RedFame

International Journal of Social Science Studies
Vol. 13, No. 1; March 2025
ISSN 2324-8033 E-ISSN 2324-8041
Published by Redfame Publishing
URL: http://ijsss.redfame.com

Skill Complementarities and Urban Labor Sorting: Evidence from Chinese Cities

Xu Guo[1], Shuhui Yuan[1], Jingyu Yu[1], Wenhao Sun[1]

[1] School of Maritime Economics and Management, Dalian Maritime University, China

Correspondence: Xu Guo, School of Maritime Economics and Management, Dalian Maritime University, China.

Received: January 13, 2025 Accepted: February 8, 2025 Available online: March 18, 2025

doi:10.11114/ijsss.v13i1.7459 URL: https://doi.org/10.11114/ijsss.v13i1.7459

Abstract

The spatial sorting of skilled labor is the key to the economic growth of cities, while skill complementarity serves as an important underlying mechanism for its formation. This study examines how skill complementarities in production influence the equilibrium allocation of labor across urban areas. We differentiate between extreme-skill complementarity, which leads to thicker-tailed skill distributions in large cities, and top-skill complementarity, which generates first-order stochastic dominance. Analyzing wage and housing price data from the 2013 CHIP survey and the CEIC database, we find robust evidence that large cities in China attract a disproportionate share of both high-skilled and low-skilled workers, while average skill levels remain unchanged across city sizes. These findings highlight the importance of providing valuable insights for developing countries to address urban development inequalities and enhance labor market efficiency.

Keywords: Labor skill, Extreme-skill complementarity, Top-skill complementarity, Thick tail

1. Introduction

The choice to reside in certain cities rather than fully autarkic locations stems from the localized aggregate increasing returns generated by the agglomeration of firms and labor. Among the key factors influencing productivity and the composition of workers within firms, complementarities play a central role, as highlighted by Marshall (1890). The presence of more productive co-workers enhances the performance of others, shaping decisions about where individuals choose to work and with whom. Marshall (1890) identified three primary sources of agglomeration economies: knowledge spillovers, linkages between input suppliers and final producers, and labor market interactions. To further conceptualize these mechanisms, Duranton & Puga (2004) proposed a taxonomy based on sharing, matching, and learning. Complementarities, akin to knowledge spillovers, influence the productivity of workers with varying skill levels, thereby generating increasing returns (Eeckhout, Pinheiro, & Schmidheiny, 2014).

The agglomeration of a heterogeneous labor force improves production efficiency. One line of reasoning stems from productivity spillovers, where a worker's productivity benefits from exposure to high-productivity co-workers. Cornelissen & Dustmann (2023); Bloom&Foster (2019) observed that, even when average productivity is held constant, teams with diverse skills are more productive, underscoring the complementarity between high- and low-productivity workers. Another perspective emphasizes how thick labor markets mitigate informational asymmetries and improve matching quality, fostering better partnerships. Recent studies have highlighted aggregate gains from positive assortative matching, where productivity increases when similar skill types are paired. In high-cost cities, high-skilled workers self-select into areas with greater affordability for their skill group, leading to improved matches and higher productivity (Shimer & Smith, 2000; Venables, 2011).

While productive complements often involve similar skill types, other combinations of skills can also be synergistic. Eeckhout et al. (2014) demonstrated the importance of extreme-skill complementarities, where high-skilled workers and providers of low-skilled services mutually enhance each other's productivity. Moreover, cities vary in many dimensions beyond size, including living costs, the frequency of idea exchange, peer effects, and skill complementarities. Davis & Dingel (2012) introduced a model in which costly idea exchange drives agglomeration, with higher-ability individuals self-selecting into larger cities where idea exchange is more frequent.

Most empirical evidence on labor sorting has been derived from industrialized countries, while transitional economies, where urbanization and industrialization lag behind, remain underexplored. This paper contributes to the literature by examining labor sorting patterns in China, the world's largest emerging economy. We propose a spatial sorting model with heterogeneously skilled workers to analyze equilibrium outcomes under extreme-skill and top-skill complementarities. Our key theoretical finding is that complementarities among workers with diverse skills shape the distribution of labor within cities.

This study also addresses the urban wage premium through the lens of house prices. While it is widely recognized that wages are higher in large cities globally, including in China, the underlying drivers—whether higher skills or higher living costs—remain unclear. By using house prices to adjust for living costs, we isolate the real wage differences across cities. Our empirical findings reveal that skill distributions in "super cities" exhibit thick tails compared to those in "big cities" and demonstrate first-order stochastic dominance (FOSD) over medium- and small-sized cities.

The remainder of this paper is structured as follows: Section 2 briefly reviews the relevant literature. Section 3 develops the spatial sorting model with heterogeneous skilled workers. Section 4 presents wage and skill distributions across cities. Section 5 discusses further results, including skill decomposition and the distribution of migrants versus stayers. Section 6 concludes with key findings and policy implications.

2. Literature Review

The urban wage premium is a central puzzle in urban economics. Numerous studies have documented a city-size wage premium, where wages in smaller cities exhibit first-order stochastic dominance (FOSD) over those in larger cities. This raises an important question: does the urban wage premium indicate that larger cities typically host workers with higher average skill levels than smaller cities? In fact, there is no clear-cut relationship between wages and skills across cities. While large cities indeed attract higher-ability workers, they also enhance productivity by accelerating human capital accumulation (Tariq & Ali, 2024). Consequently, it is challenging to disentangle the sources of higher wages—whether they arise from individual skills or city-size effects.

On one hand, thicker urban labor markets are associated with more assortative matching between workers and firms (Combes & Gobillon, 2015). Additionally, search behavior and matching efficiency tend to improve with market size and industry specialization (Luis & Pawel, 2024). On the other hand, large cities are not exclusively attractive to high-skilled workers. Public transportation, for instance, plays a crucial role in centralizing poverty and sorting lower-income populations into urban centers (Pucher & Buehler, 2012). Theoretical predictions in urban economics suggest that large cities may concentrate either high-skilled or low-skilled labor, depending on specific local factors.

Empirical studies provide mixed evidence regarding skill distributions in large cities. Some studies support the notion that skill distributions in large cities exhibit FOSD relative to smaller cities. For example, Roca (2017) used administrative data from Spain to show that migrants to large cities are positively selected in terms of education, occupational skills, and individual productivity. Similarly, Behrens et al. (2014) proposed a model of perfect sorting by talent, suggesting that large cities with significant wage premiums predominantly attract high-skilled workers, while low-skilled workers gravitate toward smaller cities with minimal wage premiums.

Other studies, however, have identified evidence of "thick tails" or broader dispersion in skill distributions in large cities. Bollinger & Hirsch (2021) measured worker intelligence using AFQT scores and found that large cities host both highly skilled and very low-skilled workers.

David (2018) argued that job search frictions lead to higher productivity losses in small regions, resulting in the self-selection of both high- and low-skilled workers into larger cities. Eeckhout et al. (2016) further reported that skill distributions in large U.S. cities exhibit thick tails, indicating that these cities disproportionately attract both the most skilled and the least skilled workers. They attributed this phenomenon to extreme-skill complementarity, a concept discussed in greater detail in Section 3.

The study of spatial sorting characteristics of labor fundamentally hinges on effectively accounting for worker heterogeneity and credibly measuring skills. Early contributions, such as Eeckhout (2004) and Davis (2011), employed multi-city models to analyze household location choices, demonstrating that equilibrium ensures equal utility for all agents, regardless of the metropolitan statistical area (MSA) they reside in. Building on this foundation, Eeckhout et al. (2014) developed a model incorporating heterogeneously skilled workers to explore spatial sorting patterns under different modes of skill complementarity. Regarding skill measurement, instead of relying on observables such as wages as proxies for skills, they argued that workers' location choices are influenced by both housing prices and earnings. They thus proposed measuring skill as the residual of wages adjusted for local housing costs.

To fully assess the characteristics of labor sorting, two primary challenges must be addressed. The first involves uncovering the theoretical mechanisms through which heterogeneous labor in cities derives productive benefits from agglomeration, thereby shaping specific patterns of spatial distribution. Eeckhout et al. (2014)considered two alternative hypotheses regarding skill complementarities and substitute abilities, focusing on workers' sorting decisions across the entire skill spectrum, including medium- and low-skilled workers. Their model derived implications for equilibrium skill distributions but did not account for city-level complementarities. In reality, agglomeration effects—such as knowledge and idea exchange—are often more pronounced in larger cities.

The second challenge lies in obtaining accurate estimates of worker skills. Early studies on urban wage premiums frequently used education level as a proxy for skills (Glaeser & Mare, 2001;Wheeler, 2001). Subsequently, some scholars defined skills as a composite of various characteristics, including intelligence, social abilities, and manual skills (Bacolod et al., 2009). More recently, skills have been modeled as a linear function of wages (Gautier & Teulings, 2009). While this method is intuitive, it neglects the impact of living costs, which vary across cities. For instance, workers with identical skill levels may earn different wages due to differences in local living expenses. Following the approach of Eeckhout et al. (2014) , this study uses housing prices as a proxy for living costs to estimate worker skills. This approach is particularly relevant in China, where housing expenditures constitute a significant portion of household consumption.

In this paper, we incorporate city-level heterogeneous skill complementarities into a spatial sorting model with heterogeneous labor to empirically investigate labor sorting patterns across Chinese cities of varying sizes. This study makes several contributions to the literature:

Theoretical Contribution: We extend the model by Eeckhout et al. (2014) by integrating city-level variations in skill complementarities and examining equilibrium labor sorting under both top-skill and extreme-skill complementarities.

Methodological Contribution: We quantify skill levels across 126 Chinese cities using an indirect measure that accounts for city-specific effects, recognizing that direct measures may be influenced by differing living expenses and consumption structures.

Empirical Contribution: We identify distinct labor sorting patterns across different city sizes in China, showing that city-level complementarities increase with city size. Specifically, we observe thick tails in skill distributions for super-cities compared to big cities, and first-order stochastic dominance (FOSD) when comparing super-cities to medium- and small-sized cities.

By addressing these challenges, this study aims to enrich the literature on labor migration and sorting within the context of Chinese urbanization. Furthermore, it offers valuable insights for crafting policies to attract talent and promote sustainable development across cities of various types.

3. The Model

3.1 Model Setup

In the spirit of the model with heterogeneously skilled workers under perfect mobility, we assume workers are indexed by a discrete skill type $i \in \mathrm{I} = \{1,...,I\}$. Without loss of generality, let {1,…,I} be an ordered sequence from low to high, with i increasing in skill level. Denote the productivity level associated with ith skill type by yi and the total measure of ith skill by M_i. Assume there are J cities $\mathrm{j} \in \mathrm{J} = \{1,...,\mathrm{J}\}$, each with a fixed amount of land H_j.

A worker of a skill type i who lives in city j has preferences over consumption c_{ij} and housing h_{ij} priced p_j. Her consumer preferences are represented by $\mathrm{u(c,h)} = \mathrm{c}^{1\text{-}\alpha} h^{\alpha}$, where $\alpha \in [0,1]$. Workers are under perfect mobility, so in equilibrium they should be indifferent among consumption-housing bundles in different cities, i.e. for cities j and j', $\mathrm{u}(\mathrm{c}_{\mathrm{ij}}, \mathrm{h}_{\mathrm{ij}}) = \mathrm{u}(\mathrm{c}_{\mathrm{ij'}}, \mathrm{h}_{\mathrm{ij'}})$.

Cities have exogenous total factor productivity denoted by A_j. For skill type i, in city j the representative firm choose a level of employment m_{ij}, wage w_{ij} and output $\mathrm{A}_j F(m_{1j},...,m_{Ij})$. Under such settings, the markets for skills clear by each skill type and the markets for housing clear in each city:

$$\sum_{j=1}^{J} C_j m_{ij} = M_i \qquad \sum_{i=1}^{I} h_{ij} m_{ij} = H_j \tag{1}$$

$$max \;\; u(c_{ij}, h_{ij}) = c_{ij}^{1-\alpha} h_{ij}^{\alpha} \quad ; \quad s.t. \;\; c_{ij} + p_j h_{ij} \le w_{ij} \tag{2}$$

The competitive equilibrium allocations for consumption and housing are: $c_{ij}^{*} = (1-\alpha) w_{ij}$ and $h_{ij}^{*} = \alpha \frac{w_{ij}}{p_j}$.

Plugging in the optimal values we derive the indirect utility for worker of skill type i as:

$$U_i = \alpha^{\alpha} (1-\alpha)^{1-\alpha} \frac{w_{ij}}{p_j^{\alpha}} \tag{3}$$

Rewrite (3) to solve for wage, $w_{ij} = U_i p_j^{\alpha} \frac{1}{\alpha^{\alpha} (1-\alpha)^{1-\alpha}}$. Now we obtain the relationship between wages across cities:

$$\frac{w_{ij}}{w_{ij'}} = \left(\frac{p_j}{p_{j'}} \right)^{\alpha} \tag{4}$$

Assuming that firms are price-takers and do not affect wages, with a given productivity A_j for city j, firms solve profit maximizing problem as:

$$\max_{\mathrm{m_{ij}}, \forall i} A_j F(m_{1j}, ..., m_{Ij}) - \sum_{i=1}^{I} w_{ij} m_{ij} \;\; ; \quad s.t. \quad m_{ij} \ge 0 \tag{5}$$

with first order condition: $A_j F_{m_{ij}}(m_{ij}) = w_{ij}$ .Here we restrict the production with a Constant Elasticity of Substitution technology, which allows variations of elasticity across skill types. Based on the CES technology, we rewrite the firm's production function as:

$$\max_{\mathrm{m_{ij}}, \forall i} A_j F(m_{1j}, ..., m_{Ij}) = \max_{\mathrm{m_{ij}}, \forall i} A_j (\sum_{i=1}^{I} m_{ij}^{\gamma} y_i) \tag{6}$$

where $\gamma < 1$. The first order condition is $A_j \gamma m_{ij}^{\gamma - 1} y_i = w_{ij}$. The equilibrium allocation shows that the wages for workers across skill types are determined by two components: the productivity of the ith skilled worker y_i and the employment of the ith skill workers m_{ij}. Without loss of generality here we assume wage is monotonic in i.

We now introduce complementarities between different skill types. For simplicity, assume there are two cities $\mathrm{j} \in \{1,2\}$ and three skill types $\mathrm{i} \in \{1,2,3\}$. Pick any two skill types i and k where there is complementarity and denote the third skill type as l. Let λ_j be the level of complementarity in city j. This allows us to write down the technology as: $\lambda_j (m_{ij}^{\gamma} y_i + m_{kj}^{\gamma} y_k) + m_{lj}^{\gamma} y_l$.

3.2 Equilibrium of City Size and Complementarities

We now proceed by defining the two cases of complementarities as such:

I. Extreme-Skill Complementarity, where high skill workers and low skill workers are complements:

$$A_j F(m_1, m_2, m_3) = A_j [\lambda_j (m_{1j}^{\gamma} y_1 + m_{3j}^{\gamma} y_3) + m_{2j}^{\gamma} y_2] \tag{7}$$

II. Top-Skill Complementarity, where high skill workers and medium skill workers are complements:

$$A_j F(m_1, m_2, m_3) = A_j[\lambda_j(m_{2j}^{\gamma} y_2 + m_{3j}^{\gamma} y_3) + m_{1j}^{\gamma} y_1] \tag{8}$$

Now we will take Case I, Extreme-Skill Complementarity as an example to derive the equilibrium conditions. The firm's problem is:

$$\max A_j[\lambda_j(m_{1j}^{\gamma} y_1 + m_{3j}^{\gamma} y_3) + m_{2j}^{\gamma} y_2] - \sum_i^3 w_{ij} m_{ij} \tag{9}$$

The first order conditions with respect to each skill type in city j are respectively:

$$\begin{cases} A_j \lambda_j \gamma m_{1j}^{\gamma-1} y_1 - w_{1j} = 0 \\ A_j \gamma m_{2j}^{\gamma-1} y_2 - w_{2j} = 0 \\ A_j \lambda_j \gamma m_{3j}^{\gamma-1} y_3 - w_{3j} = 0 \end{cases} \tag{9}$$

With perfect mobility, the wage ratio can be written in terms of housing price ratio $\frac{w_{i2}}{w_{i1}} = \left(\frac{p_2}{p_1}\right)^{\alpha}$, and this allows us to link employments of different skill types in the two cities $j = 1$ and 2:

$$\begin{cases} A_1 \lambda_1 m_{11}^{\gamma-1} = \left(\frac{p_1}{p_2}\right)^{\alpha} A_2 \lambda_2 m_{12}^{\gamma-1} \\ A_1 m_{21}^{\gamma-1} = \left(\frac{p_1}{p_2}\right)^{\alpha} A_2 m_{22}^{\gamma-1} \\ A_1 \lambda_1 m_{31}^{\gamma-1} = \left(\frac{p_1}{p_2}\right)^{\alpha} A_2 \lambda_2 m_{32}^{\gamma-1} \end{cases} \tag{10}$$

Using the market clearing conditions for skills $C_1 m_{i1} + C_2 m_{i2} = M_i$, we can solve the first order conditions for equilibrium employment of skills:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} M_1}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{21} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma-1}} M_2}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{31} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} M_3}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2} \tag{11}$$

$$m_{12} = \frac{M_1}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{22} = \frac{M_2}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{32} = \frac{M_3}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2} \tag{12}$$

Similarly, with Case II, Top-Skill Complementarity, we can obtain equilibrium allocation of skills as:

$$m_{11} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma-1}} M_1}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{21} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} M_2}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{31} = \frac{\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} M_3}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2} \tag{13}$$

$$m_{12} = \frac{M_1}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{22} = \frac{M_2}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2}, \quad m_{32} = \frac{M_3}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha} \frac{A_2}{A_1} \frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2} \tag{14}$$

Under equilibrium conditions, denote the size of city j as $S_j = \sum_{i=1}^{I} C_j m_{ij}$, and let $A_1 > A_2$ and $p_1 > p_2$. Then the size of city depends on the level of complementarity λ_1 and λ_2: (1)When $\lambda_1 = \lambda_2$, higher total factor productivity will attract all skill types to the same . With $A_1 > A_2$, city 1 will have a higher level of employment at all skill types($m_{11} > m_{21}, m_{21} > m_{22}, m_{31} > m_{32}$) and hence a bigger size, i.e. $S_1 > S_2$. (2)When $\lambda_1 \neq \lambda_2$, the size of the city is determined jointly by complementary of different skills and the city-specific productivity. With $\lambda_1 > \lambda_2$, if matched skill complements accounts for a major part of labor force, there is a great possibility that $S_1 > S_2$. With $\lambda_1 < \lambda_2$, if there is a large proportion of unmatched skill complements, total factor productivity will be the complementarity force, i.e. $S_1 > S_2$.

Corollary 1: Total factor productivity has positive influences on agglomeration, and skill complementarity is conducive to clustering of different skill types; hence higher city-level productivity leads to greater city size.

Proof is in Appendix 1.

3.3 Skill distribution under Extreme-Skill Complementarity

We now derive the skill distribution in each city based on the equilibrium allocation of skill types under

Extreme-Skill Complementarity. The density of skill type i in city j can be written as $pdf_{\text{ij}} = \frac{m_{ij}}{S_j}$. Substitute with m_{ij} and calculate the ratio of probability density of skill type i in city 1 and 2:

$$\frac{pdf_{11}}{pdf_{12}} = \frac{\frac{m_{11}}{m_{11}+m_{21}+m_{31}}}{\frac{m_{12}}{m_{12}+m_{22}+m_{32}}} = \frac{m_{12}+m_{22}+m_{32}}{m_{12}+\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{1-\gamma}} m_{22}+m_{32}} \tag{15}$$

$$\frac{pdf_{21}}{pdf_{22}} = \frac{\frac{m_{21}}{m_{11}+m_{21}+m_{31}}}{\frac{m_{22}}{m_{12}+m_{22}+m_{32}}} = \frac{\left(m_{12}+m_{22}+m_{32}\right)}{\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{1-\gamma}} m_{12}+m_{22}+\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{1}{1-\gamma}} m_{32}} \tag{16}$$

$$\frac{pdf_{31}}{pdf_{32}} = \frac{\frac{m_{31}}{m_{11}+m_{21}+m_{31}}}{\frac{m_{32}}{m_{12}+m_{22}+m_{32}}} = \frac{m_{12}+m_{22}+m_{32}}{m_{12}+\left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{1-\gamma}} m_{22}+m_{32}} \tag{17}$$

Let $A_1 > A_2$ and $0 < \gamma < 1$, when $\lambda_1 > \lambda_2$, we have $\frac{pdf_{11}}{pdf_{12}} > 1$, $\frac{pdf_{21}}{pdf_{22}} < 1$, $\frac{pdf_{31}}{\text{pdf}_{32}} > 1$; when $\lambda_1 < \lambda_2$, we have $\frac{pdf_{11}}{pdf_{12}} < 1$, $\frac{pdf_{21}}{pdf_{22}} > 1$, $\frac{pdf_{31}}{pdf_{32}} < 1$.

Corollary 2: Given Extreme-Skill Complementarity, when level of complementarity is increasing in city size($\lambda_1 > \lambda_2$), skill distribution in the larger city has thicker tails; otherwise when $\lambda_1 < \lambda_2$, skill distribution in the larger city has thinner tails.

4. The Empirical Evidence of Sorting Pattern

4.1 Data

From Equation (3), we have demonstrated that a worker's equilibrium utility depends on their wage and the city-specific housing price. As previously mentioned, while wages increase with skill level, living costs in larger cities are significantly higher. To accurately capture the skill distribution, it is necessary to adjust wages for city-specific housing price levels. This section focuses on identifying the skill distribution by accounting for observable living expenses derived from the wage distribution.

The data for this study comes from the Chinese Household Income Project Survey (CHIP), conducted by the Chinese Household Income Project. This dataset provides comprehensive information on individuals' income levels and sources, educational attainment, and household expenditures through structured interviews. The data used corresponds to the year 2013, the most recent available, and includes annual wage observations for 18,948 households and 64,777 workers across

126 cities. Following the findings of Davis & Ortalo (2011), it is well-established that the share of income allocated to housing remains relatively constant across metropolitan areas. In the context of China, researchers typically estimate this housing expenditure share at 0.3. Accordingly, we adopt this value as our estimate for α, representing the proportion of income spent on housing. Local housing price levels (p) are estimated using commercial housing sales prices obtained from the CEIC database for 2013.

To ensure the analysis is focused on urban labor sorting and maintains data quality, several restrictions are applied to the sample. First, observations from counties and county-level cities are excluded to account for China's distinct urban-rural dual economic structure and the substantial differences in economic development and labor market conditions between these areas. Second, workers who were employed for less than 12 months are excluded, as the dataset reflects annual wages. Third, Winsorization is applied to the utility sample to reduce the influence of spurious or misreported outliers. Finally, workers with negative utility values or missing data are excluded from the analysis. Descriptive statistics of wages, housing prices, and utility for the final sample are summarized in Table 1.

Table 1. Sample descriptive statistics

Variable	Count	Mean	Std	Min	Max
Wage (log)	7960	10.3573	0.6385	7.8633	12.7939
Utility (log)	7960	7.4644	0.6100	5.2073	9.8599
Housing	7960	7842.63	5599.20	2354.21	23427.42

According to the Notice of the State Council on Adjusting the Standards for Categorizing City Sizes, cities are classified into four groups based on the size of their resident population: super-cities with populations exceeding 10 million, megacities with populations between 5 and 10 million, big cities with populations between 1 and 5 million, and medium-small cities with populations below 1 million. The population data used in this study is sourced from the China Urban Construction Statistical Yearbook (2013), which includes both permanent residents and temporary residents, consistent with the definitions provided in the State Council's notice.

Due to the limited representation of megacities in the CHIP data, with only 796 workers observed, including these cities would result in biased kernel density estimations. Consequently, megacities are excluded from our analysis. Instead, we focus on comparing super-cities, such as Beijing and Shanghai, which are more developed and urbanized metropolitan areas, with big cities and medium-small cities.

Table 2. Descriptive Statistics

size	count	Wage(log)				Utility(log)			
		mean	std	10%	90%	mean	std	10%	90%
Super-cities	2501	10.559	0.637	9.831	11.381	7.459	0.611	6.750	8.229
Big cities	2844	10.348	0.589	9.655	11.002	7.486	0.584	6.833	8.206
Medium-small cities	2560	10.171	0.639	9.393	10.933	7.444	0.637	6.671	8.193

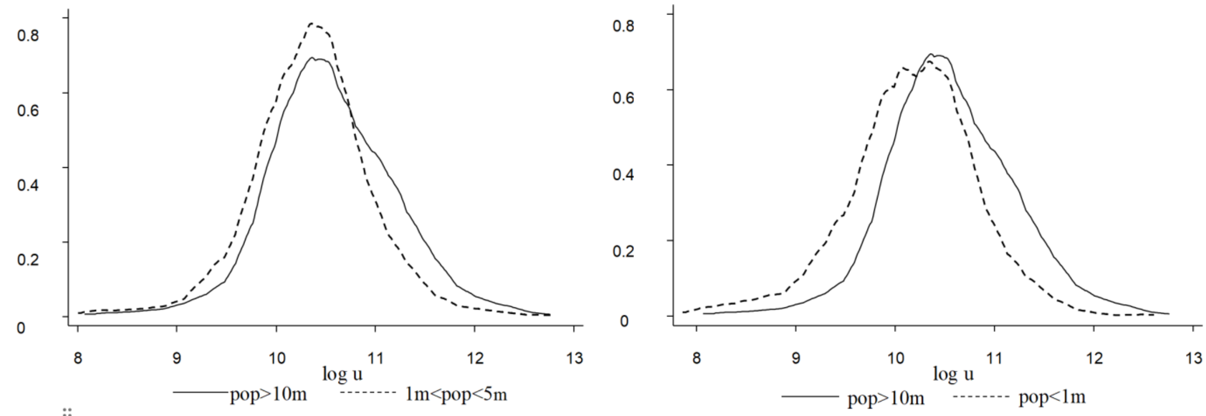
Table 2 presents the population and wage descriptive statistics for the three city groups. The sample sizes across the groups are comparable, with 2,501 workers in super-cities, 2,844 in big cities, and 2,560 in medium-small cities. As expected, the average wage increases with city size, reflecting the well-documented city-size wage premium. Workers at both the bottom and top 10th percentiles of the wage distribution earn more in larger cities than in smaller ones, consistent with the urban wage hierarchy.

The one-to-one relationship between skill and equilibrium utility provides an interpretive framework for the right panel of Table 2. The findings suggest that super-cities attract the highest proportion of high-skilled workers, while medium-small cities attract more low-skilled workers. Interestingly, big cities exhibit a higher average citywide skill level compared to both super-cities and medium-small cities. This pattern highlights the nuanced dynamics of skill sorting across city sizes and underscores the distinctive characteristics of labor distribution within urban China.

4.2 Wage Distribution

Figure 1(a) illustrates the wage distribution for workers in super-cities compared to big cities, while Figure 1(b) compares super-cities with medium-small cities. The density of wages in super-cities displays a thicker upper tail and a thinner lower tail relative to that in big cities, and the entire distribution is shifted to the right. This indicates that the cumulative density of wages in super-cities is consistently lower than in big cities, revealing a clear first-order stochastic dominance (FOSD) relationship: wages in super-cities dominate those in big cities at all levels. This suggests that a greater proportion of workers earn higher wages in larger cities compared to smaller ones.

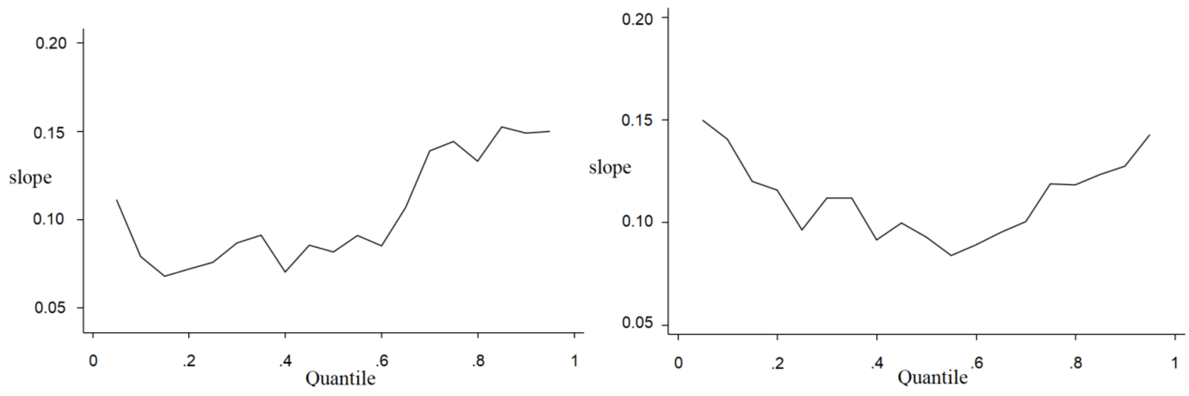
To provide a more comprehensive comparison of wage distributions across cities of different sizes, Figure 2 reports the slopes from quantile regressions of wage on city population size for all wage quantiles. Panel (a) compares super-cities with big cities, while Panel (b) contrasts super-cities with medium-small cities. The regression slopes are significantly positive across all quantiles, demonstrating that wages increase with population size at every percentile. This finding corroborates the observation that larger cities exhibit thicker upper tails and thinner lower tails in their wage distributions, while smaller cities display the opposite pattern.



(a) Super-cities and Big cities (b) Super-cities and Medium-small cities

Figure 1. Wage Distribution

The shape of the slope plot provides further insight into the city-size wage premium. In Panel (a), the slope increases sharply beyond the 60th percentile, indicating that the city-size wage premium is disproportionately higher for high-income workers. In Panel (b), comparing super-cities with medium-small cities, the slope levels are generally higher. Additionally, the slope decreases as wages move away from the lower bound before rising again near the upper bound. This suggests that workers in medium-small cities, whether at the low or high ends of the wage distribution, would benefit more than middle-income workers from moving to super-cities.



(a) Super-cities and Big cities (b) Super-cities and Medium-small Cities

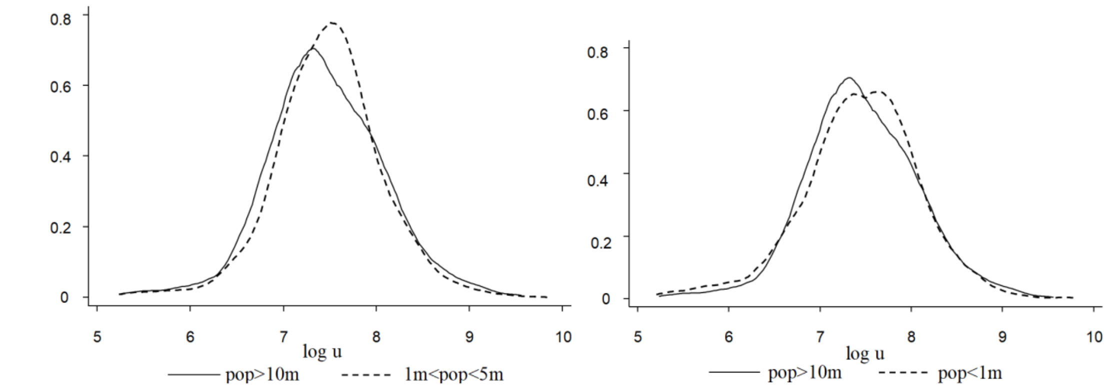
Figure 2. Slopes of Wage Quantile Regression

The FOSD relationship aligns with the well-documented urban wage premium, which posits that wages are higher in larger cities for workers with the same skill level (Glaeser & Mare, 2001). One explanation for this phenomenon lies in competitive equilibrium theory: workers are compensated according to their marginal product, which is enhanced by higher total factor productivity in larger cities. Big cities typically achieve higher productivity through technological spillovers and economies of scale, which elevate wages. However, these benefits come with higher living expenses. Under perfect mobility, wages must compensate for higher housing prices to ensure that workers remain indifferent between the consumption-housing bundles offered by large and small cities.

Regression analysis of wage and housing price on city size reveals that when city size doubles, wages increase by 11.4%, while housing prices rise by 36.5%—more than three times the rate of wage growth. This highlights the need to adjust wages for higher living expenses to accurately uncover the true skill distribution across cities.

4.3 Skill Distribution

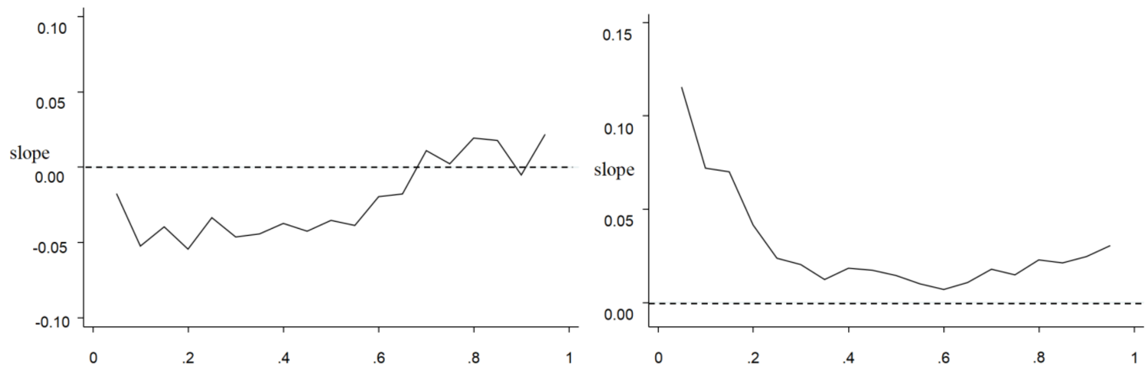
Our baseline model establishes a one-to-one relationship between skill type and a worker's equilibrium indirect utility. Workers with the same skill type are assumed to have identical characteristics under perfect mobility and achieve the same utility level at equilibrium. Using Equation (3), we employ indirect utility as an implicit measure of skill.



(a) Super-cities and Big cities (b) Super-cities and Medium-small Cities

Figure 3. Kernel Density Estimation of Skill Distribution

Figure 3(a) presents the kernel density estimation of skill distributions for super-cities (population over 10 million) and big cities (population between 1 and 5 million). The results indicate that skill distributions in super-cities exhibit thicker tails—both at the top and the bottom—and a lower density at the medium skill interval, consistent with the descriptive statistics reported in Table 2. Super-cities display a higher standard deviation and show a lower 10th percentile(Δ=-0.08) a higher 90th percentile (Δ=0.02), highlighting their thick tails. This suggests that super-cities attract a more dispersed range of workers, including both high-skilled and low-skilled individuals. In contrast, Figure 3(b) compares super-cities with medium-small cities (population below 1 million). The results show that super-cities have thinner bottom tails and thicker top tails compared to medium-small cities. Interestingly, super-cities do not exhibit a higher average skill level, and medium-skilled workers are less inclined to relocate to super-cities compared to cities of smaller sizes.



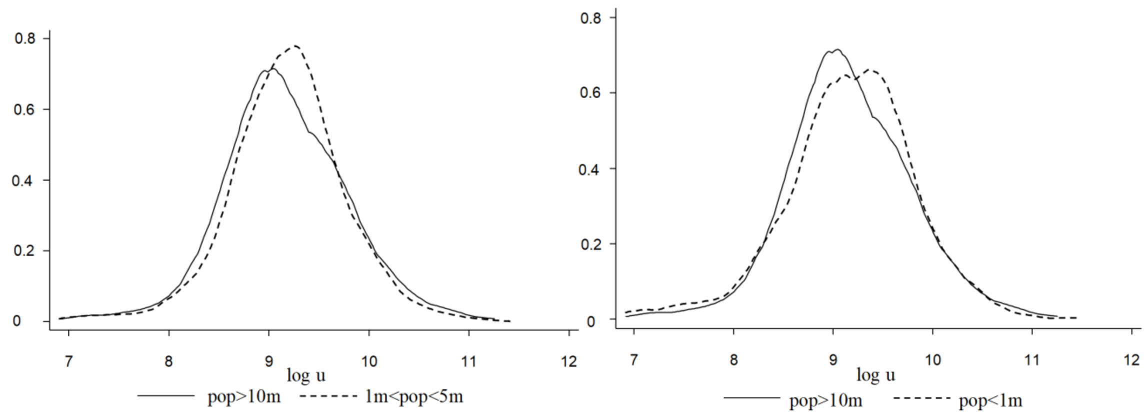
(a) Super-cities and Big cities (b) Super-cities and Medium-small Cities

Figure 4. Slope of Quantile Regression of Skill on Population

Figure 4 illustrates the slope of quantile regression of skill on city population. Panel 4(a) reveals that at the same skill percentile, workers below the 68th percentile display lower skill levels in larger cities, whereas workers above the 68th percentile exhibit increasing skill levels with city size. This finding corroborates the observation that super-cities attract both low-skilled and high-skilled workers. Panel 4(b) depicts a U-shaped relationship, with the curve positioned above zero, indicating that at each skill percentile, working in larger cities generally requires a higher skill level. The skill distribution in super-cities demonstrates first-order stochastic dominance (FOSD) over medium-small cities, while showing thicker tails compared to big cities. This nuanced sorting pattern between city sizes amplifies the findings from the comparison between super-cities and big cities, reflecting differences in urbanization and agglomeration externalities.

4.4 Replacing Rental Price

We conduct a robustness check by replacing commercial housing sales prices with rental prices as an indicator of housing expenses to re-estimate the skill distributions.Figure 5 presents the results of this robustness check. The skill distribution in super-cities continues to exhibit thicker tails compared to that in big cities, consistent with the patterns observed in our main analysis. Additionally, the skill distribution in super-cities shows first-order stochastic dominance (FOSD) over medium-small cities. These findings mirror the sorting patterns identified in the main results and confirm their robustness, thereby enhancing the reliability of our conclusions.



(a) Super-cities and Big cities (b) Super-cities and Medium-small Cities

Figure 5. Robustness Check

5. Further Discussion

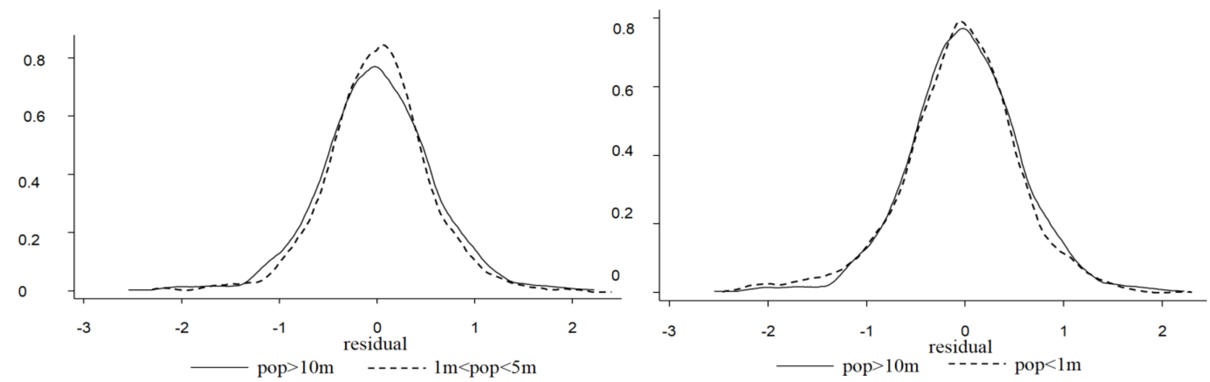
As highlighted in the existing literature, migration and labor mobility play critical roles in shaping skill distributions across cities. It is therefore essential to further analyze sorting patterns by decomposing skills and investigating the impact of job location decisions on overall skill distributions.

5.1 Sorting by Job

To begin, we use occupation as a proxy for observed skills, based on the assumption that higher-skilled workers tend to select better occupations. By leveraging this one-to-one relationship, we gain insights into observed skills by examining occupation ranks. Occupations are ranked according to the countrywide average wage of each category, which serves as a proxy for skill requirements, following the methodologies of Goos et al. (2009) and David et al. (2003). Using CHIP data, we classify the 51 occupation categories into five groups and decompose the price-based skill measure (indirect utility) into two components: an observed part, captured by occupations, and an unobserved part, represented by the residual skill that occupations cannot explain. Specifically, we estimate Equation (13), regressing skills on dummy variables indicating a worker's occupation, and use the regression residual as a proxy for unobserved skills:

$$\log A_i = \sum_i \beta_i X_i + \varepsilon_i \tag{13}$$

where X_i denotes occupation dummies and ε_i the error term. Figure 6 shows the kernel density of unobserved skills in different cities.



(a) Super-cities and Big cities (b) Super-cities and Medium-small Cities

Figure 6. Occupations Distribution

The results indicate that after controlling for observed skills, the residual skill distributions share similar upper tails but differ significantly in their lower tails. Specifically, super-cities exhibit thicker bottom tails compared to big cities, while medium-small cities have thicker bottom tails relative to super-cities. This pattern suggests that for high-paid workers, sorting is primarily determined by observable skills, whereas for low-paid workers, sorting is also influenced by unobservable skills.

These findings underscore the multifaceted nature of labor sorting and highlight the importance of considering both observed and unobserved dimensions of skills when analyzing urban labor market dynamics.

5.2 Sorting by Industry

Workers with similar job descriptions but employed in different industries can exhibit varying earnings, as marginal productivity differs across industries. High-return industries typically demand higher skills, resulting in higher average wage levels and skill requirements (Martins, 2004). Additionally, high-skilled workers tend to concentrate in regions where technology adoption is more intensive (Davis & Dingel, 2012). Given the distinct industrial compositions of Chinese cities, shaped by geographic conditions and government policies, industry-based sorting may also influence skill distributions. To observe unobservable skills, 19 industries were divided into five groups based on average wages and the residual distribution was analyzed (see Figure 7). Super-cities continue to exhibit thick tails compared to big cities, though there is no clear evidence of FOSD when comparing super-cities to medium-small cities. In Figure 7(b), super-cities display a thinner left tail while the right tails remain nearly identical to those of medium-small cities.

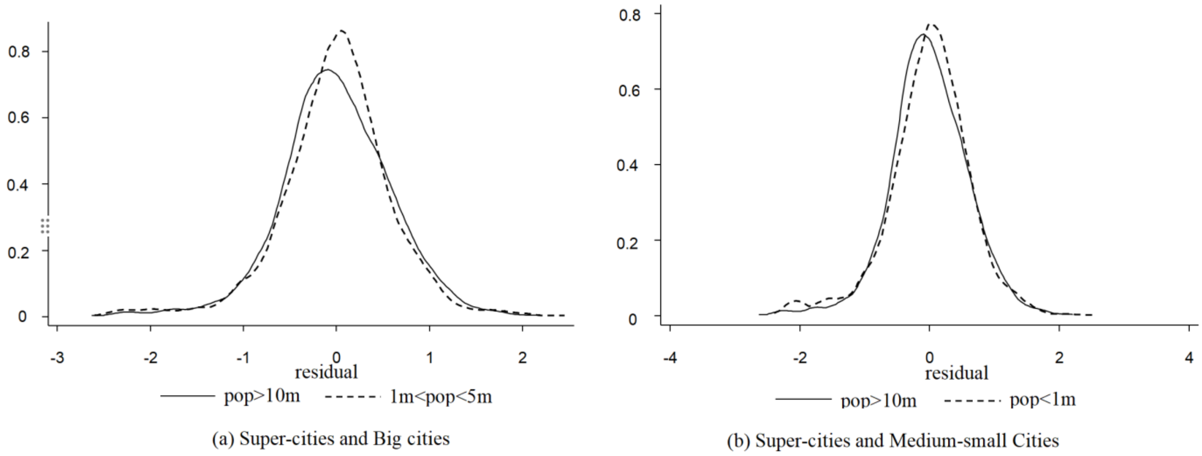


Figure 7. Industry Distribution

Combining the insights from Figures 8 and 9, we find that sorting in super-cities and big cities is influenced by both observable and unobservable skills. However, the differences in tails observed in Figures 8(b) and 9(b) indicate that in medium-small cities, the sorting of high-skilled labor relies more heavily on observable skills, whereas the sorting of low-skilled labor depends on unobservable skills.

These findings underscore the complexity of labor sorting mechanisms across industries and city sizes, highlighting the interplay between observable and unobservable dimensions of skills in shaping urban labor markets.

6. Conclusion

Urbanization fosters labor agglomeration and enhances the efficiency of factor allocation, with worker productivity benefiting from agglomeration externalities such as sharing, matching, and learning. Additionally, skill complementarity, akin to knowledge spillovers, plays a crucial role in shaping labor sorting patterns across cities. In this paper, we integrate the effects of city-level agglomeration externalities into a spatial sorting model, focusing on extreme-skill and top-skill complementarities, to examine labor sorting patterns in Chinese cities categorized as super-cities, big cities, and medium-small cities.

Using CHIP data, we provide robust empirical evidence that the skill distribution in super-cities exhibits thick tails compared to big cities and demonstrates first-order stochastic dominance (FOSD) over medium-small cities. Our results highlight the presence of extreme-skill complementarity in super-cities relative to big cities and top-skill complementarity in super-cities relative to medium-small cities, which collectively shape skill distributions across cities in distinct ways.

Furthermore, our analysis reveals that the sorting of labor is primarily driven by observable skills, such as occupation and industry, for high-skilled workers, while unobservable skills play a more significant role in the sorting of low-skilled workers.

Acknowledgments

Not applicable.

Authors' contributions

Prof. Xu Guo drafted the manuscript and Jingyu Yu revised it. Shuhui Yuan and Wenhao Sun was responsible for data collection. All authors read and approved the final manuscript.

Funding

This work was supported by Humanities and Social Science Fund of Ministry of Education of China (Grant No.20YJC790037); National Natural Science Foundation of China (Grant No.72003018)

Competing interests

Not applicable

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

References

Bacolod, M., Blum, B. S., & Strange, W. C. (2009). Skills in the city. *Journal of Urban Economics*, *65*(2), 136-153. https://doi.org/10.1016/j.jue.2008.07.004

Bauluz, L., Bukowski, P., Fransham, M., Lee, A., Lopez-Forero, M., Novokmet, F., … Verdugo, G. (2024). *Spatial Wage Inequality in North America and Western Europe: Changes Between and Within Local Labour Markets 1975-2019*. Banque de France Working Paper No. 957.

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, *122*(3), 507-553. https://doi.org/10.1086/675858

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., & Van Reenen, J. (2019). What drives differences in management practices? *American Economic Review*, *109*(5), 1648-1683. https://doi.org/10.1257/aer.20170491

Bollinger, C., & Hirsch, B. (2021). Is College a Lousy Investment? Human Capital, Job Matching, and Wage Inequality. *Journal of Labor Economics*. Retrieved from https://www.jstor.org/stable/10.1086/675858

Combes, P. P., & Gobillon, L. (2015). Spatial wage disparities – Workers, firms, and assortative matching', in Duranton, G., Henderson, J. V. and Strange, W. C. (Eds.), *Handbook of Regional and Urban Economics*, *5*, Part B, 557-580. https://doi.org/10.1016/B978-0-444-59517-8.00017-4

Cornelissen, T., Dustmann, C., & Raute, A. (2023). Peer effects in the workplace: Evidence from random group assignments. *American Economic Review*. https://doi.org/10.1257/aer.20210023

David, H., Frank, L., & Richard, J. M. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*(4), 1-44. https://doi.org/10.1163/0033-5533(03)0118-04-0001-44

Davis, D. R. (2011). Spatial Equilibrium and the Efficiency of Large Cities. *Journal of Urban Economics*, *69*(2), 147-162.

https://doi.org/10.1016/j.jue.2010.12.001

Davis, D. R., & Dingel, J. I. (2012) *A spatial knowledge economy*. NBER Working Paper No. 17874. https://doi.org/10.3386/w17874

Davis, D. R., & Dingel, J. I. (2012). *The comparative advantage of cities*. NBER Working Paper No. 17874. https://doi.org/10.3386/w17874

Davis, M. A., & Ortalo-Magné, F. (2011). 'Household expenditures, wages, rents. *Review of Economic Dynamics*, *14*(2), 248-261. https://doi.org/10.1016/j.red.2010.12.001

De la Roca, J., & Puga, D. (2017) 'Learning by working in big cities', *The Review of Economic Studies*, *84*(1), 106-142. https://doi.org/10.1093/restud/rdx017

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. *Handbook of Regional and Urban Economics*, *4*, 2063-2117. https://doi.org/10.1016/S0169-7218(04)40005-8

Eeckhout, J. (2004). Gibrat's Law for (All) Cities. *American Economic Review*, *94*(5), 1429-1451. https://doi.org/10.1257/aer.94.5.1429

Eeckhout, J., Pinheiro, P. S., & Schmidheiny, K. (2016). The skill content of urban agglomeration. *Review of Economics and Statistics*, *98*(4), 1023-1041. https://doi.org/10.1162/0034-6535.2016.00046

Eeckhout, J., Pinheiro, R., & Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, *122*(3), 554-620. https://doi.org/10.1086/675858

Gautier, P. A., & Teulings, C. N. (2009). Search and the City. *Regional Science and Urban Economics*, *39*(3), 251-265. https://doi.org/10.1016/j.regsciurbeco.2008.08.003

Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics*, *19*(2), 316-342. https://doi.org/10.1086/225775

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. *The American Economic Review, 99*(2), 58-63. https://doi.org/10.1257/aer.99.2.58

Marshall, A. (1890). *Principles of Economics*. London: Macmillan and Co.

Martins, P. S. (2004). Industry wage premia: evidence from the wage distribution. *Economics Letters, 83*(2), 157-163. https://doi.org/10.1016/j.econlet.2004.01.003

Neumark, D. (2018). The role of job search frictions in urban labor markets. *Journal of Labor Economics*, *36*(3), 435-466. https://doi.org/10.1086/675858

Pucher, J., & Buehler, R. (2012). The impact of public transportation on urban poverty: A review of the literature. *Transportation Research Part A: Policy and Practice*, *46*(10), 1447-1468. https://doi.org/10.1016/j.tra.2012.06.002

Shimer, R., & Smith, L. (2000). Assortative matching and search. *Econometrica*, *68*(2), 343-372. https://doi.org/10.1111/1467-937X.00095

Tariq, N., Ali, M., & Usman, M. (2024). Impact of human capital and natural resources on environmental quality in South Asia. *Environmental Development and Sustainability*. https://doi.org/10.1007/s10614-024-00000-x

Venables, A. J. (2011). Productivity in cities: Self-selection and sorting. *Journal of Economic Geography*, *11*(2), 241-251. https://doi.org/10.1093/jeg/lbq027

Wheeler, C. H. (2001). Search, sorting, and urban agglomeration. *Journal of Labor Economics*, *19*(4), 879-899. https://doi.org/10.1086/225775

Appendix1：

Proof process of Corollary 1:

The assumption that the larger the city size, the higher the housing price, i.e., $p_1 > p_2$,which is consistent with the theoretical reality. Combining the equilibrium condition $h_{ij} = \alpha \frac{w_{ij}}{p_j}$ when utility is maximized, and the condition $\sum_{i=1}^{3} h_{ij} m_{ij} = H$ when the market is cleared, we obtain:

$$\lambda_j (m_{1j}^{\gamma} y_1 + m_{3j}^{\gamma} y_3) + m_{2j}^{\gamma} y_2 = \frac{Hp_j}{\alpha\gamma A_j}$$

Referring to the treatment of Jan Eeckhout et al. (2014), equation (13) is collapsed using the symmetry assumption to obtain:

$$\left\{\lambda_1 \frac{A_1}{A_2}\frac{P_2}{p_1}\left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\frac{\lambda_1}{\lambda_2}\right]^{\frac{\gamma}{1-\gamma}} - \lambda_2\right\}\left[\frac{M_1}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\frac{A_2}{A_1}\frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2}\right]^{\gamma}\left(y_1 + y_3\right) +$$

$$\left\{\frac{A_1}{A_2}\frac{P_2}{p_1}\left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\right]^{\frac{\gamma}{1-\gamma}} - 1\right\}\left[\frac{M_2}{C_1\left[\left(\frac{p_1}{p_2}\right)^{\alpha}\frac{A_2}{A_1}\frac{\lambda_2}{\lambda_1}\right]^{\frac{1}{\gamma-1}} + C_2}\right]^{\gamma} y_2 = 0 \quad (2)$$

Here, $\lambda_1 \frac{A_1}{A_2}\frac{P_2}{p_1}\left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\frac{\lambda_1}{\lambda_2}\right]^{\frac{\gamma}{1-\gamma}} - \lambda_2$ is recorded as（★）, and $\frac{A_1}{A_2}\frac{P_2}{p_1}\left[\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\right]^{\frac{\gamma}{1-\gamma}} - 1$ is recorded as（◆）. For equation (2) to hold, one of （★）and（◆）must have a positive value and the other a negative value, thus obtaining the necessary condition that:

$\frac{p_2}{p_1} > (\frac{A_2}{A_1}\frac{\lambda_2}{\lambda_1})^{\frac{1}{\alpha\gamma-\gamma+1}}$ and $\frac{p_2}{p_1} > (\frac{A_2}{A_1})^{\frac{1}{\alpha\gamma-\gamma+1}}$ cannot hold simultaneously, and hence the interval ranges for the two key coefficients in the expression for m_{ij} which are:

$$\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2} \in \left(\left(\frac{A_1}{A_2}\right)^{\frac{(1-\gamma)(1-\alpha)}{\alpha\gamma+1-\gamma}}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{\alpha}{\gamma-\alpha\gamma-1}}, \left(\frac{A_1}{A_2}\right)^{\frac{(1-\gamma)(1-\alpha)}{\alpha\gamma+1-\gamma}}\right),$$

$$\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\frac{\lambda_1}{\lambda_2}\in\left(\left(\frac{A_1}{A_2}\right)^{\frac{(1-\gamma)(1-\alpha)}{\alpha\gamma+1-\gamma}}\left(\frac{\lambda_1}{\lambda_2}\right)^{\frac{\alpha}{\gamma-\alpha\gamma-1}},\left(\frac{A_1}{A_2}\right)^{\frac{(1-\gamma)(1-\alpha)}{\alpha\gamma+1-\gamma}}\right)$$

The following situations are discussed separately to compare the magnitude of S_1 and S_2.

①When $\lambda_1=\lambda_2$, both $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}$ and $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\frac{\lambda_1}{\lambda_2}$ are greater than 1, then $m_{11}>m_{21}$, $m_{21}>m_{22}$, $m_{31}>m_{32}$, $S_1=m_{11}+m_{21}+m_{31}>S_2=m_{12}+m_{22}+m_{32}$ hold.

②When $\lambda_1\neq\lambda_2$, $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}>1$ always holds, and the range of $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}\frac{\lambda_1}{\lambda_2}>1$ also depends on the magnitude of λ_1 and λ_2. $\left|\lambda_1-\lambda_2\right|$ is recorded as △.

If △ is small, so that $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}>1$, there is still $S_1>S_2$.

If △ is large, so that $\left(\frac{p_2}{p_1}\right)^{\alpha}\frac{A_1}{A_2}<1$, the size of the city at this point depends on the proportion of the unpaired labor in the city where it is located.

When $\lambda_1>\lambda_2$, the smaller the proportion of unpaired labor, the greater the probability that $S_1>S_2$. When $\lambda_1<\lambda_2$, the larger the proportion of unpaired labor, the greater the probability that $S_1>S_2$.

This situation is in line with the reality. When the unpaired labor accounts for a extremely low proportion in the urban system, the size of the two types of cities depends on the proportion of paired labor in their respective cities. At this time, cities with high productivity have great externalities, the production capacity of paired labor is improved, the degree of agglomeration is further enhanced, and the proportion of paired labor in City 1 is gradually increased, Therefore, the scale of city 1 is larger than that of city 2.

When the unpaired labor accounts for a large proportion in the urban system, the size of the two types of cities depends on the proportion of the unpaired labor force in their respective cities. At this time, the externality of high productivity cities is low, and a small proportion of the paired labor force is mainly concentrated in City 2, while the unpaired labor force is more concentrated in City 1 under the effect of urban productivity, Therefore, the scale of city 1 is still larger than that of city 2.