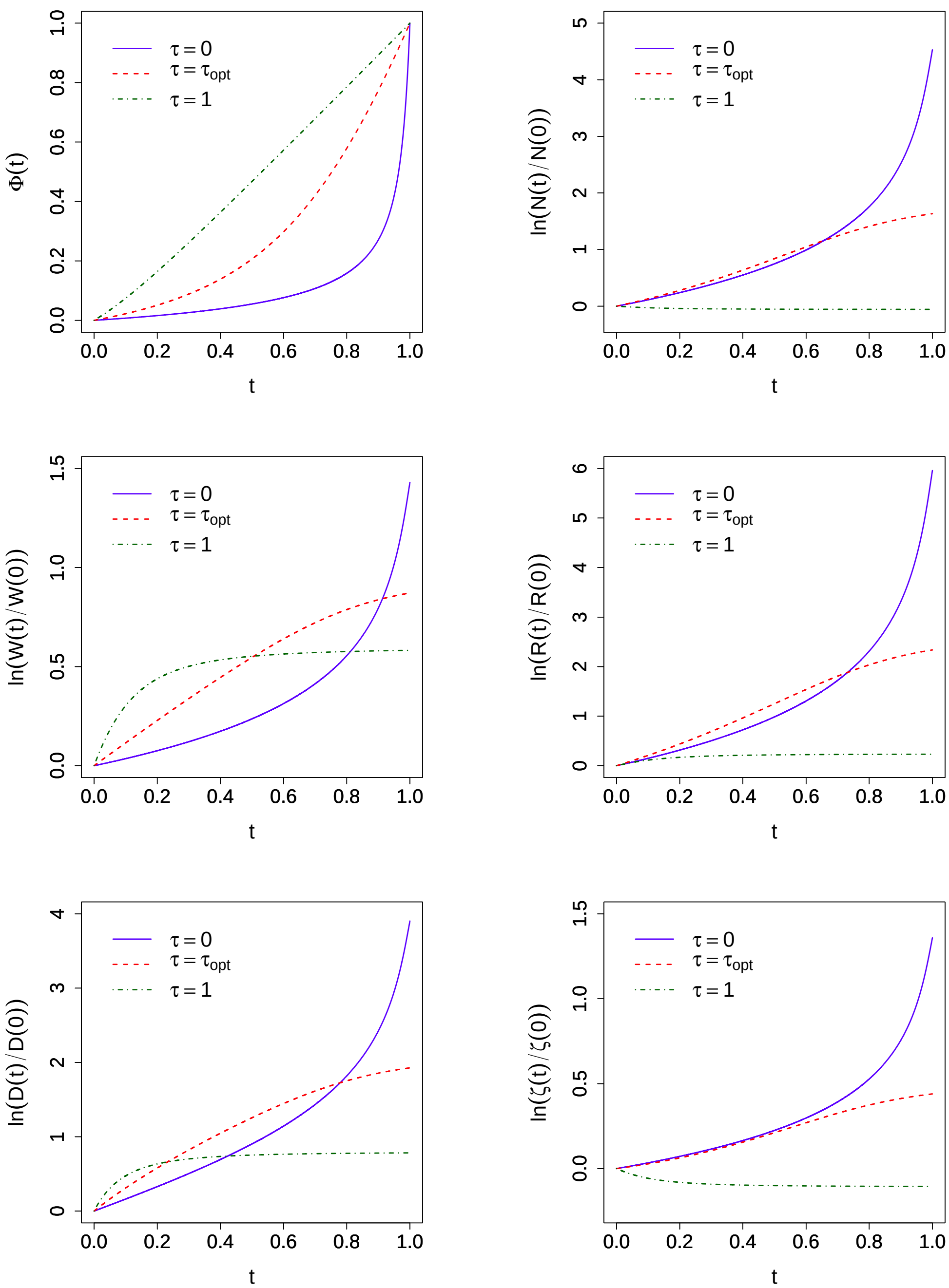

Figure 5: Lorenz Curve $\Phi(t)$ and Profiles of $N(t)$, $W(t)$, $R(t)$, $D(t)$, $\zeta(t)$