# Cities' Demand-driven Industrial Composition*

Takashi Onoda†

*Japan Bank for International Cooperation and*

*TDB-CAREE, Graduate School of Economics, Hitotsubashi University*

April 26, 2023

**Abstract**

I report a new stylized fact: large cities specialize in income-elastic sectors. I explain this by developing a model that has heterogeneous income-elasticities, heterogenous tradabilities, and mobile agents. Heterogeneous fundamental productivities generate the specialization pattern through the home market effect. The city with fundamentally better productivity becomes larger, offers a higher wage, and specialize in income-elastic sectors. Also, the model reveals that the aggregate share of tradable sectors is an important factor when we study cross-location income inequality.

**Keywords:** Home market effect, Non-homothetic preference, Cities, Industrial composition, Comparative advantage **JEL classification:** F12, F14, R12, R13, R23

*This paper is based on a chapter of my dissertation at the University of Chicago. I am grateful to my advisors, Jonathan Dingel, Rodrigo Adão, and Felix Tintelnot for their invaluable guidance. I also thank Nancy Stokey, Thomas Winberry, Kensuke Teshima, Yoichi Sugita, Kentaro Nakajima, Eiichi Tomiura, Shota Fujishima, Dao-Zhi Zeng, Tatsuhito Kono, and seminar participants in Hitotsubashi Trade/Urban Economics Workshop, Regional Science Workshop at Tohoku University, and the third-year macro research seminar and International Trade Working Group at the University of Chicago.

†Email: onoda.takashi@gmail.com

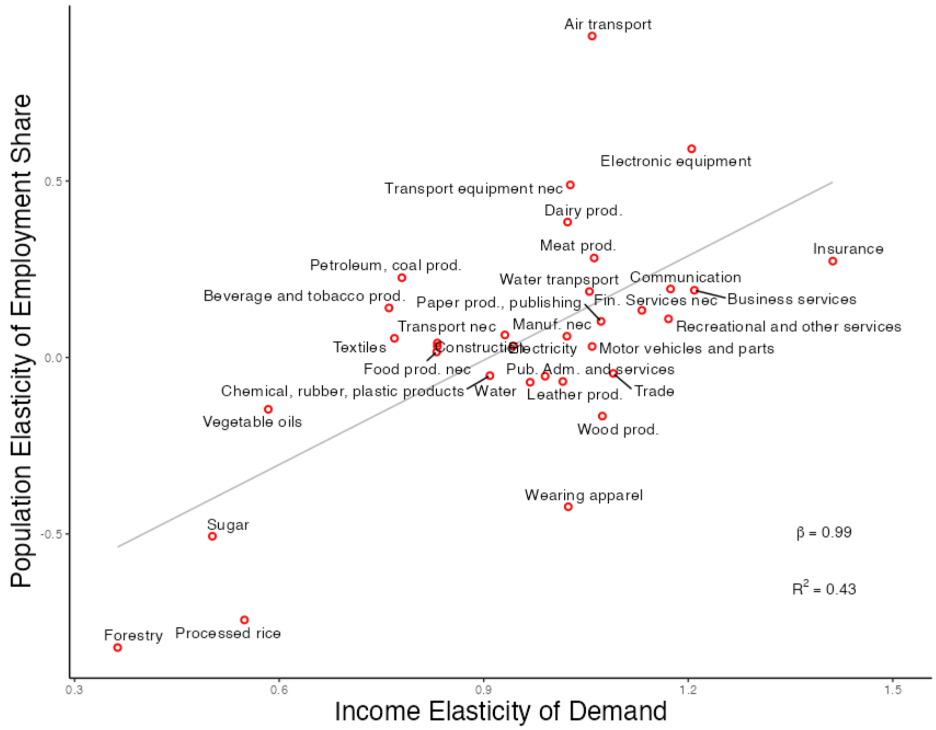
# 1 Introduction

Industrial composition varies greatly across cities: Detroit, for example, is synonymous with cars and Silicon Valley with computers. Figure 1 shows that much of the variation in employment composition across U.S. cities is linked to the income-elasticity of demand for that industries' output. Employment shares of industries with high income-elasticities of demand, such as air transport services and insurance, are higher in large Metropolitan Statistical Areas (MSAs) like New York-Newark-Jersey City Metropolitan Area. Industries with low income-elasticities of demand, such as sugar manufacturing and processed rice manufacturing, have larger employment shares in small MSAs. This pattern is not driven by the difference between food-related and non-food-related sectors. Food-related industries with high income-elasticities such as dairy products and meat products nec have relatively high employment shares in large cities.

This paper investigates this new stylized fact and the implications. I develop a model that can generate this specialization pattern as equilibrium an outcome. My model is based on Matsuyama (2019), who theoretically studies international trade patterns with heterogeneous sectoral income-elasticities and differentiated goods within a sector. I extend Matsuyama (2019) by introducing worker mobility, land consumption, and non-tradable sectors (e.g., retail, consumption amenities). In an equilibrium, a fundamental productivity difference between cities generates asymmetric population allocation and the home market effect on the wage rate and on the trade pattern, which is consistent with the specialization patterns in Figure 1. In addition to explaining the specialization pattern, the model implies that the aggregate share of tradable sectors, which is endogenously determined in my model, is related to cross-location wage patterns.

In the production patterns explained as equilibrium outcomes of my model, the home market effect plays a key role. In the home market approach first formally theorized by Krugman (1980), the effect is of two types, each of which shares the mechanism from trade costs and an increasing return to scale production. The first affects the wage rate. Other things being equal, the wage rate tends to be higher in larger markets. When firms are exposed to competition with firms in

Figure 1: Elasticity of Employment Share with respect to MSA's Population and Elasticity of Demand with respect to Income




MSA population are 2016 data from the Bureau of Economic Analysis. Employment shares are calculated from 2016 County Business Pattern data. NAICS data are mapped to GTAP sectors. When obtaining the population elasticity of employment share, region $\in \{Northeast, Midwest, South, West\}$ is controlled. Income-elasticity estimates are from Caron et al. (2020). The unweighted regression line is shown. For details of data description, see Appendix A.

other locations and sharing demands, differences in access to markets due to trade costs drive the difference in the input costs so that the firms' profits are equalized at zero in any location. The second effect affects trade pattern by generating comparative advantage. When the relative market size of sectors varies across regions, regions export goods for which they have relatively large domestic markets. This is because, in the presence of trade costs, local firms are incentivized to operate in a sector that has a relatively larger home market, and this incentive is strong enough to amplify the demand pattern to the production pattern.

In my model, a difference in cities' fundamental productivities generates these two home market effects, and this eventually produces the specialization pattern, which is consistent with Figure 1. First, a city with better fundamental productivity attracts workers, which results in a large population, and this generates the home market effect on the wage rate. Second, due to the higher wage, residents in the large city spend relatively more on income-elastic sectors and this generates the home market effect on the trade pattern. Hence, in the equilibrium, the fundamentally productive city becomes larger, offers a higher wage, and specializes in income-elastic sectors.

In addition to explaining the production pattern, the model highlights the importance of the aggregate tradable sector share when we study cross-location income inequality. In equilibrium, the wage increases not only with city size but also with the aggregate share of tradable sectors. This is because the home market effect works only through tradable sectors. The source of the home market effect is competition between firms in different locations. When a sector is non-tradable, that competition does not exist, and the home market effect does not emerge. Thus, the market size that drives the wage is the size of tradable sectors, which is the product of the overall market size and the aggregate tradable sector share. In my model, the aggregate tradable sector share of a city is endogenously determined by two factors: the prices and the income elasticities of the tradable sectors relative to the non-tradable sectors. The prices of the tradable sectors are, other things being equal, relatively expensive in the large city, which makes the aggregate tradable sector share greater (smaller) when the sectors are gross complements (substitutes). The aggregate tradable sector share of the large city becomes greater (smaller) also when the tradable sectors are relatively income-elastic (income-inelastic). These results provides us with a new perspective; we can analyze how a change in economic environment affects cross-location income patterns by changing the aggregate tradable sector shares. In Onoda (2022), I theoretically and empirically demonstrate that the model can be used to understand the evolution of the income level across cities after 1980.

Industrial composition is important from several perspectives. First, it is an important factor for local economies, and not only because an aggregate tradable sector share is a driver of local wages through the home market effect; it also is important because local economic performance is significantly affected by the industries that locate in a city (e.g., Autor et al. (2013)). To fully understand variation in local economic performance it is necessary, first, to understand what determines the mix of industry in cities. Second, it is related to higher returns to experience in big cities. Eckert et al. (2022) report that faster wage growth in big cities is explained substantially by workers sorting into the industries typically found in big cities. This suggests that understanding the determinants of industrial composition is a critical step in studying productivity growth in cities. Third, the mechanism that drives industrial composition also is important for researchers who want to exploit regional variation in the size of industries. In this paper, I report that industrial composition is related to city size. Given this relationship, regressing dependent variables on sectoral sizes or shares, while not controlling for the city size of examined locations, might lead to an omitted variable bias problem because the trend of a dependent variable could be driven by the city size of the locations rather than by sectoral sizes or shares. Understanding the mechanism can help researchers avoid this endogeneity issue.

This paper is the first research on cross-city inter-sectoral specialization patterns to be undertaken from the demand-side perspective. Most previous works on the cross-city specialization pattern has focus on non-demand side factors such as functions in production, (Duranton and Puga (2005) and Henderson and Ono (2008)), the strength of the agglomeration economy (Behrens and Robert-Nicoud (2015)), and the skill supply (Davis and Dingel (2019)). A few works focus on the demand side's effect on cross-city difference (e.g., Handbury (2019)). The closest of these is Dingel (2017), who quantifies the contributions made to the quality specialization of cities by the heterogeneous demand factor and by the skill supply factor, and he creates a model to guide the quantification. There are two major differences between his model and mine. First, we focus on different types of specialization and trade patterns. His work, and, therefore, his model, examines vertical specialization and the intra-sector trade, where different quality goods are gross substitutes. Whereas, my model examines horizontal specialization and inter-sectoral trade, where goods in different sectors are either gross complements or gross substitutes. Second, my model has mobile agents and land consumption, which his model does not. These assumptions fit the urban economy environment, and, moreover, it enables us to study the relationship between the size of a city and its industrial composition. This paper, the first work on cross-city inter-sectoral specialization patterns undertaken from the demand-side perspective, contributes to the literature by reporting the stylized fact and providing the theoretical framework for urban economies.

Furthermore, my work contributes more broadly to studies of the demand-side effect in trade patterns. The international trade literature shows theoretically that the demand-side effect plays an important role in determining trade patterns and specializations (e.g., Flam and Helpman (1987), Stokey (1991), Matsuyama (2019), Fajgelbaum et al. (2011)). The closest to mine is Matsuyama (2019) in the sense he analyzes horizontal specialization and inter-sectoral trade. My model extends Matsuyama (2019) by introducing non-tradable sectors, land consumption, and mobile agents, and it reveals a new effect that heterogeneous demand across locations generates. This is the effect on cross-location income through endogenous tradable sector shares, and it is relevant both in international and regional income inequalities.

The rest of this paper is organized as follows. Section 2 develops the model to explain the stylized fact. Section 3 provides robustness checks, and Section 4 concludes.

# 2 Model

In this section, I introduce the model of two cities ($i \in \{1, 2\}$). The cities fundamentally differ in productivity and amenity. There is a mass of $N$ workers who are freely mobile and homogeneous except for inherent taste over cities. Conditional on location, individual labor supply is inelastic. There are $K$ goods-producing sectors, and the sectors differ in relative income-elasticity ($\epsilon_k$) and tradability ($\tau_k$) as well as productivities ($\phi_k, \psi_k$) and preference shifters ($\beta_k$). In a city, land ($L_i$) is inelastically supplied and consumed by households. I start with explaining the household problem.

### Household

The problem for a worker $\theta$ is given by

$$\max_{i\in\{1,2\},U_i,C_i,H_i\{q_{ik}(\nu)\}_{\nu\in\Omega_{ik},k\in\mathcal{K}},\{Q_{ik}\}_{k\in\mathcal{K}}} U_i\cdot a_i\cdot\delta(\theta,i) \tag{1}$$

$$\begin{aligned}
s.t.\ U_i &= C_i^{1-\alpha}H_i^{\alpha}\\
C_i &= \left[\sum_{k\in\mathcal{K}}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{ik}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}\\
Q_{ik} &= \left[\int_{\Omega_{ik}}q_{ik}(\nu)^{\frac{\sigma-1}{\sigma}}d\nu\right]^{\frac{\sigma}{\sigma-1}}\\
E_i &= \sum_{k\in\mathcal{K}}\int_{\Omega_{ik}}p_{ik}(\nu)q_{ik}(\nu)d\nu+P_{iH}H_i
\end{aligned}$$

where $i$ is the city to reside in, $q_{ik}(\nu)$ is the consumption of variety $\nu$ in sector $k$, $\Omega_{ik}$ is the set of available varieties of sector $k$ in city $c$, $\mathcal{K}$ is the set of sectors, $Q_{ik}$ is the composite consumption of sector $k$, $\beta_k$ is the preference shifter of sector $k$, $E_i$ is the income in city $i$, $P_{iH}$ is the land rent in city $i$, $a_i>0$ for $i\in\{1,2\}$, $0<\eta<\sigma$, $\eta\neq1$, $1+(1-\alpha)\min_{k\in\mathcal{K}}\{\epsilon_k\}/(1-\eta)>0$, and $\sigma>1$. The utility consists of three factors: $U_c$, $a_c$, and $\delta(i,c)$.

The real consumption, $U_i$, is defined over goods consumption ($C_i$) and land consumption ($H_i$) in the form of a Cobb-Douglass utility function. The functional form of goods consumption captures the non-homothetic preference of the consumer. When $\epsilon_k=0$ for all $k\in\mathcal{K}$, this $C_i$ becomes a standard homothetic CES function. When $\epsilon_k$ varies across sectors, a higher $\epsilon_k$ corresponds to a more income-elastic sector. This is because the weight, $\beta_k^{1/\eta}U_i^{\epsilon_k/\eta}$, grows relatively more when the real consumption grows. This functional form follows Hoelzlein (2019), who incorporates housing consumption into a non-homothetic utility function that is used by Comin et al. (2021), Matsuyama (2019) and others. $1+(1-\alpha)\min_{k\in\mathcal{K}}\{\epsilon_k\}/(1-\eta)>0$ is imposed to ensure global monotonicity of $C_i$. When solving this household problem, a convenient property of the functional form of $C_i$ can be seen. The demand function derived from this preference becomes

$$Q_{ik} = \beta_l U_i^{\frac{\epsilon_k}{\eta}} P_{ik}^{-\eta} C_i^{1-\eta} E_{iC}^{\eta} \tag{2}$$

where $E_{iC}$ and $P_{ik}$ are the expenditure on goods consumption and the price index for sector $k$, respectively, in city $i$ and $P_{ik}$ is defined as $P_{ik} = \left[\int_{\nu \in \Omega_{ik}} p_{ik}(\nu)^{1-\sigma} d\nu\right]^{1/(1-\sigma)}$. This shows that the relative income-elasticity of demand ($\epsilon_k$) and price elasticity ($\eta$) are separated. I assume the price elasticity is common among sectors, and it is lower in the sector level than in the variety level within a sector ($\eta < \sigma$). Equation (2) implies the sectors are gross complements when $\eta < 1$ and gross substitutes when $\eta > 1$. Also, the expenditure share of sector $k$ is obtained as follows (see Appendix B for the derivation):

$$m_{ik} \equiv \frac{P_{ik} Q_{ik}}{\sum_{\ell \in \mathcal{K}} P_{i\ell} Q_{i\ell}} = \frac{\beta_k P_{ik}^{1-\eta} U_i^{\epsilon_k}}{\sum_{\ell \in \mathcal{K}} \beta_\ell P_{i\ell}^{1-\eta} U_i^{\epsilon_l}} \tag{3}$$

This is log-supermodular in $U_i$ and$\epsilon_k$, and it shows that, holding the price indices $\{P_k\}_{k \in \mathcal{K}}$ constant, agents with higher real consumption ($U_i$) spend relatively more on sectors with high $\epsilon_k$.[1] Another result is the price index for goods consumption ($P_i \equiv E_{iC}/C_i$), which is expressed as

$$P_{iC} = \left(\sum_{k \in \mathcal{K}} \beta_k U_i^{\epsilon_k} P_{ik}^{1-\eta}\right)^{\frac{1}{1-\eta}} \tag{4}$$

(see Appendix B for the derivation). This illustrates that as the real consumption $U_i$ rises, agents care particularly about the prices of high $\epsilon_k$ goods, on which they spend relatively more.

A worker consumes land for her residential use. $L_i$ unit of land is inelastically supplied in city $i$, and I assume $L_1 = L_2 = 1$. This is without loss of generality because changing $L_i$ for city $i$ is isomorphic to changing the utility level from an amenity $a_i$ for city $i$. I assume that the collected land rents in a city are redistributed to the residents, which implies the income of a worker is proportional to her wage in an equilibrium .

$a_i$ is the utility from an amenity offered by city $i$, such as weather, landscape, and historic

[1] $f(x, y)$ is log-supermodular in $x$ and $y$ iff $\partial^2 \log f(x, y)/\partial x \partial y > 0$.

heritage. While "amenity" generally refers to access to local services and consumer goods (e.g., restaurants), which is called consumption amenities, in this model, those local services and goods contribute to $C_i$ when they are consumed.

$\delta(\theta, i)$ is the idiosyncratic utility shock for the worker $\theta$ and city $i$ pair. This generates the heterogeneous taste over cities (following Tabuchi and Thisse (2002), Redding (2016), and others). I assume $\delta(\theta, i)$ is distributed i.i.d. across workers and cities according to the Fréchet distribution with shape parameter $1/\gamma$ ($Pr[\delta < x] = e^{-x^{-1/\gamma}}$). Each worker chooses the city that offers the higher utility, taking into account her consumption optimization. Thus, given the products of the real consumption and the utility from an amenity in the two cities, $a_1 U_1$ and $a_2 U_2$, the probability of choosing city 1 for a given agent is derived as $Pr[a_1 U_1 \delta(i) > a_2 U_2 \delta(i)] = (a_1 U_1)^{1/\gamma} / \left\{ (a_1 U_1)^{1/\gamma} + (a_2 U_2)^{1/\gamma} \right\}$. Since the shock is i.i.d., the cities' population ratio follows.

$$\frac{N_1}{N_2} = \left( \frac{a_1 U_1}{a_2 U_2} \right)^{1/\gamma} \tag{5}$$

Given the ratio of the product of the real consumption and the utility from an amenity ($a_1 U_1 / a_2 U_2$), the lower $\gamma$ is, the greater the population inequality is. This illustrates that the $\gamma$ measures the dispersion force in this economy.

### Production

The production in my model is based on Krugman (1980). For all sectors $k \in \mathcal{K}$ there are endogenous sets of varieties, homogeneous firms, and monopolistic competition. Each sector is either tradable with iceberg trade cost $\tau > 1$ or it is non-tradable. Let $\mathcal{T}$ be the set of tradable sectors and $\mathcal{N}$ be the set of non-tradable sectors ($\mathcal{K} = \mathcal{T} \cup \mathcal{N}$ and $\mathcal{T} \cap \mathcal{N} = \emptyset$). Each worker chooses the location of labor they supply. Conditional on location, individual labor supply is inelastic. I let $w_i$ denote the wage in city $i$. Each firm in sector $k$ in city $i$ needs to employ $\frac{\phi_k}{\lambda_i}$ units of labor as the fixed cost and $\frac{\psi_k}{\lambda_i}$ as the variable cost to produce a unit of variety. To let city 1 be fundamentally more productive than city 2, I assume that $\lambda = \lambda_1 > \lambda_2 = 1$. The problem for a firm that produces variety $\nu$ in sector $k$ in city $i$ is

$$\pi_{ik}(\nu) = \max_{p_{iik}(\nu), q_{iik}(\nu), p_{ijk}(\nu), q_{ijk}(\nu)} [p_{iik}(\nu) q_{iik}(\nu) - \frac{\psi_k}{\lambda_i} q_{iik}(\nu) w_i]$$
$$+ \mathbb{1}\{k \in \mathcal{T}\} \left[ p_{ijk}(\nu) q_{ijk}(\nu) - \tau \frac{\psi_k}{\lambda_i} q_{ijk}(\nu) w_i \right] - \frac{\phi_k}{\lambda_i} w_i \quad (6)$$
$$s.t.\ q_{ii}(\nu) = p_{iik}(\nu)^{-\sigma} P_{ik}^{\sigma} Q_{ik}$$
$$q_{ijk}(\nu) = p_{ijk}(\nu)^{-\sigma} P_{jk}^{\sigma} Q_{jk}$$

where $\pi_{ik}(\nu)$ is the profit by optimized production; $(i, j) \in \{(1, 2), (2, 1)\}$; $p_{iik}(\nu)$ and $p_{ijk}(\nu)$ are the prices for the markets in city $i$ and $j$, respectively; and $q_{iik}(\nu)$ and $q_{ijk}(\nu)$ are the quantities for the markets in city $i$ and $j$, respectively. In the following part, I omit $\nu$ unless it is confusing. The terms in the first bracket are the variable profits from selling products in city $i$, while those in the second are those in city $j$, which is zero if $k \in \mathcal{N}$. If sector $k$ in city $i$ has non-zero production in an equilibrium, $\pi_{ik}$ must be zero such that there is no entrant. Similarly, if sector $k$ in city $i$ has zero production in an equilibrium, $\pi_{ik}$ must be non-positive. This is the zero-profit condition in sector $k$ in city $i$.

### Definition of Competitive Equilibrium

A competitive equilibrium is $\{N_i, U_i, C_i, H_i, w_i, E_i, P_{iH}\}_{i \in \{1,2\}}$, $\{p_{ijk}, q_{ijk}\}_{(i,j,k) \in \{1,2\}^2 \times \mathcal{K}}$, and $\{\Omega_{ik}\}_{(i,k) \in \{1,2\} \times \mathcal{K}}$ such that

1. workers optimize consumption and locational choice as eq. (1) for $i \in \{1, 2\}$,
2. workers' income is given by $E_i = w_i + (N_i P_{iH} H_i)/N_i$ for $i \in \{1, 2\}$,
3. land clearing condition holds such that $N_i H_i = L_i (= 1)$ for $i \in \{1, 2\}$,
4. producers optimize production as eq. (6) for all $k \in \mathcal{K}$ and $i \in \{1, 2\}$,

5. the zero-profit condition holds such that $\pi_{ik} \leq 0$ for all $k \in \mathcal{K}$ and $i \in \{1, 2\}$ where equality holds if $q_{iik} + \tau q_{ijk} > 0$,

6. the national labor market clearing condition that $N_1 + N_2 = N$ holds, and

7. the local labor market clearing conditions,

$$\sum_{k \in \mathcal{N}} \int_{\Omega_{1k}} \left( \frac{\psi_k}{\lambda_1} q_{11k} + \frac{\phi_k}{\lambda_1} \right) d\nu + \sum_{k \in \mathcal{T}} \int_{\Omega_{1k}} \left( \frac{\psi_k}{\lambda_1} q_{11k} + \tau \frac{\psi_k}{\lambda_1} q_{12k} + \frac{\phi_k}{\lambda_1} \right) d\nu = N_1$$

and

$$\sum_{k \in \mathcal{N}} \int_{\Omega_{2k}} \left( \frac{\psi_k}{\lambda_2} q_{22k} + \frac{\phi_k}{\lambda_2} \right) d\nu + \sum_{k \in \mathcal{T}} \int_{\Omega_{2k}} \left( \frac{\psi_k}{\lambda_2} q_{22k} + \tau \frac{\psi_k}{\lambda_2} q_{21k} + \frac{\phi_k}{\lambda_2} \right) d\nu = N_2,$$

hold.

### Equilibrium Conditions

I now characterize an equilibrium by obtaining simplified conditions. I focus on on equilibria where all sectors have nonzero output in both cities, and, given the optimized production and the demand function, I impose the zero-profit condition on each sector in each city. The first result concerns non-tradable sectors.

**Proposition 1.** *Given an equilibrium, the expenditure share and the employment share of a non-tradable sector in a city are equalized.*

$$\forall i \in \{1, 2\},\ \forall k \in \mathcal{N} \quad m_{ik} = x_{ik} \tag{7}$$

where $x_{ik}$ is the employment share of sector $k$ in city $i$.

*Proof.* see Appendix D . □

This equalization follows from the zero-profit condition in the sector level. That is, each firm has zero profit and therefore the sector has zero aggregate profit. Since the revenues in the non-tradable sectors are only received from the agents in the same city and the factor payments are made only to the same agents, the zero-profit condition boils down to equation (7). Corollary 1 follows.

**Corollary 1.** *Given an equilibrium, the aggregate expenditure share and the aggregate employment share in a city are equalized both for the non-tradable sectors and for the tradable sectors.*

$$\forall i \in \{1,2\}, \sum_{k\in\mathcal{N}} m_{ik} = \sum_{k\in\mathcal{N}} x_{ik} \text{ and } \sum_{k\in\mathcal{T}} m_{ik} = \sum_{k\in\mathcal{T}} x_{ik}$$

*Proof.* $\sum_{k\in\mathcal{N}} m_{ck} = \sum_{k\in\mathcal{N}} x_{ck}$ follows from equation (7). $\sum_{k\in\mathcal{T}} m_{ck} = 1 - \sum_{k\in\mathcal{N}} m_{ck} = 1 - \sum_{k\in\mathcal{N}} x_{ck} = \sum_{k\in\mathcal{T}} x_{ck}$. □

The equalization for the tradable sectors reflects the fact that a city's trade balance must be zero in an equilibrium. This equalization between the aggregate expenditure share and the aggregate employment share in the tradable sectors is a key for the tractability of the model. Another result is Proposition 2.

**Proposition 2** (Home Market Effect on Wage)**.** *Given an equilibrium, the city sizes ($N_1$ and $N_2$), the aggregate tradable sector shares ($m_{1\mathcal{T}} \equiv \sum_{k\in\mathcal{T}} m_{1k}$, and $m_{2\mathcal{T}} \equiv \sum_{k\in\mathcal{T}} m_{2k}$), and the relative wage ($\omega = w_1/w_2$) satisfy the following equation.*

$$\frac{m_{1\mathcal{T}}}{m_{2\mathcal{T}}}\frac{N_1}{N_2} = \omega^{2\sigma-1}\left[\frac{\lambda^{-\sigma} - \rho\omega^{-\sigma}}{\lambda^{\sigma} - \rho\omega^{\sigma}}\right] \tag{8}$$

*Proof.* see Appendix D . □

This result is obtained by aggregating the zero-profit condition over tradable sectors and making use of $\sum_{k\in\mathcal{T}} m_{ik} = \sum_{k\in\mathcal{T}} x_{ik}$. This summarizes the sector-level zero-profit condition for the tradable sectors into the city level. The LHS is the relative tradable market size of city 1 to city 2. The RHS, which increases in the relative wage of city 1 ($\omega$), shows two things. First, the sector-level force--that a large local market is accompanied by a higher input cost--is carried over to the city level. This corresponds to the home market effect on the wage rate discussed in Krugman (1980) and Krugman (1991). Second, the relative aggregate expenditure share of the tradable sectors matters. For the zero-profit condition to hold in each sector-city pair, the larger city's advantage of lower trade costs for selling to consumers must be exactly offset by higher input costs. This force does not appear in non-tradable sectors because the local wage rate change affects both the demand and the input cost at the same rate.[2] Next, the expenditure shares in an equilibrium are obtained.

**Lemma 1.** *Given an equilibrium, the expenditure shares of sector $k \in \mathcal{K}$ in city 1 and city 2 can be expressed as follows:*

$$m_{1k} = \left[\frac{\lambda^{-\sigma} - \rho_k \omega^{-\sigma}}{(1-\rho_k^2)N_1}\right]^{\frac{1-\eta}{\sigma-\eta}} \left(\tilde{\beta}_k U_1^{\epsilon_k}\right)^{\mu} \left[U_1^{\frac{1}{1-\alpha}} N_1^{\frac{\alpha}{1-\alpha}}\right]^{(1-\eta)\mu} \tag{9}$$

$$m_{2k} = \left[\frac{1 - \rho_k \lambda^{-\sigma} \omega^{\sigma}}{(1-\rho_k^2)N_2}\right]^{\frac{1-\eta}{\sigma-\eta}} \left(\tilde{\beta}_k U_2^{\epsilon_k}\right)^{\mu} \left[U_2^{\frac{1}{1-\alpha}} N_2^{\frac{\alpha}{1-\alpha}}\right]^{(1-\eta)\mu} \tag{10}$$

*where*

$$\mu = \frac{\sigma-1}{\sigma-\eta} > 0, \;\; \rho_k = \begin{cases} 0 & k \in \mathcal{N} \\ \rho = \tau^{1-\sigma} & k \in \mathcal{T} \end{cases}, \;\; \textit{and} \;\; \tilde{\beta}_k = \beta_k \left(\frac{\phi_k^{\frac{1}{\sigma-1}} \psi_k}{\sigma^{\frac{\sigma}{\sigma-1}}(\sigma-1)}\right)^{1-\eta}$$

*Proof.* see Appendix D . □

Making use of Lemma 1 and $\sum_{k\in\mathcal{K}} m_{ik} = 1$, the real consumption in an equilibrium is obtained as Proposition 3.

[2]In Onoda (2022), I empirically verify that the local wage is positively associated with the tradable sector share.

**Proposition 3.** *Given an equilibrium, the real consumptions in city 1 and city 2 ($U_1$ and $U_2$) can be implicitly expressed as follows:*

$$U_1 = \left\{ \lambda^{-\sigma \frac{1-\eta}{\sigma-\eta}} \sum_{k \in \mathcal{N}} \left( \tilde{\beta}_k U_i^{\epsilon_k} \right)^{\mu} + \left[ \frac{\lambda^{-\sigma} - \rho \omega^{-\sigma}}{(1-\rho^2)} \right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k \in \mathcal{T}} \left( \tilde{\beta}_k U_i^{\epsilon_k} \right)^{\mu} \right\}^{-\frac{(1-\alpha)}{(1-\eta)\mu}} N_1^{\frac{1-\alpha}{\sigma-1} - \alpha} \tag{11}$$

$$U_2 = \left\{ \sum_{k \in \mathcal{N}} \left( \tilde{\beta}_k U_1^{\epsilon_k} \right)^{\mu} + \left[ \frac{1 - \rho \lambda^{-\sigma} \omega^{\sigma}}{(1-\rho^2)} \right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k \in \mathcal{T}} \left( \tilde{\beta}_k U_2^{\epsilon_k} \right)^{\mu} \right\}^{-\frac{(1-\alpha)}{(1-\eta)\mu}} N_2^{\frac{1-\alpha}{\sigma-1} - \alpha} \tag{12}$$

*Proof.* It follows from $1 = \sum_{k \in \mathcal{K}} m_{1k}$ and Lemma 1. □

Holding the relative wage constant, the elasticity of the real consumption with respect to the local population is given by

$$\frac{\partial U_i / U_i}{\partial N_i / N_i} = \frac{(1-\alpha)/(\sigma-1) - \alpha}{1 + (1-\alpha)\tilde{\epsilon}_k / (1-\eta)}$$

where $\tilde{\epsilon}_c = \sum_{k \in \mathcal{K}} m_{ik} \epsilon_k$. $1/(\sigma-1)$ in the numerator is the elasticity of the agglomeration economy or the positive externality in Krugman-type models with homothetic preference. The number of varieties in a location increases with market size and consumers have love-of-variety in their preferences. This elasticity is multiplied by $1-\alpha$ because it is the share of goods consumption. The negative externality $\alpha$ is caused by the inelastic supply of land. As local population grows, the size of land that a resident can consume becomes smaller, which appears as a negative externality. The externality has an additional term, $(1-\alpha)(1-\eta)/\tilde{\epsilon}_c$. To understand this, suppose all sectors have homogeneous income-elasticity, $\bar{\epsilon}$ and increase $Q_{ik}$ for all $k \in \mathcal{K}$ by the same proportion ($dQ_{ik}/Q_{ik} = dQ_{i\ell}/Q_{i\ell}$ for all $(i, \ell) \in \mathcal{K}^2$) . The elasticity of the real consumption is given by:

$$\frac{dU_i}{U_i} / \sum_{k \in \mathcal{K}} \frac{dQ_{ik}}{Q_{ik}} = \left[ \frac{1}{(1-\alpha)} + \frac{\bar{\epsilon}}{1-\eta} \right]^{-1}$$

This shows that, the greater $\bar{\epsilon}$ is, the smaller the elasticity of the real consumption is. This relationship between the elasticity of the real consumption and $\epsilon_k$ is carried over to the heterogeneous income-elasticity model. The effect on the marginal utility from aggregate consumption is summarized by the average of $\epsilon_k$ weighted with the expenditure shares.

The negative relationship between the relative wage ($\omega = w_1/w_2$) and the real consumption in city 1 ($U_1$) is related to the home market effect on the wage. Given $w_1$, a higher $w_2$ implies greater demand for tradable sectors in city 2 according to the home market effect on the wage equation (8). The greater demand is accompanied by richer varieties produced in city 2 in the tradable sectors because the trade balance must be zero. Finally, the richer varieties from city 2 reduces the price index for goods consumption in city 1 and raises the real consumption. This is the reason why $\omega$ appears only in the terms for the tradable sectors in equations (11) and (12).

Finally, an equilibrium is characterized by the seven conditions obtained above: equations (5), (8), (11), (12), the national labor clearing condition ($N_1 + N_2 = N$), and the aggregate tradable sector shares as follows:

$$m_{1\mathcal{T}} = \left[\frac{\lambda^{-\sigma} - \rho\omega^{-\sigma}}{(1-\rho^2)N_1}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \left(\tilde{\beta}_k U_1^{\epsilon_k}\right)^{\mu} \left[U_1^{\frac{1}{1-\alpha}} N_1^{\frac{\alpha}{1-\alpha}}\right]^{\frac{(1-\alpha)}{(1-\eta)\mu}} \tag{13}$$

$$m_{2\mathcal{T}} = \left[\frac{1 - \rho\lambda^{-\sigma}\omega^{\sigma}}{(1-\rho^2)N_2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \left(\tilde{\beta}_k U_2^{\epsilon_k}\right)^{\mu} \left[U_2^{\frac{1}{1-\alpha}} N_2^{\frac{\alpha}{1-\alpha}}\right]^{\frac{(1-\alpha)}{(1-\eta)\mu}} \tag{14}$$

where $m_{1\mathcal{T}} = \sum_{k\in\mathcal{T}} m_{1k}$, and $m_{2\mathcal{T}} = \sum_{k\in\mathcal{T}} m_{2k}$. These aggregate shares follow from individual tradable sector's share (equations (9) and (10)). There are seven unknown variables $\{U_1, U_2, \omega, N_1, N_2, m_{1\mathcal{T}}, m_{2\mathcal{T}}\}$ for these seven equations. By making use of the zero-profit conditions for all city-sector pairs, the number of endogenous variables here is fewer than that of the equilibrium definition. Based on these conditions, I now provide some properties of the equilibrium.

### The Existence, the Uniqueness, and the Stability of an Equilibrium

There exists a stable equilibrium subject to parameter conditions. First, I impose Assumption 1, which I assume in the rest of the paper.

**Assumption 1.** $\gamma > \frac{(1-\alpha)/(\sigma-1)-\alpha}{1+(1-\alpha)\min_{k\in\mathcal{K}}\{\epsilon_k\}/(1-\eta)}$

Assumption 1 is a sufficient condition for an equilibrium to exist as Proposition 4.

**Proposition 4** (Existence of equilibrium)**.** *Given Assumption 1, there exists an equilibrium.*

*Proof.* see Appendix D. □

Assumption 1 ensures that whatever the expenditure composition is, the agglomeration force on the real consumption is weaker than the aggregate dispersion force by inelastic land supply and heterogenous locational taste. This prevents a city from attracting all the workers. Without the term $\min_{k\in\mathcal{K}}\{\epsilon_k\}/(1-\eta)$, this becomes the condition for a unique stable equilibrium to exist in Redding (2016), who also has the dispersion forces from the heterogeneous taste over cities and inelastic land supply. Unlike Redding (2016), Proposition 4 concerns only the existence. This is because the uniqueness requires additional assumptions in my model due to the endogenous tradable sector shares. Even without the uniqueness, this paper will show that cross-city patterns in production, consumption, and wages become qualitatively the same in all the possible equilibria as I explain in the next section. In Appendix C, I explain sufficient conditions to have the uniqueness.

Next, I analyze whether there exists a stable equilibrium. When $U_1$ and $U_2$ can be expressed as functions of only $N_1$ from equations (8), (11), and (12), a stable equilibrium is defined as follows.

**Definition.** A competitive equilibrium is a stable equilibrium if and only if

$$\frac{d\left(U_1(N_1)a_1/U_2(N_1)a_2\right)^{1/\gamma}}{dN_1} < \frac{d\left(N_1/(N-N_1)\right)}{dN_1}$$

[3]This guarantees that, with the migration of infra-marginal agents who are indifferent between two cities ($U_1(N_1)a_1\delta(\theta,2) = U_2(N_1)a_2\delta(\theta,2)$), the expanding city would not experience enough of a relative gain in real consumption to support the post-migration population allocation.[4] With this

[3]With $V_1$ and $V_2$ s.t. $V_1 = a_1U_1N_1^{-\gamma}$ and $V_2 = a_2U_2N_2^{-\gamma}$, the condition of the stable equilibrium becomes $dV_1/dN_1 < dV_2/dN_1$. In the homogeneous agent's interpretation of the model, this guarantees that the expanding city does not offer a higher utility than the shrinking city.

[4]The product of the relative real consumption and amenity after the migration of $\varepsilon_1$ agents to city 1 becomes $\{U_1(N_1+\varepsilon_1)a_1\}/\{U_2(N_1+\varepsilon_1)a_2\}$. This supports population increase by $\varepsilon_2$ such that $(N_1 +$

condition, the infra-marginal agents in the expanding city find the shrinking one preferable and return to their original location. Consequently, the economy converges back to the original state. For the stability, Proposition 5 is obtained.

**Proposition 5** (Stability of equilibrium)**.** *Given Assumption 1, there exists a stable equilibrium.*

*Proof.* see Appendix D. □

When the dispersion force is strong enough, an equilibrium is stable. In the rest of the paper, I assume that Assumption 1 holds and focus on a stable equilibrium.

# 3 Cross-City Analysis

In this section, I illustrate how the two cities differ in the equilibrium depending on their fundamental productivities and amenities. I start with two partial equilibrium analyses to show the directions of the forces that the fundamental differences generate. Then, I lay out general equilibrium results in the case where the cities have the same amenity level and differ only in fundamental productivity ($a_1 = a_2$). The results of different amenities with the same productivity are briefly explained (A detailed explanation is provided in Appendix E).

## Partial Equilibrium Analyses

In the first partial equilibrium analysis, I illustrate how a difference in fundamentals generates a force that produces asymmetric population allocation and nominal wage inequality. In the second partial analysis, I explain the interplay between this force and the endogenous tradable sector shares. The second analysis reveals that the endogenous tradable sector shares amplifies the nominal wage inequality and generates a dispersion force when the sectors are gross complements.

---

$\varepsilon_2)/(N_2 - \varepsilon_2) = [\{U_1(N_1 + \varepsilon_1)a_1\}/\{U_2(N_1 + \varepsilon_1)a_2\}]^{1/\gamma}$ from equation (5). If $\epsilon_2 < \epsilon_1$, the new relative utility cannot support the migration. When $\epsilon_1 \to 0$, this is equivalent to the condition of the stability definition.

### Population and Wage

The first analysis uses two equilibrium conditions that relate the relative wage with the relative population. Making use of the population allocation with Fréchet utility shock (equation (5)), the real consumption in city 1 in equilibrium (equation (11)) and the home market effect (HME) on the wage rate (equation (8)) can be rewritten as follows:

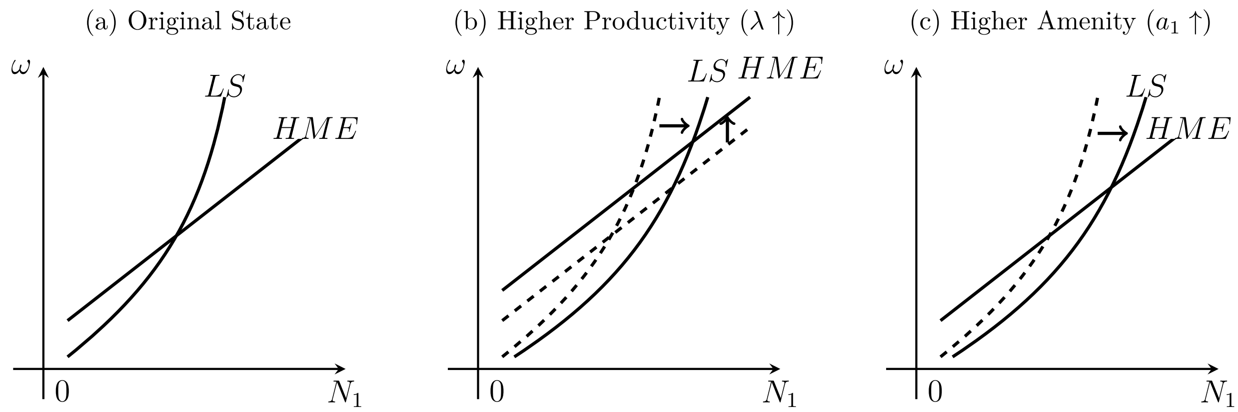
$$
\begin{aligned}
1 = & \left[\frac{1}{\lambda^{\sigma}}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu} \left[ \left(\frac{V}{a_1}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)} N_1^{\left\{\gamma\left(\epsilon_k+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{\sigma-1}\right\}} \right]^{(1-\eta)\mu} + \\
& \left[\frac{1}{\left(\lambda^{\sigma}+\rho\omega^{\sigma-1}\left(N_2/N_1\right)\left(m_{2\mathcal{T}}/m_{1\mathcal{T}}\right)\right)}\right]^{\frac{1-\eta}{\sigma-\eta}} \cdot \\
& \sum_{k\in\mathcal{K}} \tilde{\beta}_k^{\mu} \left[ \left(\frac{V}{a_1}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)} N_1^{\left\{\gamma\left(\epsilon_k+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{\sigma-1}\right\}} \right]^{(1-\eta)\mu} \qquad (15) \\
N_1 = & \frac{m_{2\mathcal{T}}}{m_{1\mathcal{T}}}\omega^{2\sigma-1}\left[\frac{\lambda^{-\sigma}-\rho\omega^{-\sigma}}{\lambda^{\sigma}-\rho\omega^{\sigma}}\right] N_2 \qquad (16)
\end{aligned}
$$

where $V = a_2 U_2 N_2^{\gamma}$ . In this partial equilibrium analysis, I focus on city 1 and fix $V(= a_2 U_2 N_2^{\gamma})$, $N_2$, and the relative tradable sector share $(m_{1\mathcal{T}}/m_{2\mathcal{T}})$. Then, equation (15) can be interpreted as the labor supply curve in city 1, and, given Assumption 1, $N_1$ increases in $\omega$. As the wage increases, the city attracts more workers. Similarly, $N_1$ increases in $\omega$ from the HME on the wage rate (equation (16)). Having a large local market requires a higher input cost to keep the profit at zero. These two curves are depicted in Figure 2a. [5]

When city 1 has a higher productivity ($\lambda \uparrow$), the HME curve shifts up and the labor supply curve shifts to the right, as in Figure 2b. The HME curve shifts for two reasons. First, the local wage linearly increases in local productivity. Consequently, the local market size expands and the

[5] How the curves intersect is not easy to tell from the equations in this part. The depiction here is based on the theoretical results in the general equilibrium. For cities with asymmetric productivity, the one with higher productivity offers higher local wage and larger population. For cities with asymmetric amenity, the one with higher amenity offers a higher local wage and has larger population if $(1-\alpha)/(\sigma-1) > \alpha$ and $\eta < 1$. These are consistent with the labor supply curve intersecting the home market effect curve from below.

Figure 2: Partial Equilibrium Analysis on $N_1$ and $\omega$




input cost, $w_1$, needs to rise to keep the zero profit. The shift of the labor supply curve is driven by additional local varieties. Given the labor supply, the higher productivity increases the mass of local varieties. Given the prices of varieties, which are linear in the wages $w_1$, this reduces the price indices and raises the local real consumption, and, therefore, it attracts more workers. Because of the shifts of the two curves, the new intersection is located where both the population and the wage are higher than before.

When city 1 has a higher amenity ($a_1 \uparrow$), this shifts the labor supply curve to the right, although it does not affect the HME curve, as in Figure 2c. This reflects the fact that the higher amenity attracts more people, but it does not affect production. As a result, the intersection moves along the HME curve, and both the population and the wage are higher than before. [6]

This analysis, which ignores what is taking place in the other city $(N_2, U_2)$ and the effect through the relative tradable sector share $(\sum_{k\in\mathcal{T}} x_{2k}/\sum_{k\in\mathcal{T}} x_{1k})$, shows the main forces generated by heterogeneous fundamental productivity and amenity. In the next analysis, I show how the relative tradable sector share behaves and related to the relative wage when population increases.

[6] While this result is in contrast to what a Rosen-Roback model (Rosen (1979), Roback (1982)) implies, this is not new in the literature. For example, Glaeser and Gottlieb (2009) point out that rising amenities can increase wages because of agglomeration economy.

### Wage and Tradable Sector Share

In the second partial equilibrium analysis, I focus on the relationship between the relative wage ($\omega$) and the aggregate tradable sector share ($m_{1\mathcal{T}}$). As in the first analysis, it uses two equilibrium conditions. First, the home market effect on wage (equation (8)) positively associates the relative wage and the aggregate tradable sector share, which is depicted as the HME-by-$m_{\mathcal{T}}$ curve in Figure 3. Second, the aggregate tradable sector share increases and decreases with the relative wage when the sectors are gross complements ($\eta < 1$) and gross substitutes ($\eta > 1$), respectively (equations (13) and (14)). This is depicted as the ZPC+Substitution curve in Figure 3.

The ZPC+Substitution curve reflects two economic forces. The first one is the zero profit conditions (ZPCs) in tradable and nom-tradable sectors. These conditions can be written as follows: (see equation (29) in Appendix D for the derivation)

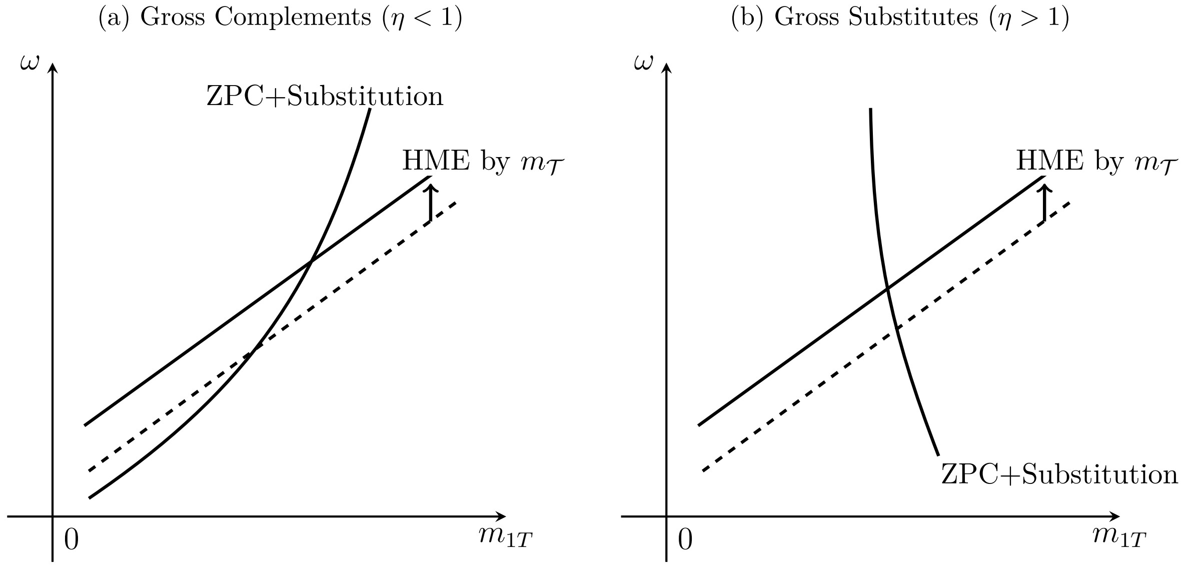
$$\text{For } k \in \mathcal{T},\ \frac{(w_i/\lambda_i)^{\sigma} - \rho\,(w_j/\lambda_j)^{\sigma}}{1-\rho^2} = N_i \tilde{\beta}_k \tilde{P}_{ik}^{\sigma-\eta} U_i^{\epsilon_k} C_i^{1-\eta} E_{iC}^{\eta}$$
$$\text{For } k \in \mathcal{N},\ \left(\frac{w_i}{\lambda_i}\right)^{\sigma} = N_i \tilde{\beta}_k \tilde{P}_{ik}^{\sigma-\eta} U_i^{\epsilon_k} C_i^{1-\eta} E_{iC}^{\eta}$$

By combining these conditions, it follows

$$\text{For } k \in \mathcal{T},\ \ell \in \mathcal{N},\ \frac{1-\rho\omega^{-\sigma}\lambda^{\sigma}}{1-\rho^2} = \frac{\tilde{\beta}_k}{\tilde{\beta}_\ell}\left(\frac{\tilde{P}_{1k}}{\tilde{P}_{1\ell}}\right)^{\sigma-\eta} U_1^{\epsilon_k-\epsilon_\ell}$$

This shows that, ignoring the effect from the heterogeneous income elasticities ($U_1^{\epsilon_k-\epsilon_\ell}$), a higher relative wage requires a higher price index in the tradable sectors relative to the non-tradable sectors ($\tilde{P}_{1k}/\tilde{P}_{1\ell}$). A higher relative wage is disadvantageous for firms only in the tradable sectors, and, therefore, it must be associated with a change that is relatively advantageous for the tradable sectors to keep the zero profits in both group. The second force is substitution between the tradable sectors and the non-tradable sectors. The higher price index in the tradable sectors implies a higher (lower) aggregate tradable sector share when the sectors are gross complements (substitutes).

Figure 3: Partial Equilibrium Analysis on $\omega$ and $m_{1\mathcal{T}}$ when $N_1$ increases




When city 1 attracts workers ($N_1 \uparrow$), the effects on the relative wage and the tradable sector share depend on whether the sectors are gross complements or gross substitutes. The larger population shifts the HME-by-$m_{\mathcal{T}}$ curve upwards because of the home market effect (equation (8)). When the sectors are gross complements (Figure 3a), workers spend more expenditures on the tradable sectors that become relatively more expensive, and this raises the relative wage further in the new intersection by the home market effect through the expenditure share. When the sectors are gross substitutes (Figure 3b), the residents spend less expenditures on the tradable sectors that become relatively more expensive. This attenuates the wage increase by the home market effect through the expenditure share.

The endogenous tradable sector share works as an additional dispersion (agglomeration) force when the sectors are gross complements (substitutes). As a result of the increasing tradable sector share in the gross-complement case, the mass of available varieties in the non-tradable sectors decreases, and the non-tradable sectors become expensive in terms of the price index. [7] The

[7] This can be verified with the price index of a non-tradable sector relative to income $\tilde{P}_{ik}/\{w_{ik}/(1-\alpha)\} = (1-\alpha)\lambda_i^{\sigma/(1-\sigma)} x_{ik}^{1/1-\sigma} N_i^{1/1-\sigma}$, which decreases in the employment share $x_{ik}$.

tradable sectors become even more expensive because the price index rises relative to that of the non-tradable sectors due to the zero profit conditions. [8] Thus, the both tradable and non-tradable sectors become expensive, and the real consumption decreases. [9] This can be interpreted as a force that counteracts the population increase–a dispersion force. On the other hand, the tradable sector share decreases when the sectors are gross substitutes. This attenuates the relative wage increase and increases the real consumption. This can be interpreted as an agglomeration force that does not appear when we fix the sectoral shares.

In the next section, I provide the results in general equilibrium for the case in which cities differ only in terms of productivity. Appendix E provides the results for the case in which there is a different amenity and a common productivity. The city that has better fundamental characteristics (productivity or amenity) becomes larger.[10] In the asymmetric productivity case, the large city offers a higher wage, and it specializes in income-elastic sectors. In the asymmetric amenity case, the same results are obtained with additional assumptions such that the sectors being gross complements or that the agglomeration force is stronger than the dispersion force from the inelastic land supply.

## General Equilibrium with Productivity Difference

In this section, I consider cities that have the same amenity level and differ only in fundamental productivity, and I provide a cross-city analysis for a stable equilibrium. The first comparison concerns population allocation and relative nominal income. The partial equilibrium analysis (Figure 2b) suggests that city 1 has the larger population and the residents receive the higher nominal

[8] It does so despite the increase in workers in the tradable sectors. This is possible if the mass of varieties shipped from city 2 decreases, which indeed happens in the general equilibrium.

[9] The change in the tradable sector shares increases the relative wage and reduces the real consumption. This relationship between the relative wage and the real consumption can be seen in equation (11). Holding $N_1$ constant, with the assumption for global monotonicity of the utility function $1+(1-\alpha)\min\{\epsilon_k\}/(1-\eta) > 0$, the real consumption $U_1$ decreases with the relative wage $\omega$.

[10] The amenity case requires either the sectors being gross complements or the sectors being gross substitutes with Assumption 2, which appears later in the main text.

income, which is proportional to the wage in an equilibrium. When we consider a general equilibrium and allow additional variables, including $N_2$, to move, the movement of $N_2$ attenuates the agglomeration economy in city 1 because of the decrease in the mass of varieties shipped from city 2. As a result, this shortens the shift of the labor supply curve in Figure 2b. On the other hand, the shrinking population in city 2 amplifies the home market effect, and it shifts further up the HME curve in Figure 2b. After all, it can be proven that the qualitative result does not change, as long as the price elasticity is not too high, and Proposition 6 is obtained.

**Proposition 6** (Cross-City Population and Wage Patterns with Gross Complements)**.** *Suppose that the productivities are asymmetric ($\lambda_1 > \lambda_2$), the amenities are symmetric ($a_1 = a_2$), and the sectors are gross complements ($\eta < 1$). Then, given an equilibrium with Assumption 1, the fundamentally productive city is larger and the relative wage is greater than the relative fundamental productivity.*

$$N_1 > N_2,\ \omega > \lambda > 1$$

*Proof.* see Appendix D. □

Although Assumption 1 is not a sufficient condition for a unique equilibrium, this result applies to all possible equilibria. The second partial equilibrium analysis (Figure 3) suggests that there is an additional dispersion force from the tradable sector share when the sectors are gross complements. This force and Assumption 1 ensure that the aggregate dispersion force is stronger than the agglomeration force, globally, and the qualitative results become the same as the first partial equilibrium analysis. In contrast, when the sectors are gross substitutes, there is an additional agglomeration force from the tradable sector share. Despite this additional force, Assumption 2 is sufficient to ensure that the additional dispersion force is not so strong that the aggregate dispersion force fails to be globally stronger than the agglomeration force.

**Assumption 2.** $\eta < 1 + (1-\alpha)\left(1 + \gamma \min_{k \in \mathcal{T}} \epsilon_k\right) / (\gamma + \alpha)$

The estimated inter-sectoral price elasticities of substitution by Comin et al. (2021) are 0.07-0.13 with 10 sectors. This indicates that the sectors being gross complements ($\eta < 1$) or the sectors being gross substitutes with Assumption 2 is not a strong assumption when the number of sectors is comparable to 10. Given Assumption 2, the same population and wage patterns can be obtained as the gross-complement case.

**Proposition 7** (Cross-City Population and Wage Patterns with Gross Substitutes)**.** *Suppose that the productivities are asymmetric ($\lambda_1 > \lambda_2$), the amenities are symmetric ($a_1 = a_2$), the sectors are gross substitutes ($\eta > 1$), and Assumption 2 holds. Then, given an equilibrium with Assumption 1, the fundamentally productive city is larger and the relative wage is greater than the relative fundamental productivity.*

$$N_1 > N_2, \ \omega > \lambda > 1$$

*Proof.* see Appendix D. □

The results of Propositions 6 and 7 are consistent with a stylized fact that nominal income is higher in larger cities, even when observable or unobservable workers' characteristics are taken into account (e.g., Behrens and Robert-Nicoud (2015),Glaeser and Mare (2001)). The next result concerns expenditure shares.

**Proposition 8** (Cross-City Expenditure Pattern)**.** *Suppose that the productivities are asymmetric ($\lambda_1 > \lambda_2$), the amenities are symmetric ($a_1 = a_2$), and the sectors are either gross complements ($\eta < 1$), or gross substitutes ($\eta > 1$) with Assumption 2 holding. Then, given an equilibrium with Assumption 1, the sectoral expenditure share ratio of city 1 to city 2 ($m_{1k}/m_{2k}$) increases in the income elasticity ($\epsilon_k$) within the non-tradable sectors and within the tradable sectors.*

*Proof.* Equations (9) and (10) provide the expenditure share ratio of a non-tradable sector and a tradable sector, respectively,

$$\forall k \in \mathcal{N}, \quad \frac{m_{1k}}{m_{2k}} = \lambda^{-\sigma\frac{1-\eta}{\sigma-\eta}} \left(\frac{U_1}{U_2}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \left(\frac{N_1}{N_2}\right)^{\left(\frac{\alpha}{1-\alpha}-\frac{1}{\sigma-1}\right)(1-\eta)\mu}$$

$$\forall k \in \mathcal{T}, \quad \frac{m_{1k}}{m_{2k}} = \left[\frac{\lambda^{-\sigma}-\rho\omega^{-\sigma}}{1-\rho\lambda^{-\sigma}\omega^{\sigma}}\right]^{\frac{1-\eta}{\sigma-\eta}} \left(\frac{U_1}{U_2}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \left(\frac{N_1}{N_2}\right)^{\left(\frac{\alpha}{1-\alpha}-\frac{1}{\sigma-1}\right)(1-\eta)\mu}$$

$N_1 > N_2$ in the equilibrium implies $U_1 > U_2$ according to equation (5), and the expenditure share ratio increases in $\epsilon_k$ within the non-tradable sectors and within the tradable sectors. □

In the equilibrium, the large city spends more expenditure in the income-elastic sectors. This difference in the expenditure shares generates the following results in the employment shares.

**Proposition 9** (Home Market Effect on Sectoral Specialization)**.** *Given an equilibrium, for a tradable sector, there is a relationship with the within-tradable employment share ratio and the within-tradable expenditure share ratio such that*

$$\forall k \in \mathcal{T}, \ \frac{\tilde{x}_{1k}}{\tilde{x}_{2k}} = \frac{(\tilde{m}_{1k}/\tilde{m}_{2k}) - \rho\lambda^{\sigma}\omega^{-\sigma}}{1-\rho\lambda^{\sigma}\omega^{-\sigma}} \frac{1-\rho\omega^{\sigma}\lambda^{-\sigma}}{1-\rho\omega^{\sigma}\lambda^{-\sigma}(\tilde{m}_{1k}/\tilde{m}_{2k})}$$

*where* $\tilde{m}_{ik} = m_{ik}/\sum_{k\in\mathcal{T}} m_{ik}$ *and* $\tilde{x}_{ck} = x_{ik}/\sum_{k\in\mathcal{T}} x_{ik}$*. This shows* $\tilde{x}_{1k}/\tilde{x}_{2k}$ *increases in* $\tilde{m}_{1k}/\tilde{m}_{2k}$*. That is, given an equilibrium, the greater the within-tradable expenditure share difference of a sector, the greater that sector's within-tradable employment share difference. Also, this implies that a city becomes the net exporter in sectors for which the city has a greater within-tradable expenditure share as follows:*

$$\forall k \in \mathcal{T}, \ \tilde{m}_{1k} > \tilde{m}_{2k} \iff x_{1k} > m_{1k}$$

*Proof.* see Appendix D . □

Then, Corollary 2 follows.

**Corollary 2** (Cross-City Trade Pattern)**.** *Suppose that the productivities are asymmetric ($\lambda_1 > \lambda_2$), the amenities are symmetric ($a_1 = a_2$), and the sectors are either gross complements ($\eta < 1$), or gross substitutes ($\eta > 1$) with Assumption 2 holding. Then, given an equilibrium with Assumption 1, there exists sector $s$ such that the fundamentally productive city becomes a net exporter for all sector $k$ whose income elasticity is not smaller than that of sector $s$ ($\epsilon_k \geq \epsilon_s$). Also, there exists sector $s$ such that the fundamentally productive city becomes a net importer for all sector $k$ whose income elasticity is not greater than that of sector $s$ ($\epsilon_k \leq \epsilon_s$).*

$$\exists s \in \mathcal{T}, \forall k \in \mathcal{T} \ s.t. \ \epsilon_k \geq \epsilon_s, \ x_{1k} > m_{1k}$$

$$\exists s \in \mathcal{T}, \forall k \in \mathcal{T} \ s.t. \ \epsilon_k \leq \epsilon_s, \ x_{1k} < m_{1k}$$

This is the home market effect in the sectoral specialization. The difference in the expenditure pattern generates comparative advantages and is amplified to that of the employment pattern. Unlike Krugman (1980), who assumes an exogenous taste difference to generate the heterogeneous relative demand, here that difference arises endogenously from the non-homothetic preference. The importance and the endogenous formation of the relative demand are the same as in Matsuyama (2019), but my definition of the relative demand is different. The result in Proposition 9 demonstrates that when non-tradable sectors exist, the relative size of demand should be measured within the tradable sectors. In Matsuyama (2019), all sectors are tradable and the relative demand size is the same whether it is within the overall economy or within the tradable sectors.

Next, I analyze the aggregate tradable and non-tradable sector expenditure shares, which are equal to their employment shares from Corollary 1. In the equilibrium, it follows from equations

(9) and (10) that the ratios of these share are

$$\frac{m_{1\mathcal{T}}}{m_{1\mathcal{N}}} = \left[\frac{1-\rho\lambda^{\sigma}\omega^{-\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \frac{\sum_{k\in\mathcal{T}}\left(\tilde{\beta}_k U_1^{\epsilon_k}\right)^{\mu}}{\sum_{k\in\mathcal{N}}\left(\tilde{\beta}_k U_1^{\epsilon_k}\right)^{\mu}}$$

$$\frac{m_{2T}}{m_{2N}} = \left[\frac{1-\rho\lambda^{-\sigma}\omega^{\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \frac{\sum_{k\in\mathcal{T}}\left(\tilde{\beta}_k U_2^{\epsilon_k}\right)^{\mu}}{\sum_{k\in\mathcal{N}}\left(\tilde{\beta}_k U_2^{\epsilon_k}\right)^{\mu}}$$

where $m_{i\mathcal{N}} = \sum_{k\in\mathcal{N}} m_{ik}$ for $i \in \{1,2\}$. This shows that the aggregate tradable sector share is determined by two forces. The first is the relative price of goods in the tradable sectors, which shows up in the form of the first factor in each RHS. Other things being equal, this force makes the large city has a greater (smaller) tradable sector share when the sectors are gross complements (substitutes) as discussed in the second partial equilibrium analysis (Figure 3). This is captured by $1-\rho\lambda^{\sigma}\omega^{-\sigma} > 1-\rho\lambda^{-\sigma}\omega^{\sigma}$ as $\omega > \lambda$ from Propositions 6 and 7. The second force is the income-elasticities of the two groups. Because the real consumption is higher in city 1 ($U_1 > U_2$) in the equilibrium, the higher the income-elasticities of tradable sectors as a whole are compared to those of the non-tradable sectors, the higher the aggregate tradable sector share in city 1 is compared to that in city 2, other things being equal. [11]

Finally, the price index of a sector differs between locations. The sector price index ratio is given by

$$\frac{P_{1k}}{P_{2k}} = \begin{cases} \frac{w_1}{w_2}\left(\frac{\lambda N_1}{N_2}\frac{x_{1k}}{x_{2k}}\right)^{-\frac{1}{\sigma-1}} & k\in\mathcal{N} \\ \frac{1}{\tau}\left(1+\frac{\tau^{2(\sigma-1)}-1}{(\tau w_2/w_1)^{\sigma-1}(\lambda N_1/N_2)(x_{1k}/x_{2k})+1}\right)^{\frac{1}{\sigma-1}} & k\in\mathcal{T} \end{cases}$$

In both groups, the sector price index ratio decreases in the employment share ratio ($x_{1k}/x_{2k}$). As the employment share ratio increases in income-elasticity within a group, the sector price index decreases in income-elasticity within a group. In city 1, the higher expenditure shares on income-

[11] This observation becomes important when we think about the effect of sector-specific trade cost reductions. When a non-tradable sector becomes a tradable, it changes the income-elasticities of the two groups and, thereby, the aggregate tradable sector share. Onoda (2022) examines how business services' trade cost reduction has affected cross-city income inequality by this mechanism.

elastic goods attract firms in those sectors, and the price indices in those sectors become relatively inexpensive, reflecting relatively more varieties.

## Connection to Stylized Fact

In summary, the fundamentally productive location becomes the large and high-income city, and within each group (tradable and non-tradable), the residents allocate their expenditure relatively more towards income-elastic sectors, which offer relatively richer varieties. On the supply side, the large and high-income city specializes in income-elastic sectors in the sense that workers are employed relatively more in income-elastic sectors within each group, which replicates the sectoral specialization pattern seen in Figure 1. While what we observe in the real world are the city size and the sectoral employment of a city, what generates the relationship between them in this model is the fundamental productivity.

Additionally, the specialization pattern is consistent with the amplification mechanism in the model. The model predicts that the employment shares of income-elastic sectors are relatively higher in large cities and that they are even higher for tradable sectors because the home market effect amplifies the demand pattern as Proposition 9 and Corollary 2. In order to see if there is such a difference between tradable and non-tradable sectors in the data, I classify sectors in Figure 1 as non-tradable that have export shares smaller than 0.03 on average in the international trade data in Caron et al. (2020) and the others as tradable. [12] The regression line between the two elasticities in Figure 1 is re-estimated for each group, separately, and shown in Figure 4. The positive relationship is clearly steeper for the tradable sectors; the slope of the regression line for the tradable sectors (solid) is 1.06 while that of the non-tradable sectors (dashed) is 0.30. This

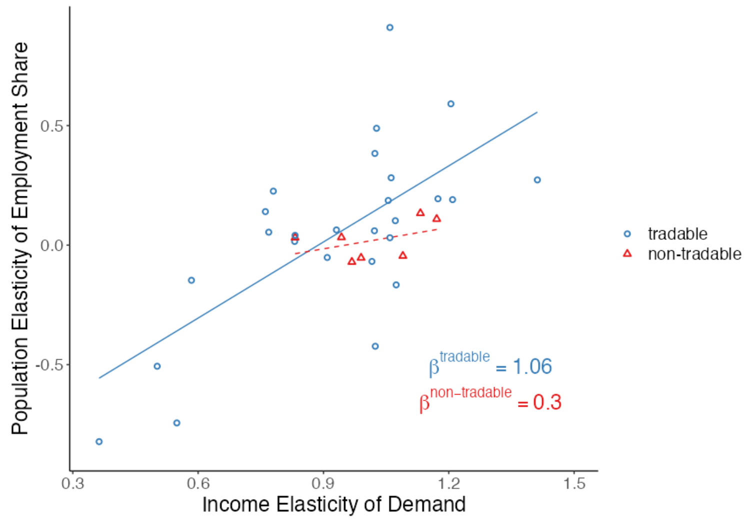
---

12

Non-tradable sectors are “Water”, “Electricity”, “Construction”, “Trade” (including wholesale trade, retail sales, hotels, and restaurants), “Recreational and other services”, “Public Administration, Defense, Education, Health”, and “Financial services nec”.

means that the employment share of an income-elastic sector increases with the city size faster for a tradable sector than for a non-tradable sector, which is consistent with the amplification mechanism in the tradable sectors.

Figure 4: Elasticity of Employment Share with respect to MSA's Population and Elasticity of Demand with respect to Income for Tradable and Non-tradable




MSA population are 2016 data from the Bureau of Economic Analysis. Employment shares are calculated from 2016 County Business Pattern data. NAICS data are mapped to GTAP sectors. When obtaining the population elasticity of employment share, region $\in \{Northeast, Midwest, South, West\}$ is controlled. Income-elasticity estimates are from Caron et al. (2020). The unweighted regression line is shown. For details of data description, see Appendix A.

# 4 Robustness Check

In this section, I address an alternative explanation of the cities' specialization patterns in Figure 1. The positive relationship is possibly driven by an omitted variable--that is, skilled-labor supply in cities. This is a reasonable concern given that skilled workers, who are at the same time high income earners, tend to reside in large cities, and those cities tend to host skill-intensive sectors (Davis and Dingel (2020)). Also, and as is well known, there is a positive correlation between the skill intensity and the income-elasticity of a sector (Caron et al. (2014) and Caron et al. (2020)).

I address this concern in two ways. First, I test if the positive relationship in Figure 1 is robust to controlling for the skill intensities of sectors. Second, I use the elasticities of employment share that are conditioned on local skilled-labor supply for Figure 1.

In Appendix F, I provide an empirical exercise with an alternative specification as an additional robustness check. The specification follows Nunn (2007), and it tests the association of the income level and the income-elasticities of the local industries. This test also shows that there is a positive relationship even when we control for the association of the skilled-labor supply and the skill intensities.

## Controlling Skill Intensities in Figure 1

In order to test if the positive relationship in Figure 1 is robust to controlling the skill intensities of sectors, I regress the elasticities of employment share on the income-elasticities of demand and the skill intensities. The regression model is given by

$$\xi_k = \alpha + \beta\epsilon_k + \gamma\theta_k + e_k$$

where

$\xi_k$ : elasticity of employment share of industry $k$ with respect to MSA's population

$\epsilon_k$ : income elasticity of demand for industry $k$ output

$\theta_k$ : skill intensity of industry $k$

$e_k$ : error term for industry $k$

$\xi_k$ is the variable of the y-axis in Figure 1. Each $\xi_k$ is obtained by regressing the log of employment shares of sector $k$ in MSAs on the log of MSAs' population sizes. The population data are from the Bureau of Economic Analysis and the employment data are from the Country Business Patterns. In theses log regressions, the region of a MSA is controlled for. $\epsilon_k$ is the variable of the x-axis in Figure 1 and $\theta_k$ is the control. For these two variables, I use the estimates by Caron et al. (2020), who obtain the estimates through a structural estimation with international trade data that have both heterogeneous skill-intensities and income-elasticities. If the cities' specialization pattern in Figure 1 only reflects that large cities specialize in skill-intensive sectors by the supply side mechanism (Davis and Dingel (2020)), $\beta = 0$ and $\gamma > 0$ are expected. I implement the cross-sectional regression for three years and the results are shown in Table 1.

Table 1: Results with Controlling Skill Intensity

| | Population elasticity of employment share | | |
|---|---|---|---|
| | 2006 | 2011 | 2016 |
| | (1) | (2) | (3) |
| Income elasticity ($\beta$) | 1.288*** | 1.201*** | 0.937*** |
| | (0.337) | (0.318) | (0.285) |
| Skill intensity ($\gamma$) | −0.145 | 0.067 | 0.132 |
| | (0.690) | (0.653) | (0.583) |
| Observations | 33 | 33 | 33 |
| $R^2$ | 0.464 | 0.486 | 0.433 |
| Adjusted $R^2$ | 0.429 | 0.452 | 0.395 |
| F Statistic (df = 2; 30) | 13.010*** | 14.183*** | 11.440*** |

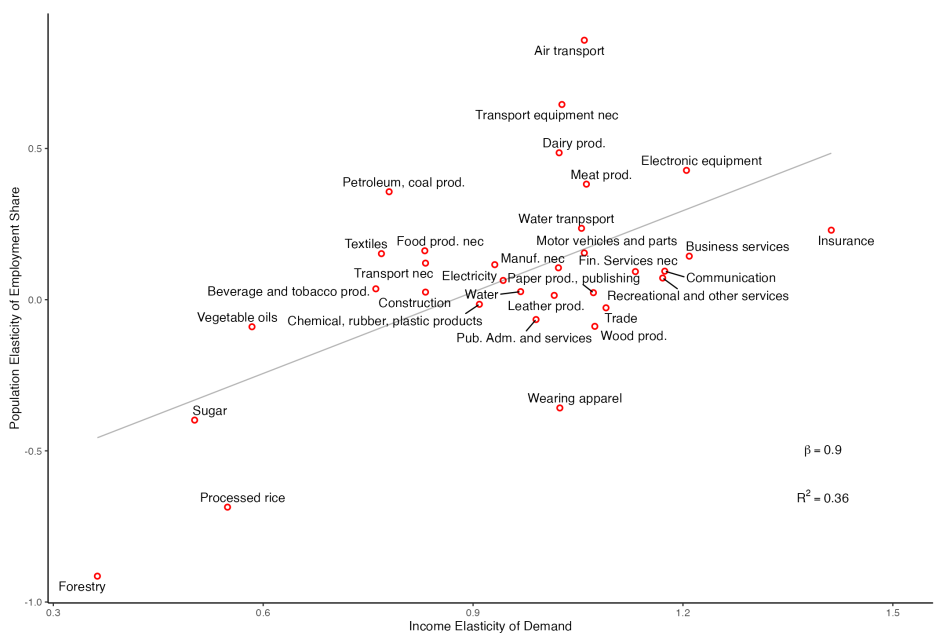
*Notes:* ***,**, and * denote significance at the 1 %, 5%, and 10% levels respectively.

In all three years, $\beta$ is significantly positive whereas $\gamma$ is not. A caveat is that all of $\xi_k$, $\epsilon_k$, and $\theta_k$ that are used in the regressions are estimated values, and, therefore, the standard errors here are not precise for the hypothesis test. Nevertheless, these results suggest that the positive relationship in Figure 1 is robust to the supply-side explanation.

## Controlling Skill Supply in Figure 1

The second robustness check is done by controlling for skill supply when the elasticities of employment share with respect to population and that of MSA's income are obtained. Specifically, the college employment ratio–the number of college graduates over non-graduates in the labor force–in a MSA is controlled for[13]. Figure 5 uses these conditional elasticities in the y-axes. As the regression line shows, the positive relationships still clearly exist. This shows that the alternative explanation cannot deny the mechanism that is proposed by my model.

Figure 5: With Population Elasticity of Employment Share Conditioned on Skill Supply




MSA population are 2016 data from the Bureau of Economic Analysis (BEA). The income of a MSA is per capita personal income in 2016; from the BEA. Employment shares are calculated from 2016 CBP data. When obtaining the income-elasticity of employment share, the college employment ratio of the labor force and the region $\in \{Northeast, Midwest, South, West\}$ are controlled for. Income-elasticity estimates are fromCaron et al. (2020). The unweighted regression lines are shown.

[13]The labor force data are from Census via IPUMS.

# 5 Conclusion

In this paper, I first report a new stylized fact: large cities specialize in income-elastic sectors. To explain this, I present a two-city model and demonstrate how a fundamental difference in productivity generates the patterns of a city's size, income level, and industrial composition, which is consistent with the production patterns. In addition, the model reveals the importance of the aggregate share of tradable sectors in local wages. This provides a new reason why we should understand the mechanism that drives the industrial composition. Because the aggregate share of tradable sectors is characterized by substantial variation across cities, this result suggests the importance of understanding a city's industrial composition. The endogenous and heterogeneous shares of tradable and non-tradable sectors deserve more investigations, and my model can be a useful framework for such studies both in international trade and in urban economics.

# References


Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, 103(6):2121–68.

Behrens, K. and Robert-Nicoud, F. (2015). Agglomeration theory with heterogeneous agents. *Handbook of regional and urban economics*, 5:171–245.

Caron, J., Fally, T., and Markusen, J. (2020). Per capita income and the demand for skills. *Journal of International Economics*, 123:103306.

Caron, J., Fally, T., and Markusen, J. R. (2014). International trade puzzles: A solution linking production and preferences. *The Quarterly Journal of Economics*, 129(3):1501–1552.

Carrico, C., Jones, L., and Tsigas, M. E. (2012). Disaggregate us labor statistics for the usage 2.0 and gtap applied general equilibrium models. *Available at SSRN 2169415*.

Comin, D., Lashkari, D., and Mestieri, M. (2021). Structural change with long-run income and price effects. *Econometrica*, 89(1):311–374.

Davis, D. R. and Dingel, J. I. (2019). A spatial knowledge economy. *American Economic Review*, 109(1):153–70.

Davis, D. R. and Dingel, J. I. (2020). The comparative advantage of cities. *Journal of International Economics*, 123:103291.

Dingel, J. I. (2017). The determinants of quality specialization. *The Review of Economic Studies*, 84(4):1551–1582.

Duranton, G. and Puga, D. (2005). From sectoral to functional urban specialisation. *Journal of urban Economics*, 57(2):343–370.

Eckert, F., Hejlesen, M., and Walsh, C. (2022). The return to big-city experience: Evidence from refugees in denmark. *Journal of Urban Economics*, 130:103454.

Fajgelbaum, P., Grossman, G. M., and Helpman, E. (2011). Income distribution, product quality, and international trade. *Journal of political Economy*, 119(4):721–765.

Flam, H. and Helpman, E. (1987). Vertical product differentiation and north-south trade. *The American Economic Review*, pages 810–822.

Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, 47(4):983–1028.

Glaeser, E. L. and Mare, D. C. (2001). Cities and skills. *Journal of labor economics*, 19(2):316–342.

Handbury, J. (2019). Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities. Technical report, National Bureau of Economic Research.

Henderson, J. V. and Ono, Y. (2008). Where do manufacturing firms locate their headquarters? *Journal of Urban Economics*, 63(2):431–450.

Hoelzlein, M. (2019). Two-sided sorting and spatial inequality in cities.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3):483–499.

Matsuyama, K. (2019). Engel's law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica*, 87(2):497–528.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics*, 122(2):569–600.

Onoda, T. (2022). Heterogenous tradable sector shares and wages. *Unpublished Manuscript*.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics*, 101:148–167.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6):1257–1278.

Rosen, S. (1979). *Wage-Based Indexes of Urban Quality of Life*. Baltimore: Johns Hopkins University Press.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, 88(4):641–658.

Stokey, N. L. (1991). The volume and composition of trade between rich and poor countries. *The Review of Economic Studies*, 58(1):63–80.

Tabuchi, T. and Thisse, J.-F. (2002). Taste heterogeneity, labor mobility and economic geography. *Journal of Development Economics*, 69(1):155–177.

# Appendix

## A Data

To create Figure 1, I borrow estimates of income-elasticities from Caron et al. (2020). Using 2007 international trade data for 109 countries, these scholars estimate the elasticities for 49 sectors. The elasticity varies from 0.137 for "Processed rice" to 1.311 for "Financial services nec". For employment data, I use datasets from Country Business Patterns (CBP). CBP provides employment data for sectors classified annually according to the North American Industry Classification System (NAICS) in metropolitan areas. The classification in Caron et al. (2020), which is different from NAICS, is called GTAP. In most cases, one GTAP code corresponds to multiple 3-digit or 4-digit NAICS codes. Following Carrico et al. (2012), and mapping NAICS data to GTAP, I create employment data by GTAP. For income level of MSAs, I use per capita personal income available in the Bureau of Economic Analysis. Table 2 displays the distribution of income levels and population in 2016. The largest MSA in 2016 in this sample is New York-Newark-Jersey City, (NY-NJ-PA), which had a population of 20,275,179, while the smallest is Parkersburg-Vienna, WV, which had a population of 91,488. The college employment ratio, which I use in a robustness check, varies substantially across MSAs: 0.176 in Hanford-Corcoran, CA, is the lowest and 1.194 in San Jose-Sunnyvale-Santa Clara, CA, is the highest.

Table 2: Distribution of Population and College Employment Ratio across MSAs in 2016

| | Min | Q1 | Median | Mean | Q3 | Max |
|---|---|---|---|---|---|---|
| Population | 91,488 | 177,627 | 373,568 | 984,042 | 850,187 | 20,275,179 |
| College employment ratio | 0.176 | 0.319 | 0.437 | 0.455 | 0.549 | 1.194 |

The elasticities of employment share with respect to MSA's population that I estimated and the income elasticities and skill intensities that I borrow from Caron et al. (2020) are summarized in Table 3.

Table 3: Estimates for Industries

| GTAP code | Industry | Income elasticity | Skill intensity | Population elasticity of employment share | |
|---|---|---|---|---|---|
| | | | | Estimate | S.E. |
| atp | Air transport | 1.06 | 0.30 | 0.91 | 0.09 |
| ele | Electronic equipment | 1.21 | 0.38 | 0.59 | 0.11 |
| otn | Transport equipment nec | 1.03 | 0.34 | 0.49 | 0.12 |
| mil | Dairy prod. | 1.02 | 0.24 | 0.38 | 0.13 |
| omt+cmt | Meat prod. | 1.06 | 0.22 | 0.28 | 0.12 |
| isr | Insurance | 1.41 | 0.52 | 0.27 | 0.04 |
| p_c | Petroleum, coal prod. | 0.78 | 0.35 | 0.23 | 0.11 |
| cmn | Communication | 1.17 | 0.50 | 0.19 | 0.03 |
| obs | Business services | 1.21 | 0.49 | 0.19 | 0.01 |
| wtp | Water transport | 1.06 | 0.32 | 0.19 | 0.11 |
| b_t | Beverage and tobacco prod. | 0.76 | 0.28 | 0.14 | 0.08 |
| ofi | Fin. Services nec | 1.13 | 0.53 | 0.13 | 0.02 |
| ros | Recreational and other services | 1.17 | 0.48 | 0.11 | 0.02 |
| ppp | Paper prod., publishing | 1.07 | 0.34 | 0.10 | 0.03 |
| ome | Machinery and equipment nec | 0.93 | 0.37 | 0.06 | 0.06 |
| omf | Manuf. nec | 1.02 | 0.27 | 0.06 | 0.05 |
| tex | Textiles | 0.77 | 0.23 | 0.05 | 0.07 |
| otp | Transport nec | 0.83 | 0.29 | 0.04 | 0.03 |
| ely | Electricity | 0.94 | 0.37 | 0.03 | 0.06 |
| cns | Construction | 0.83 | 0.30 | 0.03 | 0.02 |
| mvh | Motor vehicles and parts | 1.06 | 0.34 | 0.03 | 0.11 |
| ofd | Food prod. nec | 0.83 | 0.26 | 0.02 | 0.06 |
| trd | Trade | 1.09 | 0.30 | -0.05 | 0.01 |
| crp | Chemical, rubber, plastic products | 0.91 | 0.36 | -0.05 | 0.06 |
| osg | Pub. Adm. and services | 0.99 | 0.50 | -0.05 | 0.01 |
| lea | Leather prod. | 1.02 | 0.20 | -0.07 | 0.10 |
| wtr | Water | 0.97 | 0.37 | -0.07 | 0.08 |
| vol | Vegetable oils | 0.58 | 0.22 | -0.15 | 0.11 |
| lum | Wood prod. | 1.07 | 0.25 | -0.17 | 0.07 |
| wap | Wearing apparel | 1.02 | 0.23 | -0.42 | 0.08 |
| sgr | Sugar | 0.50 | 0.20 | -0.51 | 0.08 |
| pcr | Processed rice | 0.55 | 0.12 | -0.74 | 0.06 |
| frs | Forestry | 0.36 | 0.14 | -0.82 | 0.10 |

# B Derivation of Demand Function and Indirect Utility

### Derivation of Demand Function

The household problem is given by

$$
\begin{aligned}
\max_{U_i,H_i,C_i,\left\{Q_{i,k}\right\}_{k\in\mathcal{K}},\left\{q_{i,k}(\nu)\right\}_{\nu\in\Omega_{i,k},(i,k)\in(1,2)\times\mathcal{K}}} \quad & U_i \\
& +\lambda\left(U_i - H_i^{\alpha}C_i^{1-\alpha}\right) \\
& +\pi\left(C_i - \left[\sum_{k\in\mathcal{K}}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}\right) \\
& +\sum_{k\in\mathcal{K}}\xi_k\left(Q_{i,k} - \left[\int_{\Omega_{i,k}} q_{i,k}(\nu)^{\frac{\sigma-1}{\sigma}}d\nu\right]^{\frac{\sigma}{\sigma-1}}\right) \\
& +\phi\left(E_i - \sum_{k\in\mathcal{K}}\int_{\Omega_{i,k}} p_{i,k}(\nu)q_{i,k}(\nu)d\nu - P_i(H)H_i\right)
\end{aligned}
$$

where $\lambda$, $\pi$, $\{\xi_k\}_{k\in\mathcal{K}}$, and $\phi$ are the lagrange multipliers. The FOCs are:

$$
\begin{aligned}
U_i &: 1+\lambda-\mu\frac{\eta}{\eta-1}\left(\sum_{k\in\mathcal{K}}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}\right)^{\frac{\eta}{\eta-1}-1}\left(\sum_{k\in\mathcal{K}}\epsilon_k\frac{1}{\eta}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}U_i^{-\frac{1}{\eta}}Q_{i,k}\right)=0 \\
C_i &: \lambda(1-\alpha)H^{\alpha}C_i^{-\alpha}-\pi=0 \\
Q_{i,k} &: \pi\left(\sum_{k\in\mathcal{K}}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}-1}\left(\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}^{\frac{\eta-1}{\eta}-1}\right)-\xi_j=0 \\
q_{i,k}(\nu) &: \xi_k\left[\int_{\Omega_{i,k}} q_{i,k}(\nu)^{\frac{\sigma-1}{\sigma}}d\mu\right]^{\frac{\sigma}{\sigma-1}-1}q_{i,k}(\nu)^{\frac{\sigma-1}{\sigma}-1}-\phi p_{i,k}(\nu)=0 \\
H_i &: \lambda\alpha H_i^{\alpha-1}C_i^{1-\alpha}-\phi P_{i,H}=0
\end{aligned}
$$

First, derive the usual result with CES from the FOC w.r.t. $q_{i,k}(\nu)$ and the definition of $Q_{i,k}$.

$$
P_{i,k}Q_{i,k}^{\frac{1}{\sigma}} = p_{i,k}(\nu)q_{i,k}(\nu)^{\frac{1}{\sigma}}
$$

where $P_{i,k}=\left[\int_{\Omega_{i,k}(j)}\left(p_{i,k}(\nu)\right)^{1-\sigma}d\omega\right]^{1/(1-\sigma)}$It follows from this result, the FOC w.r.t $Q_{i,k}$, and that w.r.t $q_{i,k}(\nu)$,

$$\pi C_i\frac{\left(\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}^{\frac{\eta-1}{\eta}}\right)}{\left(\sum_{k\in\mathcal{K}}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}^{\frac{\eta-1}{\eta}}\right)^{\frac{1}{\eta-1}}}=\phi P_{i,k}Q_{i,k}$$

$$\Longrightarrow\frac{\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,k}^{\frac{\eta-1}{\eta}}}{\beta_l^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{i,\ell}^{\frac{\eta-1}{\eta}}}=\frac{P_{i,k}Q_{i,k}}{P_{i,l}Q_{i,\ell}} \tag{17}$$

Next, I obtain the expenditure on good consumption by using eq. (17).

$$\begin{aligned}E_{iC}\equiv\sum_{k\in\mathcal{K}}\int_{\Omega_{ik}}p_{ik}(\nu)q_{ik}(\nu)d\nu&=\sum_{k\in\mathcal{K}}P_{ik}Q_{ik}\\&=\sum_{k\in\mathcal{K}}P_{i\ell}Q_{i\ell}\frac{\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k}{\eta}}Q_{ik}^{\frac{\eta-1}{\eta}}}{\beta_l^{\frac{1}{\eta}}U_i^{\frac{\epsilon_l}{\eta}}Q_{i\ell}^{\frac{\eta-1}{\eta}}}\\&=\frac{P_{i\ell}Q_{i\ell}}{\beta_l^{\frac{1}{\eta}}U_i^{\frac{\epsilon_l}{\eta}}Q_{i\ell}^{\frac{\eta-1}{\eta}}}\sum_{k\in\mathcal{K}}\beta_k^{\frac{1}{\eta}}U_i^{\frac{\epsilon_k(1-\eta)}{\eta}}Q_{ik}^{\frac{\eta-1}{\eta}}\\&=\frac{P_{i\ell}Q_{i\ell}}{\beta_l^{\frac{1}{\eta}}U_i^{\frac{\epsilon_l}{\eta}}Q_{i\ell}^{\frac{\eta-1}{\eta}}}C_i^{\frac{\eta-1}{\eta}}\end{aligned}$$

The demand function immediately follows as:

$$Q_{i\ell}=\beta_\ell U_i^{\epsilon_k}P_{i\ell}^{-\eta}C_i^{1-\eta}E_{iC}^{\eta} \tag{18}$$

The expenditure on goods consumption follows from multiplying eq. (18) by $P_{i\ell}$ and aggregating it across sectors.

$$E_{iC}=\sum_{k\in\mathcal{K}}\beta_\ell U_i^{\epsilon_k}P_{i\ell}^{1-\eta}C_i^{1-\eta}E_{iC}^{\eta}$$

Then,

$$m_{ik} = \frac{\beta_k U_i^{\epsilon_k} P_{ik}^{1-\eta}}{\sum_{\ell \in \mathcal{K}} \beta_\ell U_i^{\epsilon_k} P_{i\ell}^{1-\eta}}$$

### Derivation of Price Index for Goods Consumption

Substitute eq. 18 for $Q_{i,\ell}$ in the definition of $C_i$.

$$\begin{aligned}
C_i &= \left( \sum_{k \in \mathcal{K}} \beta_k^{\frac{1}{\eta}} U_i^{\frac{\epsilon_k}{\eta}} \left[ \beta_k^{\frac{1}{\eta}} U_i^{\frac{\epsilon_k}{\eta}} P_{ik}^{-1} C_i^{\frac{1-\eta}{\eta}} E_{iC} \right]^{\eta-1} \right)^{\frac{\eta}{\eta-1}} \\
\iff E_{iC} &= \left( \sum_{k \in \mathcal{K}} \beta_k U_i^{\epsilon_k} P_{ik}^{1-\eta} \right)^{\frac{1}{1-\eta}} C_i \\
&= P_{iC} C_i
\end{aligned}$$

where $P_{i,C} \equiv \left( \sum_{k \in \mathcal{K}} \beta_k U_c^{\epsilon_k} P_{c,k}^{1-\eta} \right)^{1/(1-\eta)}$, which is the price index for goods consumption.

### Derivation of Indirect Real Consumption

It follows from the FOC w.r.t. $Q_{i,k}$ and $P_{i,k} Q_{i,k}^{\frac{1}{\sigma}} = p_{i,k}(\nu) q_{i,k}(\nu)^{\frac{1}{\sigma}}$ that

$$\xi_k P_{ck}^{-1} = \phi$$

The FOC w.r.t $q_{i,k}(\nu)$ can be transformed with this equation.

$$\begin{aligned}
\xi_k =& \pi C_i^{\frac{1}{\eta}} \left( \beta_k^{\frac{1}{\eta}} U_i^{\frac{\epsilon_k}{\eta}} Q_{ik}^{\frac{\eta-1}{\eta}-1} \right) \\
\implies -\phi \sum_{k \in \mathcal{K}} P_{ik} Q_{ik} =& \pi C_i^{\frac{1}{\eta}} \sum_{k \in \mathcal{K}} \left( \beta_k^{\frac{1}{\eta}} U_i^{\frac{\epsilon_k}{\eta}} Q_{ik}^{\frac{\eta-1}{\eta}} \right) \\
\iff -\phi E_{iC} =& \pi C_i
\end{aligned}$$

Then, I obtain the usual Cobb-Douglas result by combining this with the FOCs w.r.t. $C_i$ and $H_i$.

$$\frac{\alpha}{1-\alpha} = \frac{P_{iH}H_i}{P_{iC}C_i}$$

I can implicitly express the indirect real consumption by plugging this into the definition of $U_i$,

$$U_i = \alpha^\alpha(1-\alpha)^{1-\alpha}\frac{E_i}{P_{iC}^{1-\alpha}P_{iH}^{\alpha}} = \alpha^\alpha(1-\alpha)^{1-\alpha}\frac{E_i}{P_{iH}^{\alpha}}\left(\sum_{k\in\mathcal{K}}\beta_k U_i^{\epsilon_k}P_{ik}^{1-\eta}\right)^{\frac{\alpha-1}{1-\eta}}$$

**Derivation of Price Index for Good Consumption as Function of $E_i$, $L_i$, $N_i$, and $U_i$**

It can be obtained from the real good consumption with the land clearing condition.

$$\begin{aligned} U_i &= (1-\alpha)^{1-\alpha}\left(\frac{L_i}{N_i}\right)^{\alpha}\left(\frac{E_i}{P_{iC}}\right)^{1-\alpha} \\ \implies P_{iC} &= (1-\alpha)E_i\left(\frac{L_i}{N_i}\right)^{\frac{\alpha}{1-\alpha}}U_i^{-\frac{1}{1-\alpha}} \end{aligned} \tag{19}$$

# C Uniqueness of Equilibrium

The uniqueness of an equilibrium is not necessary to obtain cross-city patterns that are consistent with empirical facts as I explain in the main text. Nevertheless, it can be obtained with additional assumptions. I start with the case of gross complements. Proposition 10 requires an additional assumption.

**Assumption 3.** $(1-\alpha)/(\sigma-1) > \alpha$

**Proposition 10** (Uniqueness of equilibrium with gross complements)**.** *Given Assumptions 1 and 3, if the sectors are gross complements ($\eta < 1$), then there exists a unique equilibrium*

*Proof.* see Appendix D. □

Proposition 10 concerns only gross complements, and this is related to the endogenous tradable sector shares. When a city attracts new workers, the mass of locally produced varieties increase both in the non-tradable sectors and in the tradable sectors. However, the effects on the price index are different. In the tradable sectors, the mass of varieties that were produced in and shipped from the other city decreases. Also, the share of local varieties in the price index is smaller than 1. For these reasons, the price index reduction is relatively smaller in the tradable sectors. When the sectors are gross complements, this generates a force that increases the tradable sector shares in the city that attracts new workers. This is captured by the first term of the RHS in eq. (8) and (9). Larger population $N_1$, holding the expenditure shares constant, increases the relative wage $\omega$ by the home market effect (eq. (8)), which in turn increases the first term of eq. 9 only for the tradable sectors.

The increasing tradable sector shares works as a dispersion force. When residents spend more on the tradable sectors, workers move from the non-tradable sectors to the tradable sectors as Corollary 1. This reduces the price index of the city's tradable sectors by producing additional varieties, whereas it raises that of the non-tradable sectors by losing varieties. However, there is another effect. The new varieties in the tradable sectors become available in the other city, and it reduces the price index of the tradable sectors without costing locally-produced varieties in the non-tradable sectors. By this mechanism, the increase of the tradable sector shares decreases the relative real consumption of the city ($U_i/U_j$). Having an additional dispersion force only strengthens the aggregate dispersion force that is already stronger than the agglomeration force by Assumption 1. This is the reason why the uniqueness can be obtained with a relatively weak additional assumption. Assumption 3 ensures that the agglomeration force is stronger than the dispersion force by inelastic land supply, and it is a sufficient condition to obtain a relative wage that increases in population, which provides tractability.

When the sectors are gross substitutes, tradable sector shares works as an agglomeration force. Given $\eta > 1$, expenditures and employments are reallocated to non-tradable sectors when a city attracts workers because they become relatively inexpensive. This generates a force that raises the relative real consumption. With heterogeneous income-elasticities, it is not feasible to prove that this force does not reverse the relationship, globally, between the agglomeration force and the dispersion force in Assumption 1 globally. Hence, I impose Assumption 4 to obtain the uniqueness.

**Assumption 4.** $|\mathcal{N}| = |\mathcal{T}|$ *, and* $\forall k \in \mathcal{T}$*,* $\exists \ell \in \mathcal{N}$ *such that* $\epsilon_l = \epsilon_k$*,* $\beta_\ell = \kappa_1 \beta_k$*,* $\phi_\ell = \kappa_2 \phi_k$*, and* $\psi_\ell = \kappa_3 \psi_k$ *where* $0 \leq \kappa_1$ *and* $0 < \kappa_2, \kappa_3$

**Proposition 11** (Uniqueness of equilibrium with symmetric sectors)**.** *Given Assumptions 1, 4, there exists a unique equilibrium*

*Proof.* see Appendix D. □

In Proposition 11, $\kappa_1$ can be 0, and, therefore, it includes the case in which there is no non-tradable sector. Also, there is no condition on $\eta$ in Assumption 4. In other words, sectors can be either gross substitutes or gross complements. Finally, Proposition 11 does not use Assumption 3. This result can be obtained because the aggregate tradable sector share becomes tractable with Assumption 4.

# D Proof of Propositions and Lemma

### Proof of Proposition 1

I begin by substituting the optimized production into the zero-profit condition. As for the optimized production, the optimized prices are as follows.

$$\forall k \in \mathcal{K},\ p_{ij,k} = p_{ik} = \frac{\sigma}{\sigma - 1} \frac{\psi_k}{\lambda_i} w_i$$
$$\forall k \in \mathcal{T},\ p_{ij,k} = \tau p_{ik} = \frac{\sigma}{\sigma - 1} \tau \frac{\psi_k}{\lambda_i} w_i$$

Then, the zero-profit condition implies $\pi_{i,k} = 0$ for all $k$ in $\mathcal{K}$ and $i \in \{1,2\}$. It follows for all $(i,j)$ in $\{(1,2),(2,1)\}$,

$$\forall k \in \mathcal{N},\ q_{ii,k}\frac{\psi_k}{\lambda_i}w_i\left[\frac{1}{\sigma-1}\right] - \frac{\phi_k}{\lambda_i}w_i = 0$$
$$\forall k \in \mathcal{T},\ \left(q_{ii,k} + q_{ij,k}\tau\right)\frac{\psi_k}{\lambda_i}w_i\left[\frac{1}{\sigma-1}\right] - \frac{\phi_k}{\lambda_i}w_i = 0$$

The total labor demand by a firm producing variety $\nu$ in sector $k$ in city $i$, $N_{ik}(\nu)$, is pinned down as

$$N_{ik}(\nu) = \begin{cases} q_{ii,k}\frac{\psi_k}{\lambda_i} + \frac{\phi_k}{\lambda_i} = \sigma\frac{\phi_k}{\lambda_i} & k \in \mathcal{N} \\ \left(q_{ii,k} + q_{ij,k}\tau\right)\frac{\psi_k}{\lambda_i} + \frac{\phi_k}{\lambda_i} = \sigma\frac{\phi_k}{\lambda_i} & k \in \mathcal{T} \end{cases} \tag{20}$$

The labor demand is determined by the fixed cost and the productivity levels. Now, I use a normalization. It can be shown that $\beta_k$, $\phi_k$ and $\psi_k$ affect the equilibrium values of $N_1, N_2, w_1, w_2, C_1, C_2, H_1, H_2, U_1$, and $U_2$ only through $\beta_k^{\frac{1}{1-\eta}}\phi_k^{\frac{1}{\sigma-1}}\psi_k$. Therefore, given any set of parameters, replacing $\{\beta_k, \phi_k, \psi_k\}_{k\in\mathcal{K}}$ by $\left\{\tilde{\beta}_k, 1/\sigma, (\sigma-1)/\sigma\right\}_{k\in\mathcal{K}}$ where $\tilde{\beta}_k = \left[\beta_k^{\frac{1}{1-\eta}}\phi_k^{\frac{1}{\sigma-1}}\psi_k / \left\{(1/\sigma)^{\frac{1}{\sigma-1}}(\sigma-1)/\sigma\right\}\right]^{1-\eta}$ does not affect the equilibrium values of those variables. One caveat is that the price index is affected by this change. Let $\tilde{P}_k$ be the new price index given $\left\{\tilde{\beta}_k, 1/\sigma, (\sigma-1)/\sigma\right\}_{k\in\mathcal{K}}$. Then, $P_k = (1/\sigma)^{\frac{1}{\sigma-1}}\{\sigma/(\sigma-1)\}\tilde{P}_k$. Following Matsuyama (2019), I set $\psi_k = (\sigma-1)/\sigma$ and $\phi_k = 1/\sigma$ so that $p_{ik} = w_i/\lambda_i$ for all $k \in \mathcal{K}$ and $N_{i,k}(\nu) = 1/\lambda_i$, which requires that $\beta_k$ is replaced by $\tilde{\beta}_k = \left[\beta_k^{\frac{1}{1-\eta}}\phi_k^{\frac{1}{\sigma-1}}\psi_k / \left\{(1/\sigma)^{\frac{1}{\sigma-1}}(\sigma-1)/\sigma\right\}\right]^{1-\eta}$. Then, it follows from eq. (20) and the normalization that the aggregate supply of goods by a firm is given by

$$\forall k \in \mathcal{N},\ q_{ii,k} = 1$$
$$\forall k \in \mathcal{T},\ q_{ij,k} + q_{ij,k}\tau = 1$$

Next, to equate demand to supply, I derive the aggregate demand for a variety in sector $k$ in city $i$. I let $D_{ck}$ denote the aggregate demand, and it is as follows:

$$D_{ik} = p_k^{-\sigma} A_{ik}$$

where

$$A_{ik} = N_i \tilde{P}_{ik}^{\sigma} Q_{ik} + \rho_k N_j \tilde{P}_{jk}^{\sigma} Q_{jk}, \; and \; \rho_k = \begin{cases} 0 & k \in \mathcal{N} \\ \rho = \tau^{1-\sigma} & k \in \mathcal{T} \end{cases}$$

Equating demand ($D_{ik}$) and supply ($q_{ii,k}$ for $k \in \mathcal{N}$ and $q_{ii,k} + q_{ij,k}\tau$ for $k \in \mathcal{T}$) with $p_{ik} = w_i/\lambda_i$ gives, $\forall i \in \{1, 2\}, \forall k \in \mathcal{K}$

$$1 = \left(\frac{w_i}{\lambda_i}\right)^{-\sigma} A_{ik} \tag{21}$$

This equates supply and demand, and it reflects the zero-profit condition. This illustrates that, given a sector ($k$) and productivity levels ($\lambda_i$) being equal, the city with greater aggregate demand has a higher wage. Also, given a city ($i$), this equates the aggregate demands across sectors. To make use of eq. (21), I use two different expressions of the demand function to substitute for the $Q_{ik}$ that is contained in $A_{ik}$.

$$Q_{ik} = \begin{cases} \tilde{P}_{ik}^{-1} E_{iC} m_{ik} \\ \tilde{\beta}_k \tilde{P}_{ik}^{-\eta} U_i^{\epsilon_k} C_i^{1-\eta} E_{iC}^{\eta} \end{cases} \tag{22}$$

The first follows from eq. (3) and the second from eq. (2). With the first expression, eq. (21) becomes

$$\left(\frac{w_i}{\lambda_i}\right)^{-\sigma} = N_i \tilde{P}_{ik}^{\sigma-1} E_{iC} m_{ik} + \rho_k N_j \tilde{P}_{jk}^{\sigma-1} E_{iC} m_{jk} \tag{23}$$

Proposition 1 follows from eq. (23) for a non-tradable sector $k \in \mathcal{N}$

$$\begin{aligned}\left(\frac{w_i}{\lambda_i}\right)^{\sigma} &= N_i \tilde{P}_{ik}^{-1} E_{iC} m_{ik} \\ &= \frac{N_i w_i m_{ik}}{\lambda_i^{\sigma} x_{ik} N_i w_i^{1-\sigma}} \\ \implies x_{1k} &= m_{1k}\end{aligned} \tag{24}$$

where I use $E_{iC} = w_i$ and the price index $\tilde{P}_{ik}^{1-\sigma} = \lambda_i^{\sigma} x_{ik} N_i w_i^{1-\sigma}$ for $k \in \mathcal{N}$ where $x_{ik}$ is the employment share in sector $k$ in city $i$.

## Proof of Corollary 1

This follows from eq. (24) as follows:

$$\begin{aligned}&\sum_{k\in\mathcal{N}} x_{ik} = \sum_{k\in\mathcal{N}} m_{ik} \iff 1 - \sum_{k\in\mathcal{T}} x_{ik} = 1 - \sum_{k\in\mathcal{T}} m_{ik} \\ &\iff \sum_{k\in\mathcal{T}} x_{ik} = \sum_{k\in\mathcal{T}} m_{ik}\end{aligned} \tag{25}$$

## Proof of Proposition 2

This follows from zero-profit conditions for a tradable sector $k \in \mathcal{T}$ (eq. (23)) for the two cities that

$$\begin{aligned}\frac{(w_i/\lambda_i)^{\sigma} - \rho\,(w_j/\lambda_j)^{\sigma}}{1-\rho^2} &= \tilde{P}_{ik}^{\sigma-1} N_i E_{iC} m_{ik} \\ &= \frac{N_i w_i m_{ik}}{\lambda_i^{\sigma} x_{ik} N_i w_i^{1-\sigma} + \rho \lambda_j^{\sigma} x_{jk} N_j w_j^{1-\sigma}}\end{aligned} \tag{26}$$

where I use the price index $\tilde{P}_{ik}^{1-\sigma} = \lambda_i^\sigma x_{ik} N_i w_i^{1-\sigma} + \rho \lambda_j^\sigma x_{jk} N_j w_j^{1-\sigma}$, $E_{iC} = (1-\alpha)E_i = w_i$, and $x_{ik}$ is the employment share in sector $k$ in city $i$. It follows that

$$\lambda_i^\sigma x_{ik} N_i w_i^{1-\sigma} + \rho \lambda_j^\sigma x_{jk} N_j w_j^{1-\sigma} = (1-\rho^2) \frac{N_i w_i m_{ik}}{(w_i/\lambda_i)^\sigma - \rho\, (w_j/\lambda_j)^\sigma} \tag{27}$$

For city 1, aggregate over tradable sectors and use the income (wage) ratio $\omega = E_1/E_2 = w_1/w_2$,

$$\sum_{k \in \mathcal{T}} x_{1k} \lambda^\sigma N_1 \omega^{1-\sigma} + \rho \sum_{k \in \mathcal{T}} x_{2k} N_2 = (1-\rho^2) \frac{N_1 \omega \sum_{k \in \mathcal{T}} m_{1k}}{\lambda^{-\sigma} \omega^\sigma - \rho}$$

Transforming this making use of eq. (25), the condition that characterizes the home market effect on wage is obtained.

$$\frac{\sum_{k \in \mathcal{T}} m_{1k} N_1}{\sum_{k \in \mathcal{T}} m_{2k} N_2} = \omega^{2\sigma-1} \left[ \frac{\lambda^{-\sigma} - \rho \omega^{-\sigma}}{\lambda^\sigma - \rho \omega^\sigma} \right]$$

### Proof of Lemma 1

I use the second form of the demand function (22). eq. (21) becomes

$$\left( \frac{w_i}{\lambda_i} \right)^\sigma = N_i \tilde{\beta}_k \tilde{P}_{ik}^{\sigma-\eta} U_i^{\epsilon_k} C_i^{1-\eta} E_{iC}^\eta + N_j \rho_k \tilde{\beta}_k \tilde{P}_{jk}^{\sigma-\eta} U_j^{\epsilon_k} C_j^{1-\eta} E_{jC}^\eta \tag{28}$$

It follows from eq. 28 of sector $k \in \mathcal{K}$ for the two cities (eq. (28)),

$$\frac{(w_i/\lambda_i)^\sigma - \rho_k\, (w_j/\lambda_j)^\sigma}{1-\rho_k^2} = N_i \tilde{\beta}_k \tilde{P}_{ik}^{\sigma-\eta} U_i^{\epsilon_k} C_i^{1-\eta} E_{iC}^\eta \tag{29}$$

I eliminate $\tilde{P}_{ik}$ and $C_i$. First, from eq. (3) and (19),

$$
\begin{aligned}
\tilde{P}_{ik}^{1-\eta} &= \frac{m_{ik} \sum_{\ell \in \mathcal{K}} \beta_l U_i^{\epsilon_k} \tilde{P}_{i\ell}^{-\eta}}{\beta_k U_i^{\epsilon_k}} \\
&= \frac{m_{ik} \left[ (1-\alpha) E_i \left( L_i / N_i \right)^{\frac{\alpha}{1-\alpha}} U_i^{-\frac{1}{1-\alpha}} \right]^{1-\eta}}{\beta_k U_c^{\epsilon_k}} \\
\iff \tilde{P}_{ik} &= \left( \frac{m_{ik}}{\beta_k U_i^{\epsilon_k}} \right)^{\frac{1}{1-\eta}} (1-\alpha) E_i \left( \frac{L_i}{N_i} \right)^{\frac{\alpha}{1-\alpha}} U_i^{-\frac{1}{1-\alpha}}
\end{aligned}
$$

Next, I derive $\tilde{P}_{iC}$ by aggregating $\tilde{P}_{ik}$ to obtain $C_i$.

$$
\begin{aligned}
\tilde{P}_{iC} &= \left( \sum_{k \in \mathcal{K}} \beta_k U_i^{\epsilon_k} \tilde{P}_{ik}^{1-\eta} \right)^{\frac{1}{1-\eta}} \\
&= \left( \sum_{k \in \mathcal{K}} \beta_k U_i^{\epsilon_k} \left( \frac{m_{i,k}}{\beta_k U_i^{\epsilon_k}} \right) \left[ (1-\alpha) E_i \left( \frac{L_i}{N_i} \right)^{\frac{\alpha}{1-\alpha}} U_i^{-\frac{1}{1-\alpha}} \right]^{1-\eta} \right)^{\frac{1}{1-\eta}} \\
&= (1-\alpha) E_i \left( \frac{L_i}{N_i} \right)^{\frac{\alpha}{1-\alpha}} U_i^{-\frac{1}{1-\alpha}} \\
\iff C_i &= U_i^{\frac{1}{1-\alpha}} \left( \frac{L_i}{N_i} \right)^{-\frac{\alpha}{1-\alpha}}
\end{aligned}
$$

By plugging these into eq. (29),

$$
\begin{aligned}
\frac{(w_i/\lambda_i)^\sigma - \rho_k\,(w_j/\lambda_j)^\sigma}{1-\rho_k^2} &= N_i \tilde{\beta}_k U_i^{\epsilon_k}\left[\left(\frac{m_{i,k}}{\beta_k U_c^{\epsilon_k}}\right)^{\frac{1}{1-\eta}}(1-\alpha)E_i\left(\frac{L_i}{N_i}\right)^{\frac{\alpha}{1-\alpha}}U_c^{-\frac{1}{1-\alpha}}\right]^{\sigma-\eta} \\
&\quad \cdot \left[U_i^{\frac{1}{1-\alpha}}\left(\frac{L_i}{N_i}\right)^{-\frac{\alpha}{1-\alpha}}\right]^{1-\eta} E_{i,C}^{\eta} \\
&= N_i \tilde{\beta}_k U_i^{\epsilon_k}\left[\left(\frac{m_{i,k}}{\beta_k U_i^{\epsilon_k}}\right)^{\frac{1}{1-\eta}}\right]^{\sigma-\eta}\left[U_i^{\frac{1}{1-\alpha}}\left(\frac{L_i}{N_i}\right)^{-\frac{\alpha}{1-\alpha}}\right]^{1-\sigma}[(1-\alpha)E_i]^{\sigma} \\
&= N_i\left(\tilde{\beta}_k U_i^{\epsilon_k}\right)^{\frac{1-\sigma}{1-\eta}} m_{ik}^{\frac{\sigma-\eta}{1-\eta}}\left[U_i^{\frac{1}{1-\alpha}}\left(\frac{L_i}{N_i}\right)^{-\frac{\alpha}{1-\alpha}}\right]^{1-\sigma} w_i^{\sigma} \\
m_{ik}^{\frac{\sigma-\eta}{1-\eta}} &= \frac{\lambda_i^{-\sigma} - \rho_k \lambda_j^{-\sigma}\,(w_j/w_i)^\sigma}{(1-\rho_k^2)N_i}\left(\tilde{\beta}_k U_i^{\epsilon_k}\right)^{\frac{\sigma-1}{1-\eta}}\left[U_i^{\frac{1}{1-\alpha}}\left(\frac{L_i}{N_i}\right)^{-\frac{\alpha}{1-\alpha}}\right]^{\sigma-1} \\
m_{ik} &= \left[\frac{\lambda_i^{-\sigma} - \rho_k \lambda_j^{-\sigma}\,(w_j/w_i)^\sigma}{(1-\rho_k^2)N_i}\right]^{\frac{1-\eta}{\sigma-\eta}}\left(\tilde{\beta}_k U_i^{\epsilon_k}\right)^{\frac{\sigma-1}{\sigma-\eta}}\left[U_i^{\frac{1}{1-\alpha}}\left(\frac{N_i}{L_i}\right)^{\frac{\alpha}{1-\alpha}}\right]^{(\sigma-1)\frac{1-\eta}{\sigma-\eta}}
\end{aligned}
$$

## Proof of Proposition 9

For a tradable sector in city 1, from (27),

$$x_{1k}\lambda^\sigma N_1 + \rho x_{2k} N_2 \omega^{\sigma-1} = (1-\rho^2)\frac{N_1 m_{1k}}{\lambda^{-\sigma} - \rho\omega^{-\sigma}} \tag{30}$$

The counterpart of this with $m_{2k}$ can be obtained as

$$\rho x_{1k}\lambda^\sigma N_1 + x_{2k} N_2 \omega^{\sigma-1} = (1-\rho^2)\frac{N_2 m_{2k}}{\omega^{-\sigma} - \rho\lambda^{-\sigma}}\omega^{-1} \tag{31}$$

Solve (30) and (31) for $x_{1k}$ and $x_{2k}$ using 8.

$$\frac{x_{1k}}{\sum_{k\in\mathcal{T}} x_{1k}} = \frac{\left(m_{1k}/\sum_{k\in\mathcal{T}} m_{1k}\right) - \rho\lambda^{\sigma}\omega^{-\sigma}\left(m_{2k}/\sum_{k\in\mathcal{T}} m_{2k}\right)}{1-\rho\lambda^{\sigma}\omega^{-\sigma}}$$
$$\frac{x_{2k}}{\sum_{k\in\mathcal{T}} x_{2k}} = \frac{\left(m_{2k}/\sum_{k\in\mathcal{T}} m_{2k}\right) - \rho\omega^{\sigma}\lambda^{-\sigma}\left(m_{1k}/\sum_{k\in\mathcal{T}} m_{1k}\right)}{1-\rho\omega^{\sigma}\lambda^{-\sigma}}$$

The employment share ratio immediately follows from this.

## Proof of Proposition 4 and Proposition 5

I introduce new variables, $V_1$ and $V_2$, defined as

$$V_1 = a_1 U_1 N_1^{-\gamma}$$

$$V_2 = a_2 U_2 N_2^{-\gamma}$$

Using them, the equilibrium conditions can be rewritten as follows:

$$N_1^{\left(\frac{1}{1-\sigma}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} = \lambda^{-\sigma\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu} \left(\frac{V_1}{a_1}N_1^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} + \left[\frac{\lambda^{-\sigma}-\rho\omega^{-\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(\frac{V_1}{a_1}N_1^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \tag{32}$$

$$N_2^{\left(\frac{1}{1-\sigma}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} = \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu} \left(\frac{V_2}{a_2}N_2^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} + \left[\frac{1-\rho\lambda^{-\sigma}\omega^{\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(\frac{V_2}{a_2}N_2^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \tag{33}$$

$$\left(\frac{N_2}{N_1}\right)^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} = \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V_1 N_1^{\gamma}/a_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V_2 N_2^{\gamma}/a_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}} \lambda^{\sigma} \left[\frac{1-\rho\lambda^{-\sigma}\omega^{\sigma}}{\lambda^{-\sigma}-\rho\omega^{-\sigma}}\right]^{\mu} \omega^{1-2\sigma} \tag{34}$$

$$V_1 = V_2$$

$$N = N_1 + N_2$$

I prove the existence of an equilibrium by showing that $V_1$ and $V_2$ can be expressed as continuous functions of $N_1$ and that they have an intersection by the intermediate value theorem.

**(i) $V_1$ and $V_2$ can be expressed as functions of $N_1$** First, given the assumption for global monotonicity of $C_i$, $1+(1-\alpha)\min\{\epsilon_k\}/1-\eta > 0$, it follows from eq. (32) and (33) that

$$\frac{\partial V_1(N_1,\omega)}{\partial\omega} < 0, \frac{\partial V_2(N_2,\omega)}{\partial\omega} > 0$$

Then, notice that, given $N_1, N_2$, the RHS of eq. (34) decreases in $\omega$, taking into account $\partial V_1(\omega, N_1)/\partial\omega$ and $\partial V_2(\omega, N_2)/\partial\omega$. Also, given $N_1, N_2$,

$$\lim_{\omega\to\rho^{\frac{1}{\sigma}}\lambda} \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V_1 N_1^{\gamma}/a_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V_2 N_2^{\gamma}/a_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}} \lambda^{\sigma} \left[\frac{1-\rho\lambda^{-\sigma}\omega^{\sigma}}{\lambda^{-\sigma}-\rho\omega^{-\sigma}}\right]^{\mu} \omega^{1-2\sigma} \to +\infty \tag{35}$$

$$\lim_{\omega\to\rho^{-\frac{1}{\sigma}}\lambda} \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V_1 N_1^{\gamma}/a_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V_2 N_2^{\gamma}/a_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}} \lambda^{\sigma} \left[\frac{1-\rho\lambda^{-\sigma}\omega^{\sigma}}{\lambda^{-\sigma}-\rho\omega^{-\sigma}}\right]^{\mu} \omega^{1-2\sigma} \to 0 \tag{36}$$

Therefore, by the intermediate value theorem, given $(N_1, N_2)$, it uniquely pins down $\omega_1$ and, consequently, $V_1$ and $V_2$ from eq. (32) and (33), respectively. With the workers clearing condition, $V_1$ and $V_2$ are functions of $N_1$.

**(ii) $V_1$ and $V_2$ are continuous in $N_1$** Given $N_1 \in (0, N)$ (and therefore $N_2$), eq. (32) and (33) imply that $V_1 N_1^{\gamma}$ and $V_2 N_2^{\gamma}$ are continuous in $\omega$ on $(\rho^{\frac{1}{\sigma}}\lambda, \rho^{-\frac{1}{\sigma}})$, decrease and increase in $\omega$, respectively, and are nonzero. Then, given $N_1 \in (0, N)$, the RHS of eq. (34) is a continuous decreasing function of $\omega$. Combined with eq. (35) and (36), it follows that for all $N_1 \in (0, N)$ there exists $\omega \in (\rho^{\frac{1}{\sigma}}\lambda, \rho^{-\frac{1}{\sigma}}\lambda)$. Also, $\lim_{N_1\to x+} \omega = \lim_{N_1\to x-} \omega$ for all $x \in (0, N)$. This implies that $\omega$ is continuous on $N_1 \in (0, N)$. It immediately follows from eq. (32) and (33) that $V_1$ and $V_2$ are continuous in $N_1 \in (0, N)$.

**(iii) $V_1 > V_2$ when $N_1 \to 0$ and $V_1 < V_2$ when $N_1 \to N$ and intersection exits** Transform eq. (32) and (33),

$$\begin{aligned} 1 = &\sum_{k\in\mathcal{N}} \lambda^{-\sigma\frac{1-\eta}{\sigma-\eta}} \tilde{\beta}_k^{\mu} \left(\frac{V_1}{a_1}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{1-\sigma}\right](1-\eta)\mu} \\ &+ \sum_{k\in\mathcal{T}} \left[\frac{\lambda^{-\sigma}-\rho\omega^{-\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \tilde{\beta}_k^{\mu} \left(\frac{V_1}{a_1}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{1-\sigma}\right](1-\eta)\mu} \end{aligned} \tag{37}$$

$$
\begin{aligned}
1 = &\sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu} \left(\frac{V_2}{a_2}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{1-\sigma}\right](1-\eta)\mu} \\
&+\sum_{k\in\mathcal{T}} \left[\frac{1-\rho\lambda^{-\sigma}\omega^{\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \tilde{\beta}_k^{\mu} \left(\frac{V_2}{a_2}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{1-\sigma}\right](1-\eta)\mu}
\end{aligned}
\tag{38}
$$

Notice $0 \le (\lambda^{-\sigma}-\rho\omega^{-\sigma})/(1-\rho^2),\ (1-\rho\lambda^{-\sigma}\omega^{\sigma})/(1-\rho^2) < \infty$. and, for all $k \in \mathcal{K}$, $0 < \tilde{\beta}_k$. Given Assumption 1, $\gamma\left(\epsilon_k/(1-\eta)+1/(1-\alpha)\right)+\alpha/(1-\alpha)-1/(1-\sigma) > 0$. Also, remember $\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha} > 0$ for all $k \in \mathcal{K}$. Thus, eq. (37) implies $\lim_{N_1\to 0} V_1 = \infty$, and eq. (38) implies $\lim_{N_1\to 0} V_2 < \infty$. The same can be done when $N_1 \to N$ or $N_2 \to 0$.

Since $V_1$ and $V_2$ are continuous functions of $N_1$, respectively, there exists an intersection at $N_1 \in (0, N)$ by the intermediate value theorem. At the intersection, it satisfies the condition of a stable equilibrium as

$$
\begin{aligned}
&\frac{dV_1}{dN_1} < \frac{dV_2}{dN_1} \implies \frac{d\ln V_1}{dN_1} < \frac{d\ln V_2}{dN_1}\ (\because V_1 = V_2) \\
\iff &\frac{d\ln\left(U_1(N_1)a_1/U_2(N_1)a_2\right)}{dN_1} < \frac{d\ln\left(N_1/(N-N_1)\right)^{\gamma}}{dN_1} \\
\implies &\frac{d\left(U_1(N_1)a_1/U_2(N_1)a_2\right)^{1/\gamma}}{dN_1} < \frac{d\left(N_1/(N-N_1)\right)}{dN_1}\ (\because U_1(N_1)a_1N_1^{-\gamma} = U_2(N_1)a_2N_2^{-\gamma})
\end{aligned}
$$

## Proof of Proposition 10

I prove this in the following two steps.

**(i) $\omega$ increases in $N_1$** Given Assumption 3 and $\eta < 1$, the LHS of eq. (32) and (33) strictly increase and strictly decreases in $N_1$, respectively. Suppose, for $\exists x \in [0, N]$, $\omega$ weakly decreases in $N_1$ at $N_1 = n$. Then, at $n$, $V_1N_1^{\gamma}$ must strictly increase and $V_2N_2^{\gamma}$ strictly decreases in $N_1$ to satisfy eq. (32) and (33), respectively. Then, $\omega$ needs to strictly increase in $N_1$ to satisfy eq. (34). This contradicts that $\omega$ weakly decreases in $N_1$. Therefore, $\omega$ strictly increases in $N_1$ ($d\omega/dN_1 > 0$) for all $N_1 \in [0.N]$.

**(ii) $V_1$ decreases and $V_2$ increases in $N_1$ and unique intersection** Given Assumption 1, it follows that

$$\frac{\partial V_1(N_1,\omega)}{\partial N_1} < 0, \ \frac{\partial V_2(N_2,\omega)}{\partial N_2} < 0$$

Combining the results so far, the signs of the total derivatives can be obtained.

$$\begin{aligned}
\frac{dV_1(N_1,\omega)}{dN_1} &= \frac{\partial V_1(N_1,\omega)}{\partial N_1} + \frac{\partial V_1(N_1,\omega)}{\partial \omega_1}\frac{d\omega}{dN_1} < 0 \\
\frac{dV_2(N_2,\omega)}{dN_1} &= \frac{\partial V_2(N_2,\omega)}{\partial N_2}\frac{dN_2}{dN_1} + \frac{\partial V_2(N_2,\omega)}{\partial \omega}\frac{d\omega}{dN_1} > 0
\end{aligned}$$

Since $V_1$ and $V_2$ are monotonically decreasing and increasing continuous functions of $N_1$, respectively, the intersection is unique in the proof of Proposition 4.

## Proof of Proposition 11

It follows from eq. (9) and (10)

$$\begin{aligned}
\sum_{k\in\mathcal{T}} m_{1k} &= \frac{\left\{(\lambda^{-\sigma} - \rho\omega^{-\sigma})/(1-\rho^2)\right\}^{\frac{1-\eta}{\sigma-\eta}}}{\kappa^{\mu}\lambda^{-\sigma\frac{1-\eta}{\sigma-\eta}} + \left\{(\lambda^{-\sigma} - \rho\omega^{-\sigma})/(1-\rho^2)\right\}^{\frac{1-\eta}{\sigma-\eta}}} \\
\sum_{k\in\mathcal{T}} m_{2k} &= \frac{\left[(1-\rho\lambda^{-\sigma}\omega^{\sigma})/(1-\rho^2)\right]^{\frac{1-\eta}{\sigma-\eta}}}{\kappa^{\mu} + \left[(1-\rho\lambda^{-\sigma}\omega^{\sigma})/(1-\rho^2)\right]^{\frac{1-\eta}{\sigma-\eta}}}
\end{aligned}$$

where $\kappa = \kappa_1\kappa_2^{\frac{1-\eta}{\sigma-1}}\kappa_3^{1-\eta}$. Substituting these for eq. (8) gives:

$$\frac{N_2}{N_1} = \frac{\kappa^{\mu}\left(1-\rho\lambda^{-\sigma}\omega^{\sigma}\right)^{\mu} + (1-\rho^2)^{-\frac{1-\eta}{\sigma-\eta}}\left(1-\rho\lambda^{-\sigma}\omega^{\sigma}\right)}{\kappa^{\mu}\lambda^{-\sigma\frac{1-\eta}{\sigma-\eta}}\left(\lambda^{-\sigma}-\rho\omega^{-\sigma}\right)^{\mu} + (1-\rho^2)^{-\frac{1-\eta}{\sigma-\eta}}\left(\lambda^{-\sigma}-\rho\omega^{-\sigma}\right)}\lambda^{\sigma}\omega^{1-2\sigma}$$

This shows that $\omega$ increases in $N$. Equilibrium conditions can be rewritten as follows:

$$1 = \Phi_1(\omega) \sum_{k \in \mathcal{N}} \tilde{\beta}_k^\mu \left(\frac{V_1}{a_1}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{1-\sigma}\right](1-\eta)\mu} \tag{39}$$

$$1 = \Phi_2(\omega) \sum_{k \in \mathcal{N}} \tilde{\beta}_k^\mu \left(\frac{V_2}{a_2}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\frac{\alpha}{1-\alpha}-\frac{1}{1-\sigma}\right](1-\eta)\mu} \tag{40}$$

where $\Phi_1(\omega) \equiv \kappa^\mu \lambda^{-\sigma\frac{1-\eta}{\sigma-\eta}} + \left[(\lambda^{-\sigma} - \rho\omega^{-\sigma})/(1-\rho^2)\right]^{\frac{1-\eta}{\sigma-\eta}}$ and $\Phi_2(\omega) \equiv \kappa^\mu + [(1-\rho\lambda^{-\sigma}\omega^\sigma)/(1-\rho^2)]^{\frac{1-\eta}{\sigma-\eta}}$. Given $\omega$ increasing in $N_1$, the assumption of $1+(1-\alpha)\min\{\epsilon_k\}/(1-\eta) > 0$ for global monotonicity of $C_i$, and Assumption (1), eq. (39) and (40) imply $V_1$ and $V_2$ decreases and increases in $N_1$, respectively. The rest of the proof follows Proof of Proposition 4 and Proposition 5.

## Proof of Proposition of 6

I prove this by contradiction. Suppose $N_1 \leq N_2$ in an equilibrium. Then, eq. (37) and (38) and $V_1 = V_2 = V$ imply that it is necessary

$$\lambda^{-\sigma} - \rho\omega^{-\sigma} > 1 - \rho\lambda^{-\sigma}\omega^\sigma$$

It follows $\omega > \lambda$. Then, it follows from Equation 34

$$1 = \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^\mu \left(VN_1^\gamma\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^\mu \left(VN_2^\gamma\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}} \lambda^\sigma \left[\frac{1-\rho\lambda^{-\sigma}\omega^\sigma}{\lambda^{-\sigma}-\rho\omega^{-\sigma}}\right]^\mu \omega^{1-2\sigma} \tag{41}$$

$$< \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^\mu V^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)\right](1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^\mu V^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)\right](1-\eta)\mu}} \lambda^{-\sigma} \tag{42}$$

When $\eta < 1$, it is obvious that the RHS of the inequality is smaller than 1 when $N_1 \leq N_2$. Thus, a contradiction. When $\eta > 1$, given Assumption 2

$$\frac{\min_{k\in\mathcal{T}} \epsilon_k + 1}{1-\eta} + \frac{\gamma+\alpha}{1-\alpha} < 0$$

It follows that

$$\forall k \in \mathcal{T},\ \left[\gamma\left(\frac{\epsilon_k}{1-\eta} + \frac{1}{1-\alpha}\right) + \left(\frac{1}{1-\eta} + \frac{\alpha}{1-\alpha}\right)\right](1-\eta)\mu > 0$$

This implies that the RHS of inequality (42) is smaller than 1 when $N_1 \leq N_2$. Thus, a contradiction.

As for the wage level, think about eq. (41) in the equilibrium and evaluate the RHS with $\omega = \lambda$

$$\frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(VN_1^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(VN_2^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}} \lambda^{\eta\mu}$$

With either (i) $\eta < 1$ or (ii) $\eta > 1$ and Assumption 2, this is greater than 1 with $N_1 > N_2$. The RHS of eq. (41) decreases in $\omega$, and, therefore, $\omega > \lambda$ in the equilibrium.

# E Equilibrium with Asymmetric Amenity

In this appendix, I impose symmetric fundamental productivity ($\lambda = \lambda_1 = \lambda_2 = 1$) and show how the asymmetric amenity ($a_1 \neq a_2$) generates a cross-city difference. I let city 1 have a higher amenity ($a_1 > a_2$) without loss of generality. The same result on the population pattern can be obtained as the asymmetric productivity case.

**Proposition 12** (Cross-City Population Patterns with Asymmetric Amenity)**.** *Suppose that the* *amenities are asymmetric ($a_1 > a_2$), the productivities are symmetric ($\lambda_1 = \lambda_2$), and the sectors are either gross complements ($\eta < 1$), or gross substitutes ($\eta > 1$) with Assumption 2 holding. Then, given an equilibrium with Assumption 1, the city with better amenity is larger ($N_1 > N_2$).*

*Proof.* Suppose $N_1 \leq N_2$ in an equilibrium. Then, from eq. (37) and (38), it is necessary that

$$1 - \rho\omega^{-\sigma} > 1 - \rho\omega^{\sigma}$$

This implies $\omega > 1$. Then, it follows from eq. 34

$$1 = \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(VN_1^{\gamma}/a_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(VN_2^{\gamma}/a_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}} \left[\frac{1-\rho\omega^{\sigma}}{1-\rho\omega^{-\sigma}}\right]^{\mu} \omega^{1-2\sigma} \tag{43}$$

$$< \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V/a_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_1^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)\right](1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} \left(V/a_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} N_2^{\left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)\right](1-\eta)\mu}} \tag{44}$$

When $\eta < 1$, it is obvious that the RHS of eq. (44) is smaller than 1 when $N_1 \leq N_2$. Thus, a contradiction. When $\eta > 1$, given Assumption 2

$$\frac{\min_{k\in\mathcal{T}} \epsilon_k + 1}{1-\eta} + \frac{\gamma+\alpha}{1-\alpha} < 0$$

It follows that

$$\forall k \in \mathcal{T},\ \left[\gamma\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)+\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)\right](1-\eta)\mu > 0$$

Thus, the RHS of eq. (44) is smaller than 1. Thus, a contradiction. □

Next, the wage pattern requires an additional assumption, which is introduced in Appendix C for the equilibrium uniqueness.

**Proposition 13** (Cross-City Wage Patterns with Asymmetric Amenity)**.** *Suppose that the amenities are asymmetric ($a_1 > a_2$), the productivities are symmetric ($\lambda_1 = \lambda_2$), the agglomeration force is stronger than the dispersion force from inelastic land supply (Assumption 3), and the sectors are gross complements ($\eta < 1$). Then, given an equilibrium with Assumptions 1, the city with better amenity offers a higher wage ($w_1 > w_2$).*

*Proof.* The equilibrium conditions can be rewritten as

$$
\begin{aligned}
N_1^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} &= \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu}\left(\frac{V_1}{a_1}N_1^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \\
&\quad + \left[\frac{1-\rho\omega^{-\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu}\left(\frac{V_1}{a_1}N_1^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \\
N_2^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} &= \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu}\left(\frac{V_2}{a_2}N_2^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \\
&\quad + \left[\frac{1-\rho\omega^{\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu}\left(\frac{V_2}{a_2}N_2^{\gamma}\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \\
\left(\frac{N_2}{N_1}\right)^{\left(\frac{1}{1-\eta}+\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} &= \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu}\left(VN_1^{\gamma}/a_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu}\left(VN_2^{\gamma}/a_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}\left[\frac{1-\rho\omega^{\sigma}}{1-\rho\omega^{-\sigma}}\right]^{\mu}\omega^{1-2\sigma} \\
V_1 &= V_2 \\
N &= N_1 + N_2
\end{aligned}
$$

Suppose $\omega < 1$ at $N_1 = N_2 = N/2$. Given Assumption 3 and $\eta < 1$, the first and the second equations imply $V_1/a_1 > V_2/a_2$. However, the third implies $V_1/a_1 < V_2/a_2$. These contradict each other. So, when $N_1 = N_2 = N/2$, the first three equations imply $\omega \geq 1$. Given Assumption 3 and $\eta < 1$, $\omega$ increases in $N_1$ as Proof of Proposition 10, and, therefore, the wage at an equilibrium, where $N_1 > N/2$, is higher in city 1 than in city 2 ($\omega = w_1/w_2 > 1$) □

Assumption 3 ensures that the larger city has a greater goods consumption ($C_1$), which increases the relative wage through the home market effect.

Similarly, the expenditure pattern generally depends on which force is stronger, the agglomeration force or the dispersion force by inelastic land supply. First, I analyze the case without tradable sectors, which requires no additional strong assumption to obtain a clear result.

**Proposition 14** (Cross-City Expenditure Pattern with Asymmetric Amenity and without Tradable Sectors)**.** *Suppose that the amenities are asymmetric ($a_1 > a_2$), the productivities are symmetric ($\lambda_1 = \lambda_2$), there is no tradable sector ($\mathcal{K} = \mathcal{N}$), and either the sectors are gross complements ($\eta < 1$) or the sectors are gross substitutes ($\eta > 1$) with Assumption 2 holding. Then, given an equilibrium with Assumption 1, if Assumption 3 holds, the sectoral expenditure share ratio of city 1 to city 2 ($m_{1k}/m_{2k}$) increases in the income elasticity ($\epsilon_k$), and, it decreases, otherwise.*

*Proof.* First, Proposition 12 applies and $N_1 > N_2$. Second, the equilibrium conditions characterizing real consumption become

$$
\begin{aligned}
N_1^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} &= \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu} \left(U_1\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \\
N_2^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} &= \sum_{k\in\mathcal{N}} \tilde{\beta}_k^{\mu} \left(U_2\right)^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}
\end{aligned}
$$

Given $N_1 > N_2$, these immediately imply the results. □

The large city offers a relatively higher (lower) real consumption when the agglomeration is relatively stronger (weaker), and this relative real consumption determines the relative expenditure pattern.

When the model has both non-tradable and tradable sectors, it becomes difficult to obtain results because of the endogenous aggregate tradable sector share, which is related to the relative wage. In the rest of paper, I focus on the cases without non-tradable sectors, and provide two contrasting results in special cases.The first case is given by Proposition 15.

**Proposition 15** (Cross-City Expenditure Pattern with Asymmetric Amenity without Non-Tradable Sectors Case 1)**.** *Suppose that the amenities are asymmetric ($a_1 > a_2$), the productivities are symmetric ($\lambda_1 = \lambda_2$), there is no non-tradable sector ($K = \mathcal{T}$), and there is no housing expenditure ($\alpha = 0$). Then, given an equilibrium with Assumption 1, the sectoral expenditure share ratio of city 1 to city 2 ($m_{1k}/m_{2k}$) increases in the income elasticity ($\epsilon_k$).*

*Proof.* The three of equilibrium conditions are given by

$$N_1^{\frac{1-\eta}{\sigma-\eta}} = \left[\frac{1-\rho\omega^{-\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_1^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \tag{45}$$

$$N_2^{\frac{1-\eta}{\sigma-\eta}} = \left[\frac{1-\rho\omega^{\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_2^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \tag{46}$$

$$\frac{N_1}{N_2} = \omega^{2\sigma-1}\left[\frac{1-\rho\omega^{-\sigma}}{1-\rho\omega^{\sigma}}\right] \tag{47}$$

It follows from eq. (45) and (46) that

$$\frac{N_1^{\frac{1-\eta}{\sigma-\eta}}}{N_2^{\frac{1-\eta}{\sigma-\eta}}}\left[\frac{1-\rho\omega^{\sigma}}{1-\rho\omega^{-\sigma}}\right]^{\frac{1-\eta}{\sigma-\eta}} = \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_1^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_2^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}$$

With eq. (47),

$$\omega^{(2\sigma-1)\frac{1-\eta}{\sigma-\eta}} = \frac{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_1^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_2^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}$$

In the equilibrium, $\omega > 1$ as 13. It implies $U_1 > U_2$, and the expenditure share pattern immediately follows. □

The housing consumption is eliminated here to ensure that the real consumption ($U$) becomes higher with larger population. When the sectors are tradable, the benefit of agglomeration economy in the large city can be enjoyed also by residents in the small city, although partially, as they can access the rich varieties in the large city by paying the trade cost. Therefore, it is not possible to generally show that the real consumption becomes higher even with Assumption 3[14]. However, when the model does not have the dispersion force from inelastic land supply at all, the real consumption necessarily increases with population by the agglomeration force, and the city with better amenity offers a higher real consumption. Consequently, the workers in the city spend relatively more expenditure on income-elastic sectors.

On the other hand, when the dispersion force from inelastic land supply is stronger than the agglomeration force, the opposite result is obtained as Proposition 16.

**Proposition 16** (Cross-City Expenditure Pattern with Asymmetric Amenity Case 2)**.** *Suppose that the amenities are asymmetric ($a_1 > a_2$), the productivities are symmetric ($\lambda_1 = \lambda_2$), there is no non-tradable sector ($\mathcal{K} = \mathcal{T}$), the dispersion force from inelastic land supply is stronger than the agglomeration force ($(1-\alpha)/(\sigma-1) < \alpha$), and either the sectors are gross complements ($\eta < 1$) or the sectors are gross substitutes ($\eta > 1$) with Assumption 2 holding. Then, given an equilibrium with Assumption 1, the sectoral expenditure share ratio of city 1 to city 2 ($m_{1k}/m_{2k}$) decreases in the income elasticity ($\epsilon_k$).*

*Proof.* First, Proposition 12 applies and $N_1 > N_2$. Second, the three of equilibrium conditions are given by

$$N_1^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} = \left[\frac{1-\rho\omega^{-\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_1^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \tag{48}$$

$$N_2^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu} = \left[\frac{1-\rho\omega^{\sigma}}{1-\rho^2}\right]^{\frac{1-\eta}{\sigma-\eta}} \sum_{k\in\mathcal{T}} \tilde{\beta}_k^{\mu} U_2^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu} \tag{49}$$

$$\frac{N_1}{N_2} = \omega^{2\sigma-1}\left[\frac{1-\rho\omega^{-\sigma}}{1-\rho\omega^{\sigma}}\right] \tag{50}$$

[14] When the driver of the city size difference is a difference in productivities, it is possible as in the main text.

It follows from eq. 50 that $\omega > 1$. Finally, it follows from eq. (48) and (49) that

$$\left(\frac{N_1}{N_2}\right)^{\left(\frac{1}{\sigma-1}-\frac{\alpha}{1-\alpha}\right)(1-\eta)\mu}\left[\frac{1-\rho\omega^{\sigma}}{1-\rho\omega^{-\sigma}}\right]^{\frac{1-\eta}{\sigma-\eta}} = \frac{\sum_{k\in\mathcal{T}}\tilde{\beta}_k^{\mu}U_1^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}{\sum_{k\in\mathcal{T}}\tilde{\beta}_k^{\mu}U_2^{\left(\frac{\epsilon_k}{1-\eta}+\frac{1}{1-\alpha}\right)(1-\eta)\mu}}$$

Given$(1-\alpha)/(\sigma-1) \leq \alpha$ and $N_1 > N_2$, the LHS is smaller (greater) than 1 when $\eta < 1$ ($> 1$), which implies $U_1 \leq U_2$. The expenditure share pattern immediately follows. $\square$

In this case, a majority of workers choose to live in a city with better amenity, but they have a lower real consumption because of the high land price and the high expenditure share of land consumption. Consequently, the residents spend relatively more on income-inelastic sectors.

# F Alternative Specification

I test the association of the income level and the income-elasticities of the local industries with alternative identification. This test follows Nunn (2007), and the regression model is given by

$$\begin{aligned} y_{mk} =& \alpha \cdot \exp(\beta_{income} \cdot \epsilon_k \cdot \log(Population_m) + \beta_{skill} \cdot \theta_k \cdot \log(College_m) \\ &+ \sum_m \gamma_m D_m + \sum_{region} \gamma_{k,region} D_{k,region})) \cdot e_{mk} \end{aligned} \tag{51}$$

where

$$y_{mk}: \text{employment level of industry } k \text{ in MSA } m$$

$$\epsilon_k: \text{income elasticity of demand in industry } k$$

$$Population_m: \text{population in MSA } m$$

$$\theta_k: \text{skill intensity of goods in industry} k$$

$$College_m: \text{college employment ratio in labor force in MSA } m$$

$$D_m: \text{MSA dummy variable}$$

$$D_{k,region}: (\text{industry } k \times \text{region} \in \{Northeast, Midwest, South, West\}) \text{ dummy variable}$$

$$e_{mk}: \text{error term for MSA } m \times \text{industry } k$$

The coefficient of interest is $\beta_{income}$. When $\beta_{income}$ is positive, the employment level rises more for high $\epsilon_k$ as the population rises, which is consistent with the stylized fact and the model prediction. Two points are noteworthy. First, as the regression model shows, I implement level regressions by the Poisson Pseudo Maximum Likelihood (PPML) estimation. As discussed in Silva and Tenreyro (2006), log-linear regressions require a very specific condition on error terms to obtain consistent estimators. Moreover, in log-linear estimations, it is problematic when zeros are contained in the data. On the other hand, PPML provides consistent estimators that do not require this condition, and it is efficient with various error term patterns. For this reason, PPML is very common in gravity equation estimations in international trade where zeros are prevalent and error terms show heteroskedasiticy. In my dataset, 17% of the sample is zero. To address these zeros, I use PPML, and I set employment levels (instead of employment shares) as the dependent variable so that the error terms show a heteroskedasticity pattern that is suitable for PPML in terms of efficiency. Second, I control the supply side effect using the skill-intensity of sectors, as in the main empirical work. This time, the interaction term, $\theta_k \cdot \log(College_m)$, is used to implement the control, and the estimates of skill intensities are again borrowed from Caron et al. (2020), who obtained them through a structural estimation that had both heterogeneous skill-intensities and income-elasticities.

Table 4: Regression Result of Alternative Specification

| | employment (PPML) | | | log(employment+1) | | |
|---|---|---|---|---|---|---|
| | 2006 | 2011 | 2016 | 2006 | 2011 | 2016 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\beta_{income}$ | 0.294*** | 0.370*** | 0.320*** | 1.039*** | 1.037*** | 0.892*** |
| | (0.030) | (0.029) | (0.029) | (0.080) | (0.082) | (0.074) |
| $\beta_{skill}$ | 2.006*** | 2.320*** | 2.351*** | 1.459*** | 1.426*** | 1.793*** |
| | (0.152) | (0.143) | (0.153) | (0.379) | (0.363) | (0.384) |
| Observations | 8,382 | 8,415 | 7,986 | 8,382 | 8,415 | 7,986 |
| $R^2$ | | | | 0.855 | 0.858 | 0.860 |
| Adjusted $R^2$ | | | | 0.848 | 0.851 | 0.853 |

*Notes:* ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

The regression results are consistent with the stylized fact and the model prediction. Because this regression is cross-sectional, I implement it for three different years separately to make use of the annual data. The results are shown in Table 4. The results with PPML are shown in the left three columns, while those in the right three columns show log-linear regressions after 1 is added to every observation to take care of zeros. In all of the results, $\beta_{income}$ is significantly positive. The size of $\beta_{income}$ is between 0.888 and 1.121 in PPML. For example, to see the impact of this $\beta_{income}$ using 2006 data, suppose, first, that City A has a population that is greater than the sample mean in log point by one standard deviation in log point, which corresponds to a population of 1,197,504 in level, second, that City B has a population of sample mean in log point, which is 412,808, and, third, that City A and City B have the same college employment ratio. Then, the employment ratio of a sector with $\epsilon_k = 1.2$ over that with $\epsilon_k = 0.8$ ($N_{\epsilon_k=1.2}/N_{\epsilon_k=0.8}$) is 1.14 times greater in

City A than in City B.

The income-elasticities have a greater explanatory power than the skill supply. It is not straightforward to measure the explanatory power of the PPML estimators because technically they are obtained by a maximum likelihood estimation. To construct a measurement for the PPML estimators, I calculate the estimated employment shares implied by the fitted values for the employment levels and obtain the residual sum of squares (RSS) from the difference between the estimated shares and the actual data. Using this RSS, a measurement which analogous to $R^2$ in linear regressions is constructed and summarized in Table 5. In the first row, $RSS_{FE}$ is the $RSS$ when the regression has only the FE effects of MSA and industry $\times$ region and $TSS$ is the residual sum of squares by the unconditional mean of the employment shares. Table 5 shows that the FE effects explain over 90% of TSS in all of the three years, which is natural given the rich set of the fixed effects. In the second row, $RSS_{FE+skill}$ is obtained by the regression that has the skill interaction term as an additional control. Given an industry and a region, the skill supply effect explains 10% of the variation at most. Similarly, in the third row, $RSS_{FE+income}$ is obtained by the regression, here with the income interaction term instead of the skill interaction, and the income effect explains 7%-13% of the variation in all the three years. Finally, in the last row, $RSS_{FE+skill+income}$ is by the regression that has the same controls as (51), and the skill and income effects jointly capture 8%-20% of the variation. These results suggest that, given an industry and a region, the demand effect can explain a significant portion of the variation in the employment share, even after controlling the skill supply. Moreover, the explanatory power is not smaller than that of the skill supply .

Table 5: Explanatory Power

| | 2006 | 2011 | 2016 |
|---|---|---|---|
| $1 - \frac{RSS_{FE}}{TSS}$ | 0.93 | 0.91 | 0.90 |
| $1 - \frac{RSS_{FE+skill}}{RSS_{FE}}$ | 0.10 | 0.00 | 0.00 |
| $1 - \frac{RSS_{FE+income}}{RSS_{FE}}$ | 0.13 | 0.08 | 0.07 |
| $1 - \frac{RSS_{FE+skill+income}}{RSS_{FE}}$ | 0.20 | 0.06 | 0.08 |