AI and Spatial Sorting

Ziqian Song[†]

Sciences Po

This Draft: Mar 2026

Artificial intelligence (AI) affects cognitive tasks performed by skilled labor, potentially altering the skill-biased productivity advantages of large cities and spatial sorting. In this paper, we quantify the aggregate and distributional effects of AI-induced spatial sorting using a spatial equilibrium model with heterogeneous labor. We measure AI software by isolating variation in firm-level software assets induced by shocks to public AI R&D grants in China. Our structural estimates indicate that AI is a relative substitute for skilled labor and a relative complement to unskilled labor. Counterfactual analyses show that AI automation overturns canonical spatial sorting of skilled labor, with the skill ratio and skill premium declining most in large cities. We also find that AI automation increases aggregate productivity but lowers the welfare of skilled workers. The key mechanism is the interaction between local agglomeration forces and AI-driven relocation of unskilled labor toward large cities.

Keywords: artificial intelligence, spatial sorting, agglomeration, migration, productivity
JEL Codes: R12, R23, O33, J24

[†]Email: ziqian.song@sciencespo.fr

1 Introduction

A dominant pattern in urban economics in recent decades is the disproportionate spatial sorting of skilled labor into large cities, accompanied by a dramatic increase in the skill premium (Combes et al., 2008; Diamond and Gaubert, 2022). Emerging technologies such as artificial intelligence (AI) may disrupt this pattern by automating cognitive tasks that underlie urban agglomeration economics (Combes et al., 2012; Davis and Dingel, 2019), thereby reshaping the comparative advantage of skilled labor in large cities and fundamentally altering patterns of spatial sorting. This raises two key questions: How does AI influence the spatial distribution of workers? And what are the resulting welfare and distributional implications? Addressing these questions is critical for understanding how to enhance welfare and reduce inequality in the era of AI.

In this paper, we document AI-induced spatial sorting in China and quantify its aggregate and distributional effects. China provides a unique setting for this research due to several reasons. First, China features widespread AI adoption across regions and industries. Second, the country has undergone rapid urbanization in recent decades, marked by large-scale internal migration and a disproportionate concentration of skilled labor in large cities. Together, these factors generate the substantial variation needed to analyze the effects of AI on spatial sorting. Additionally, China's hukou system imposes high barriers to inter-regional migration. This context enables us to evaluate migration policies in the context of AI and derive normative implications for policymakers.[1]

We integrate a task-based model of artificial intelligence (AI) into a Rosen-Roback spatial equilibrium framework with heterogeneous labor (Acemoglu and Restrepo, 2018; Rosen, 1979; Roback, 1982). Representative firms in heterogeneous cities optimally choose input combinations and task allocations, given the elasticity of substitution between AI software and each type of labor. Workers of different skill types make optimal migration decisions based on migration costs and idiosyncratic preferences over locations. We also explicitly model the skill-biased productivity advantage in large cities by assuming a specific functional form for endogenous agglomeration forces (Combes et al., 2012; Davis and Dingel, 2019). This specification allows us to clearly isolate the relationship between AI and spatial sorting via the knowledge spillover channel.

We measure AI software by combining software asset data from Chinese publicly listed firms with AI-related research grants from the National Natural Science Foundation of China (NSFC). Specifically, our measure isolates the variation in firms' software assets explained by their exposure to AI-related public R&D shocks. We then estimate the elas-

[1]The hukou system is a household registration system in China based on an individual's family background. It determines where people are officially permitted to live, work, and access public services. (Song, 2014) In particular, the hukou system restricts migrants' access to public goods in urban areas, with local governments varying the strictness of enforcement over time. (Zhang et al., 2019)

ticity of substitution between AI and both skilled and unskilled labor. Structural estimation indicate that AI is a relative substitute for skilled labor and a relative complement for unskilled labor, in contrast to the predictions of canonical capital–skill complementarity (Krusell et al., 2000). Our findings suggest that AI exhibits the preliminary characteristics of a skill-substituting technology, with potential implications for reducing labor market inequality.

To quantify the spatial equilibrium effects of AI automation, we calibrate the growth in AI software productivity using relative human–AI performance data from Kiela et al. (2021, 2023). In our model, increases in AI software productivity generate two distinct effects: (1) The extensive margin of AI automation, which represents an expansion of the task automation frontier. As AI software becomes more productive, it becomes more profitable to adopt, raising the endogenous automation threshold and increasing the set of tasks performed by AI. (2) The intensive margin of AI automation, which is denoted by a uniform increase in the productivity of AI software. This deepening of automation enhances efficiency in tasks already performed by AI, without necessarily expanding the range of automated tasks.

We conduct a series of counterfactual simulations of AI automation within our spatial equilibrium model. We find that AI automation counteracts the spatial sorting of skilled labor into large cities, resulting in the largest declines in both the skill ratio and the skill premium within these urban cognitive hubs. We also show that while AI automation increases aggregate output, it reduces the welfare of skilled workers. By shutting down the endogenous skill-biased agglomeration force, we find that AI automation has negligible aggregate effects. This highlights that knowledge spillovers are a key mechanism driving the aggregate productivity gains from AI-induced spatial sorting.

While the extensive and intensive margins of AI automation exert similar effects on the spatial sorting of skilled labor, they differ fundamentally in their impact on unskilled labor. At the extensive margin, the substitution effect of AI automation dominates the productivity effect. As AI displaces skilled labor without generating sufficient compensating productivity gains, it causes unskilled labor to out-migrate from large cities. In contrast, the intensive margin generates strong productivity effects that attract unskilled labor to sort into large cities. We find that the aggregate impact is quantitatively dominated by the intensive rather than the extensive margin, driven by the concentration of unskilled workers in highly productive large cities.

This research relates to several strands of literature. First, this paper contributes to the literature on the spatial sorting of workers across cities. Seminal works by Combes et al. (2008) and Diamond (2016) document the increasing concentration of high-skilled workers in high-wage, high-amenity cities, contributing to the dramatic increase in the skill premium and regional inequality. A key mechanism underlying spatial sorting is knowl-

edge spillovers (Combes et al., 2012; Gaubert, 2018; Davis and Dingel, 2019). Theoretically, Eeckhout et al. (2014) explain this pattern through extreme skill complementarity in large cities, while Behrens et al. (2014) propose a framework disentangling how agglomeration economies, sorting, and selection jointly determine labor allocation. Our framework builds on Giannone (2017), who study how the interaction between skill-biased technical change (SBTC) and agglomeration contributes to regional skill divergence. We extend this literature by incorporating AI automation into a production function characterized by skill-biased agglomeration. While existing studies emphasize the disproportionate sorting of skilled labor into large cities, we are among the first to show that AI automation can reverse this pattern, thereby contributing to regional skill convergence. We also show that AI-induced spatial sorting raises aggregate productivity. A key mechanism is the interaction between local agglomeration forces and AI-driven relocation of unskilled labor toward large cities.

Second, this paper complements the growing literature on the labor market effects of technological change. Krusell et al. (2000) estimate capital-skill complementarity within a production function to explain the increase in the skill premium. Autor and Dorn (2013) document that the automation of routine tasks displaces middle-skill labor and generates job polarization, while Eeckhout et al. (2026) extend this routine-biased technical change (RBTC) framework to the urban context, showing that IT investment disproportionately reduces the routine-cognitive employment share in large cities. More recent work has focused on AI. Webb (2019) and Acemoglu et al. (2022) study the early labor-market impacts of AI and show that skilled workers are more exposed to AI. Experimental evidence on large language models (LLMs) (Noy and Zhang, 2023; Brynjolfsson et al., 2023; Peng et al., 2023; Eloundou et al., 2024) finds large productivity gains, with effects that are larger for less experienced workers. We complement this literature by extending the analysis of Eeckhout et al. (2026) to AI software capital. Our structural estimation shows that AI software is a relative substitute for skilled labor and a relative complement to unskilled labor. While Eeckhout et al. (2026) show that IT causes a larger decrease in the routine share and a rise in non-routine cognitive share in large cities, our results highlight a new substitution: AI disproportionately displaces skilled workers (without AI skills), driving the sharpest decline in the aggregate skill ratio in large cities.[2] We also characterize AI-induced spatial sorting and quantify its aggregate and distributional implications

Finally, this paper contributes to the development literature on internal migration. While the prior work has focused on migration barriers (Bryan and Morten, 2019; Tombe and Zhu, 2019), insurance barriers (Munshi and Rosenzweig, 2016), pollution (Khanna

[2]Our results can also be interpreted as AI-skilled labor replacing cognitive workers who lack AI skills. We argue that the effective deployment of AI software requires a complementary bundle of AI investment and AI-skilled labor (Acemoglu et al., 2022). Preliminary empirical evidence showing that pre-LLM AI can substitute for non-routine cognitive workers who lack AI skills is provided by Copestake et al. (2023).

et al., 2025), land security (Adamopoulos et al., 2024), and family migration mode (Imbert et al., 2023) in shaping spatial distributions of workers, we propose a new perspective on how AI affects internal migration decisions and their aggregate implications. Our findings indicate that, in developing economies with ongoing urbanization, AI-induced sorting may lower the skill ratio in large cities and weaken skill-biased agglomeration in the steady state relative to developed economies. Our paper also contributes to the literature on China's "Great Migration" (Tombe and Zhu, 2019; Facchini et al., 2019b; Fan, 2019; Imbert et al., 2022) and draws normative implications for migration policy designed to manage the aggregate productivity gains from AI-induced spatial sorting in China.

The remainder of the paper is organized as follows. Section 2 introduces the data, measurement, and descriptive evidence. Section 3 presents the spatial equilibrium model with AI and heterogeneous skills. Section 4 describes the structural estimation and model fit. Section 5 conducts counterfactual analyses. Section 6 concludes.

2 Data, Measurement, and Descriptive Evidence

In this section, we describe our main data sources and measures of AI software and provide descriptive evidence on the relationship between AI and spatial sorting. To empirically evaluate how AI affects the geography of skills, we address the lack of standardized firm-level measures of AI adoption and combine these measures with migration-flow and wage data from Chinese censuses and surveys.

2.1 Main Data Sources

We measure AI software by combining software-asset data for publicly listed Chinese firms from CSMAR with public grant data from the National Natural Science Foundation of China. Details on the construction of the AI-software measure are provided in Section 2.2.

On the worker side, we use microdata from the 2000, 2010, and 2015 population censuses to construct city-level skill ratios and to measure intercity migration flows. We also use the China Migrants Dynamic Survey (CMDS), which contains detailed information on wages, locations, and industries, to estimate city-level skill premium. Additional details on the data sources are provided in the Appendix B.1.

2.2 Measurement of AI Software

A key empirical challenge is that AI-specific software investments are rarely disaggregated in standard corporate balance sheets. To overcome this, we introduce a novel measure of AI software by isolating the component of firms' software assets driven by AI-related pub-

lic R&D shocks.[3] The identifying assumption is that when AI-related public knowledge generated by grant-funded research diffuses to listed firms and induces incremental software investment, that incremental component reflects AI-related software assets. Because public grants are allocated through a highly competitive peer-review process, the resulting variation is plausibly orthogonal to firms' software investment decisions. We discuss the institutional criteria for the exogeneity of AI-related public R&D grant shocks in subsection 2.2.1.

2.2.1 Criteria of NSFC

The National Natural Science Foundation of China (NSFC) is the primary national agency funding basic research in the natural sciences. Its grant allocation follows a highly competitive, multi-stage peer-review process. Each year, principal investigators submit proposals that undergo an initial administrative screening, followed by a rigorous evaluation by three to five external academic experts.

Funding decisions are based on three core criteria: the scientific merit and novelty of the proposed research, methodological feasibility, and the principal investigator's prior academic record. Because this national allocation process is primarily driven by scientific criteria, the distribution of NSFC grants provides a plausibly exogenous source of variation in public R&D shocks.

2.2.2 Construction of the AI Software Measure

We begin by extracting software asset data from the balance sheets of all publicly listed firms. Next, we identify AI-related public research grants from China's National Natural Science Foundation using a curated set of keywords.[4] These grant records contain rich information, including project titles, abstracts, and funding amounts, which allow us to clearly identify AI-specific R&D grants. Descriptive evidence on public AI R&D grants and publicly listed firms is provided in Appendix B.2.

We then compute the technological proximity between each firm and university using the cosine similarity of their respective technology portfolios. This proximity allows us to isolate the exogenous variation in firms' software assets driven by public AI grant shocks, yielding a firm-level measure of AI software. Finally, we aggregate the firm-level measures to the city level and combine them with local wage-bill data to construct our city-level AI-capital measure. Details of the construction are as follows.

We assume that AI software is a structural component of a firm's total software assets,

[3] Our measure covers software-based AI applications used by firms, such as machine learning, natural language processing, large language models, and related applications.

[4] The keyword list includes both English and Chinese terms, such as: 人工智能 (Artificial Intelligence), 机器学习 (Machine Learning), 监督学习 (Supervised Learning), 深度学习 (Deep Learning), 强化学习 (Reinforcement Learning), 神经网络 (Neural Networks), 自然语言处理 (Natural Language Processing, NLP), 计算机视觉 (Computer Vision, CV), 语音识别 (Speech Recognition), 知识图谱 (Knowledge Graphs).

governed by the following relationship:

$$\text{Software}_{ft}^{\text{AI}} = \left(\text{Software}_{f}\right)^{\delta}$$

where Software_{ft} denotes the total software assets of firm f in year t, $\text{Software}_{ft}^{\text{AI}}$ denotes the AI-relevant software component, and δ represents the elasticity of AI software with respect to total software.

The production of AI software at the firm level is modeled as a Cobb–Douglas function of public AI grants and private R&D investment:

$$\text{Software}_{ft}^{\text{AI}} = \left[\prod_{u \in ll} \left(G_{ut}\right)^{s_{fu}}\right]^{\gamma} I_{ft}$$

where $\text{Software}_{ft}^{\text{AI}}$ denotes the AI-specific software of firm f in year t, I_{ft} represents the firm's private R&D investment on AI, and G_{ut} is the amount of public AI-related grant funding received by university u in year t. The term s_{fu} captures the technological distance between firm f and university u, and γ denotes the elasticity of AI software production with respect to public R&D spillovers.

We measure the predetermined technological proximity between universities and firms using the following metric:

$$s_{fu} = \frac{V_f \cdot V_u}{\|V_f\| \cdot \|V_u\|} \cdot \omega_{l(f),l(u)}, \qquad \omega_{l(f),l(u)} = \frac{1}{d_{l(f)l(u)}}$$

where V_f and V_u denote the technological stock vectors of public firm f and university u, respectively. Following Bloom et al. (2013) and Dyevre (2023), we assume entities with higher technological similarity benefit more from knowledge spillovers.[5] The technological vector for each firm is constructed using its patent portfolio prior to 2010, while the university vector is based on historical public research funding received before 2010.[6] In addition, $d_{l(f)l(u)}$ denotes the geographic distance between the cities in which firm f and university u are located. This term captures the localized nature of knowledge spillovers (Jaffe et al., 1993).

We estimate the elasticity of firm software assets with respect to public AI grants by Poisson Pseudo-Maximum Likelihood (PPML) using the following specification:

$$\ln \text{Software}_{ft} = \frac{\gamma}{\delta} \cdot \sum_{u \in ll} s_{fu} \ln G_{u,t-1} + \eta_{dj} + \mu_t + \epsilon_{ft}, \tag{1}$$

[5] We abstract from the possibility that these spillovers generate business-stealing effects across regions.

[6] We map Chinese academic subject classifications to IPC classes using standardized definitions to ensure consistency in the technological space.

where $\ln \text{Software}_{ft}$ denotes the log of software assets for firm f in year t, while $\sum_{u \in \mathcal{U}} s_{fu} \ln G_{u,t-1}$ captures the firm's exposure to lagged AI-related public grants, weighted by the technological and geographic proximity s_{fu} to universities. We include sector-city fixed effects η_{dj} and year fixed effects μ_t.[7] Finally, ϵ_{ft} denotes the idiosyncratic error term.

To isolate the component of software capital driven by the public AI R&D shock, we compute predicted AI software capital as follows:

$$\widehat{\text{Software}}_{ft}^{\text{AI}} = \exp\left(\frac{\hat{\gamma}}{\hat{\delta}} \sum_{u \in \mathcal{U}} s_{fu} \ln G_{u,t-1} + \hat{\eta}_{dj} + \hat{\mu}_t\right) - \exp\left(\hat{\eta}_{dj} + \hat{\mu}_t\right) \tag{2}$$

Specifically, we compute the AI software capital for each firm f in year t by taking the difference between its total predicted software assets and the counterfactual predicted software assets where the public AI grants are set to zero.

Table 1: Measurement of AI Software

Dependent Variable	Software_{ft}			
	(1)	(2)	(3)	(4)
$\sum_{u \in \mathcal{U}} s_{fu} \ln G_{u,t-1}$	0.049	0.048	0.058	0.057
	(0.003)	(0.003)	(0.009)	(0.009)
Variance Share of AI	0.028	0.020	0.027	0.023
Variance Share of FE	0.906	0.928	0.919	0.929
Variance Share of Covariance	0.066	0.052	0.054	0.048
Control × Year Dummy	NO	YES	NO	YES
Year FE	YES	YES	YES	YES
City-Sector FE	NO	YES	NO	YES
City FE	YES	NO	YES	NO
Sector FE	YES	NO	NO	NO
Method	PPML	PPML	PPML	PPML
Observations	14,249	14,249	14,249	14,249
R-squared	0.705	0.736	0.815	0.831

Notes: Standard errors are clustered at the firm level.

Next, we aggregate AI software to the city level. First, we define two distinct periods: Period 1 covers the years 2012 and 2013, while Period 2 covers 2017 and 2018. Using our public firm-level sample, we compute the AI-to-labor ratio (R_{jt}) for each city j in period t by aggregating the predicted AI software and firm-level wage bills:

$$R_{jt} = \frac{\sum_{f \in \mathcal{F}_j} \widehat{\text{Software}}_{ft}^{\text{AI}}}{\sum_{f \in \mathcal{F}_j} \text{Wage Bill}_{ft}},$$

[7] Unlike most reduced-form specifications, we do not include firm fixed effects. Because AI adoption is highly heterogeneous across firms, including these fixed effects would absorb the critical between-firm variation necessary to identify our effects.

where $\mathcal{F}_j$ denotes the set of firms in city j observed in our sample, and Wage Bill$_{ft}$ is the total wage bill of firm f in period t.

Finally, we scale this estimated AI-to-labor ratio by each city's total wage bill to recover the aggregate city-level AI software (K_{jt}):

$$K_{jt} = R_{jt} \times \text{Total Wage Bill}_{jt}.$$

Since our measure relies primarily on publicly listed firms' software assets and public R&D grants, it serves as a conservative proxy for total AI investment. In particular, it omits complementary inputs such as hardware infrastructure, unobserved private R&D spillovers, and industrial policies that may also contribute to AI investment.

Nonetheless, our estimates remain informative. We find that approximately 2% of the variation in public firms' software assets can be explained by exposure to public AI funding shocks. More importantly, public grants are positively correlated with private investment in AI, such that $\text{Cov}\left(\left[\prod_{u \in ll} (G_{ut})^{s_{fu}}\right]^{\gamma}, I_{ft}\right) > 0$). Therefore, our measure also captures the endogenous response of private investment to public AI funding.

Spatial Distribution of AI Software We then describe the spatial distribution of AI software investment across cities.

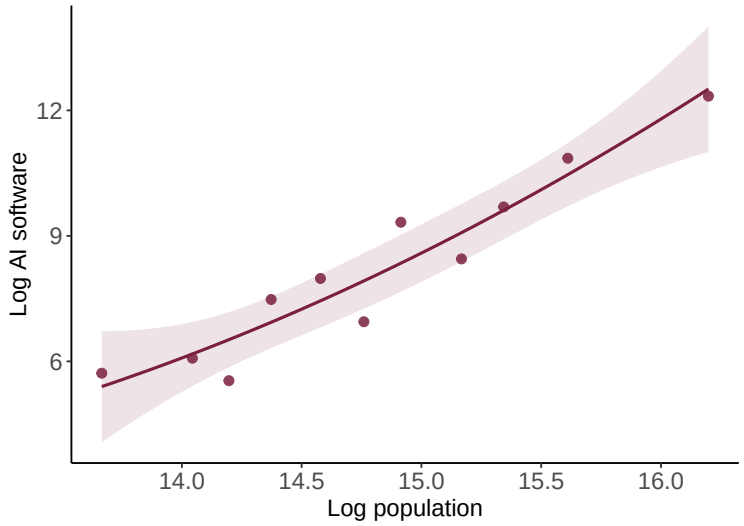


Figure 1: AI Software Investment and City Size.

Notes: Cities are grouped into 11 bins based on population size, with each point representing the average of 10 cities.

Figure 1 plots city AI software investment against city population. Large cities exhibit substantially higher AI-software investment, consistent with concentration of intangible capital adoption in larger urban centers.

2.2.3 Robustness

We conduct a series of robustness checks for the construction of AI software.

Alternative distance decay We consider alternative exponential distance-decay specifications:

$$\omega_{l(f),l(u)} = \exp\left(-\nu\, d_{l(f),l(u)}\right),$$

where

$$\nu = \frac{\log 2}{h}, \qquad h \in \{100, 200, 500\} \text{ km}.$$

We also drop the distance term entirely, allowing knowledge spillovers that do not decay with distance.

Dropping technological similarity As an additional robustness check, we set $\omega_{l(f),l(u)}$ to depend only on geographic distance. This specification allows knowledge spillovers to depend solely on spatial proximity, independent of technological proximity between firms and universities. We then compare the resulting AI-software measure with our baseline measure.

$$\omega_{l(f),l(u)} = \frac{1}{d_{l(f),l(u)}}.$$

Endogenous allocation to innovation clusters A potential concern is that NSFC grants may be endogenously concentrated in innovation centers, which could generate localized knowledge spillovers to nearby private firms. In that case, our AI grant-induced software measure could be biased by endogenous co-location of public funding and private technology adoption.

Although NSFC allocation is primarily based on scientific criteria and we do not find direct evidence of systematic targeting, we address this concern using a leave-one-out strategy:

$$\sum_{u \in \mathcal{U} \cap l(u) \neq l(f)} s_{fu} \ln G_{u,t-1}. \tag{3}$$

Specifically, when constructing the public-grant shock for firms in a given city, we exclude all grants allocated to universities located in the same city.

2.3 Descriptive Evidence

2.3.1 City AI Exposure

To measure city-level AI exposure, we leverage occupational composition data from the 2000 China Population Census. We construct a shift-share measure by interacting these predetermined local occupational shares with a occupation-specific AI exposure index (Felten et al., 2021). This approach exploits cross-city variation in initial employment structures to identify the local impact of AI shocks.

The city-level AI exposure is defined as:

$$\text{AI}_j = \sum_o \left(\frac{L_{jo,2000}}{L_{j,2000}} \right) \times \text{AIOE}_o \tag{4}$$

where AIOE_o represents the AI Occupational Exposure score for occupation o developed by Felten et al. (2021). $L_{jo,2000}$ denotes the employment in occupation o in city j in the base year 2000, and $L_{j,2000}$ is the total employment in city j in the same year.

2.3.2 Can AI Exposure Predict AI Investment?

With this AI exposure measure, we first conduct a validity check for our AI-software investment measure. We examine whether cities with higher exposure experience larger increases in AI-software investment.

Due to data limitations, we focus on long-difference changes between 2012–2013 and 2017–2018. We measure changes in software investment using three outcomes: (1) $\Delta_t \ln K_j$, the log change in AI software; (2) $\Delta_t \ln K_j^{\text{Total}}$, the log change in total software; and (3) $\Delta_t \ln K_j^{\text{Non-AI}}$, the log change in non-AI software. The latter two outcomes serve as placebo tests.

Table 2: Changes in AI Software and AI Exposure

Dependent Variable	$\Delta_t \ln K_j$		$\Delta_t \ln K_j^{\text{Total}}$		$\Delta_t \ln K_j^{\text{Non-AI}}$	
	(1)	(2)	(3)	(4)	(5)	(6)
AI_j	1.238	1.118	0.404	-0.092	-0.834	-1.210
	(0.424)	(0.392)	(0.432)	(0.571)	(0.422)	(0.632)
Mean of Dep. Variable	0.605	0.605	-0.011	-0.011	0.325	0.325
Control	NO	YES	NO	YES	NO	YES
Method	OLS	OLS	OLS	OLS	OLS	OLS
Observations	123	123	123	123	123	123
R-squared	0.095	0.085	0.105	0.114	0.092	0.077

Notes: Standard errors are clustered at the city level. Controls include log city population and the skill ratio in 2000.

Table 2 reports the estimates. Columns (1)–(2) show that higher AI exposure is associated with larger increases in AI-software investment. Columns (3)–(6) show that AI exposure does not predict changes in total software investment or non-AI software investment, supporting the validity of our AI-software measure as capturing variation specific to AI investment.

2.3.3 Skill Convergence and AI Exposure

We next examine how AI affects regional convergence in skill ratios. In the United States, regional skill levels converged between 1940 and 1980 (Giannone, 2017). After 1980, a

"Great Divergence" emerged, with regions that had higher initial skill ratios experiencing faster subsequent growth in skill ratios (Moretti, 2012). China's 2010–2020 period shares several features with the post-1980 United States, especially rising human capital and information-intensive industries. This period also featured rapid urbanization and large-scale internal migration, with skilled workers disproportionately sorting into large cities. A key difference is that, whereas the U.S. "Great Divergence" is associated with the rise of information technology (IT), China's large-scale spatial sorting coincides with rapid AI development. We therefore ask two questions. First, did China experience regional skill convergence or divergence? Second, does AI adoption reinforce or counteract regional skill divergence?

We address these questions by incorporating city-level AI exposure into the following equation:

$$\frac{1}{t-t_0}\Delta_t \ln\left(\frac{S_j}{U_j}\right) = \alpha + \beta_1 \ln\left(\frac{S_j}{U_j}\right)_{t_0} + \beta_2 \mathrm{AI}_j + \beta_3 \left[\ln\left(\frac{S_j}{U_j}\right)_{t_0} \times \mathrm{AI}_j\right] + \mathrm{X}_j'\gamma + \varepsilon_j \quad (5)$$

where $\Delta_t \ln\left(\frac{S_j}{U_j}\right)$ denotes the log change in the skill ratio between two periods. Because our analysis compares two periods of unequal length (a 10-year period from 2000 to 2010 and a 5-year period from 2010 to 2015), we standardize the dependent variable by dividing the total change by the number of years, $(t-t_0)$. This transformation yields an annualized average growth rate, ensuring that the point estimates are comparable across both horizons. The term $\ln\left(\frac{S_j}{U_j}\right)_{t_0}$ denotes the baseline log skill ratio at the start of the respective period (2000 or 2010). X_j is a vector of controls including log population of city j in 2000. Finally, ϵ_j denotes the error term.

Table 3: Skill Convergence and AI Exposure

Dependent Variable	$\frac{1}{t-t_0}\Delta_t \ln\left(\frac{S_j}{U_j}\right)$					
	2000–2010			2010–2015		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln\left(\frac{S_j}{U_j}\right)_{2000}$	-0.017	-0.014	-0.012			
	(0.002)	(0.002)	(0.003)			
$\ln\left(\frac{S_j}{U_j}\right)_{2010}$				-0.031	-0.022	-0.027
				(0.005)	(0.005)	(0.006)
AI_j		-0.023	0.031		-0.087	-0.170
		(0.010)	(0.059)		(0.014)	(0.054)
$\left[\ln\left(\frac{S_j}{U_j}\right)_{t_0} \times \text{AI}_j\right]$			0.017			-0.042
			(0.018)			(0.024)
Mean of Dep. Variable	0.098	0.098	0.098	0.050	0.050	0.050
Control	YES	YES	YES	YES	YES	YES
Year FE	YES	NO	NO	NO	NO	NO
Method	OLS	OLS	OLS	OLS	OLS	OLS
Observations	294	294	294	294	294	294
R-squared	0.214	0.234	0.237	0.172	0.297	0.305

Notes: Standard errors are clustered at the city level. The sample is constructed from the 2010 and 2015 population censuses. Skilled labor is defined as workers with a junior-college degree or above, and unskilled labor is defined as workers with a high school degree or below. The sample includes individuals aged 15–64. Controls include log city population in 2000. Columns (1)–(3) report estimates using changes over 2000–2010, while Columns (4)–(6) report estimates using changes over 2010–2015.

Table 3 presents the main results using long-difference specifications to characterize regional convergence patterns. Columns (1)–(3) report estimates for the 2000–2010 subsample, while Columns (4)–(6) report estimates for the 2010–2015 subsample.

First, the estimates indicate robust conditional convergence in skill ratios: regions with higher initial skill ratios experience systematically lower subsequent skill-ratio growth. This differs from the post-1980 U.S. "Great Divergence" documented in Moretti (2012), where skill accumulation became increasingly concentrated in innovation hubs and large cities. Second, the negative coefficient on AI exposure indicates that higher AI exposure is associated with a declining skill ratio. More importantly, the coefficient on the interaction between AI exposure and the initial skill ratio (β_3) is significantly negative in 2010–2020 but positive in 2000–2010. In contrast to the "Great Divergence" associated with skill-biased technological change and the agglomeration of innovative hubs in the United States (Giannone, 2017), our evidence suggests that AI in China disproportionately weakens skill concentration in large cities and contributes to a convergence in regional skill levels.

A natural concern is that our exposure measure proxies for a predetermined industrial mix that independently predicts changes in the skill ratio. We address this concern by using the 2000–2010 estimate in Column (3) as a falsification test. Because AI was not

widely adopted during this earlier period, the estimated coefficient is statistically insignificant and exhibits the opposite sign relative to the baseline estimate in Column (6). This falsification test indicates that our AI exposure measure is unlikely to be confounded by pre-existing trends in the regional industrial mix.

Our findings provide preliminary evidence on AI and spatial sorting. We highlight two points. (1) AI exposure is associated with a declining skill ratio, consistent with skilled labor being more exposed to AI-driven task displacement (Acemoglu et al., 2022; Copestake et al., 2023). (2) Despite rising human capital and technological innovation, China exhibits regional skill convergence rather than divergence after 2010. This spatial pattern is partly associated with uneven regional adoption of AI.

3 A Quantitative Model with AI and Heterogeneous Skills

In this section, we incorporate a task-based framework of artificial intelligence (Acemoglu and Restrepo, 2018) into a Rosen–Roback spatial equilibrium model (Rosen, 1979; Roback, 1982). This framework provides a micro-foundation for capturing how AI displaces tasks traditionally performed by skilled labor.

The economy consists of a set of origin cities O, and a set of destination cities $\mathcal{D}$. We index origins by $i \in O$ and destinations by $j \in \mathcal{D}$. Representative firms operate in each destination city j, employing workers of heterogeneous skill types to produce a local output. Workers are born in an origin city i and are endowed with an exogenous skill type (skilled or unskilled).

3.1 Production

In each city j, a representative firm operates under perfect competition and takes the price of the final good as the numeraire. Output Y_j is produced using a nested Constant Elasticity of Substitution (CES) production function that combines unskilled labor (U_j) with a composite input of skilled labor and AI software (ES_j). In the inner nest, we incorporate a task-based framework in which tasks are performed by AI software (K_j) and skilled labor (S_j). These inputs are then aggregated into a CES composite (ES_j).

Outer CES Nest The outer nest aggregates unskilled labor and the AI–skilled labor composite as follows:

$$Y_j = A_j \left[\eta_j^u (U_j)^\sigma + \eta_j^s (ES_j)^\sigma \right]^{\frac{1}{\sigma}} \tag{6}$$

where U_j is unskilled labor, ES_j is the AI–skilled labor composite, and η_{uj} and η_{sj} are city-specific productivity terms for unskilled labor and the AI–skilled composite, respectively. The elasticity of substitution between unskilled labor and the AI–skilled composite is $1/(\sigma - 1)$.

Building on the literature on skill-biased agglomeration (Combes et al., 2012; Gian-

none, 2017; Davis and Dingel, 2019), we model agglomeration forces as a function of the local skill composition, total population, and exogenous fundamentals:

$$\eta_j^s = \left(\frac{S_j}{U_j}\right)^{\lambda^s} (S_j + U_j)^{\phi^s} \exp(\xi_j^s) \tag{7}$$

$$\eta_j^u = \left(\frac{S_j}{U_j}\right)^{\lambda^u} (S_j + U_j)^{\phi^u} \exp(\xi_j^u) \tag{8}$$

where ξ_j^s and ξ_j^u denotes fundamental productivity of city j for the AI-skilled labor composite and unskilled labor, respectively. S_j and U_j represents skilled and unskilled labor. The parameters (λ^s, λ^u) and (ϕ^s, ϕ^u) govern the agglomeration force associated with the skill ratio and total population, respectively.

Inner CES Nest The inner nest is assembled from a continuum of tasks $i \in [0, 1]$. The AI–skilled labor composite ES_j is defined as:

$$ES_j = \left[\int_0^1 z_j(i)^\rho di\right]^{\frac{1}{\rho}} \tag{9}$$

where $z_j(i)$ denotes the intermediate input for task i in city j, and ρ governs the elasticity of substitution across tasks.

The intermediate good $z_j(i)$ can be produced using either skilled labor or AI software. Skilled labor can perform all tasks $i \in [0, 1]$, whereas only tasks in $[0, \beta^k]$ are automatable and can be performed by AI software. The task-level production function is given by:

$$z_j(i) = \begin{cases} A^s l(i), & \text{if } i \in [0, 1] \\ A^k \psi(i) k(i), & \text{if } i \in [0, \beta^k] \end{cases} \tag{10}$$

where $l(i)$ denotes skilled-labor input in task i, $k(i)$ denotes AI-capital input in task i, and $\psi(i)$ denotes the task-specific productivity term. A^s and A^k denote the factor-augmenting terms for skilled labor and AI software, respectively.

For tractability in our quantitative analysis, we impose the following functional form for the task-specific productivity term:

$$\psi(i) = i^{-1/\theta} \tag{11}$$

where $\theta > 0$. In Appendix D.1, we report numerical simulations of $\psi(i)$ for alternative values of θ.

Under this specification, task-level AI productivity is strictly decreasing and convex

in the task index i: low-index tasks are easier to automate, whereas high-index tasks are harder to automate. This aligns with our conception that AI primarily expands the automation frontier for both routine cognitive tasks and the non-routine cognitive tasks performed by workers without AI skills.

Because tasks are produced competitively, the price of each task $p_j(i)$ equals the marginal cost of its production input:

$$p_j(i) = \begin{cases} \dfrac{w_j^s}{A^s}, & \text{if} \quad i \in [\beta^k, 1], \\ \min\left\{\dfrac{w_j^s}{A^s}, \dfrac{v}{A^k \psi(i)}\right\}, & \text{if} \quad i \in [0, \beta^k]. \end{cases} \tag{12}$$

where w_j^s denotes the skilled wage in city j, and v denotes the price of AI software, which is normalized to one.

Consistent with the framework of Acemoglu and Restrepo (2018, 2022), given that skilled labor holds a strict comparative advantage in higher-indexed tasks, there exists a unique threshold $\widetilde{\beta}_j^k \in (0, 1)$ satisfying the following cost-parity condition:

$$\frac{w_j^s}{A^s} = \frac{v}{A^k \psi(\widetilde{\beta}_j^k)} \tag{13}$$

With our functional form for task-level AI productivity $\psi(i)$, the unique automation threshold is given by:

$$\widetilde{\beta}_j^k = \left(\frac{A^k w_j^s}{A^s v}\right)^{\theta} \tag{14}$$

The city-specific threshold $\widetilde{\beta}_j^k$ denotes the task index at which the unit cost of producing with AI software equals that of skilled labor. For all tasks $i \in [0, \widetilde{\beta}_j^k]$, we have $\frac{1}{\psi(\widetilde{\beta}_j^k)} < \frac{A^k w_j^s}{A^s v}$, implying that these tasks are cheaper to perform using AI software. The derivation of this automation threshold is provided in Appendix C.1.

We impose the following assumption to guarantee that the local automation threshold is unconstrained by the global technological frontier:

Assumption 1 (Slack Automation Frontier)**.** *The economic profitability threshold is strictly below the technological automation frontier, $\widetilde{\beta}_j^k$, for all cities:*

$$\widetilde{\beta}_j^k < \beta^k, \quad \textit{for all } j \in \mathcal{D}$$

Under Assumption 1, any expansion in $\widetilde{\beta}_j^k$ driven by cost-minimization immediately reallocates tasks from skilled labor to AI software.

Solving the firm's cost minimization problem and integrating over tasks yields the composite AI–skilled input as:

$$ES_j = \left[\left(\left(\widetilde{\Gamma}_j^k \right)^{\frac{1-\rho}{\rho}} K_j \right)^{\rho} + \left((1 - \widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j \right)^{\rho} \right]^{\frac{1}{\rho}} \tag{15}$$

where $\widetilde{\Gamma}_j^k = \frac{\left(\widetilde{\beta}_j^k\right)^{1-\frac{\rho}{\theta(1-\rho)}}}{1-\frac{\rho}{\theta(1-\rho)}}$ is a closed-form expression for the aggregate task share of AI software, with the derivation provided in Appendix C.1.

Substituting this expression into the outer CES function gives the final aggregate production function:

$$Y_j = A_j \left\{ \eta_j^u (U_j)^{\sigma} + \eta_j^s \left[\left(\left(\widetilde{\Gamma}_j^k \right)^{\frac{1-\rho}{\rho}} A^k K_j \right)^{\rho} + \left((1 - \widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j \right)^{\rho} \right]^{\frac{\sigma}{\rho}} \right\}^{\frac{1}{\sigma}} \tag{16}$$

where $K_j = \int_0^{\widetilde{\beta}_j^k} k(i)\, di$ and $S_j = \int_{\widetilde{\beta}_j^k}^1 l(i)\, di$. The parameter $1/(1-\rho)$ represents the elasticity of substitution between AI and skilled labor, while $1/(1-\sigma)$ denotes the elasticity of substitution between unskilled labor and the AI–skilled labor composite.

By nesting AI software with skilled labor, our specification allows for differential substitution between AI and heterogeneous labor. We then structurally estimate the two key elasticities, ρ and σ, in Section 4.1. If $\rho < \sigma$, AI exhibits capital-skill complementarity, consistent with technological change such as IT. In contrast, $\rho > \sigma$ implies that AI acts as a relative substitute for skilled labor and a relative complement to unskilled labor. Emerging evidence shows that skilled labor is more exposed to AI (Copestake et al., 2023; Restrepo, 2024), which implies that AI could be fundamentally different from prior skill-biased technical change.

3.2 Worker

Workers of skill type $e \in \{u, s\}$ born in city i make a nested migration decision (Imbert et al., 2023). In the upper nest, they choose whether to stay in their origin city i or migrate to another city j, drawing idiosyncratic preference shocks from a generalized extreme value distribution with shape parameter γ^e. In the lower nest, conditional on migrating, they choose a specific destination city j, drawing from a generalized extreme value distribution with shape parameter η^e. Figure 2 illustrates the nested structure of workers' location choices.

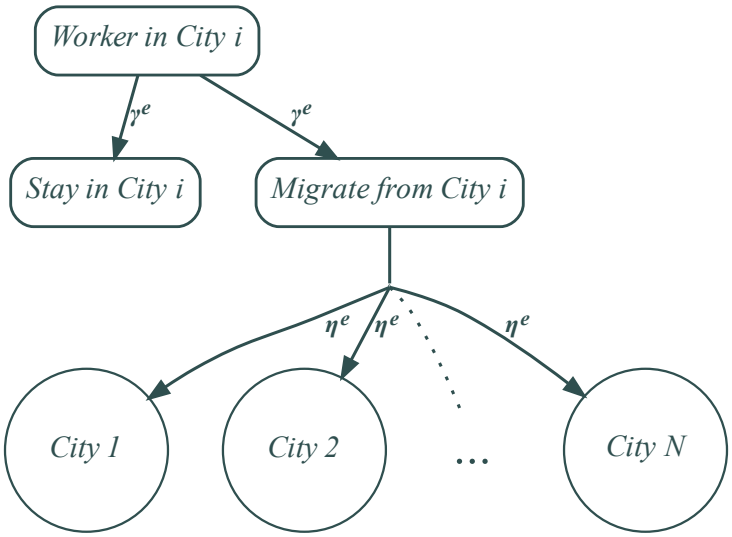


Figure 2: Nested Location Choice

The indirect utility of a type e worker from origin city i who chooses to migrate to destination city j is given by:

$$\ln V_{ij}^e = \ln w_j^e - \alpha_h^e \ln r_j - \tau_{ij} + \epsilon_{ij}^{e,\text{lower}} \tag{17}$$

where w_j^e denotes the wage of type e workers in city j, α_h^e is the share of expenditures on housing for type e workers, r_j is the housing price in city j, $\epsilon_{ij}^{e,\text{lower}}$ denotes the idiosyncratic preference shock associated with choosing destination city j, conditional on migrating from origin city i.

The indirect utility of a type e worker who stays in their origin city i is given by:

$$\ln V_{ii}^e = \ln w_i^e - \alpha_h^e \ln r_i + \epsilon_{ii}^{e,\text{upper}} \tag{18}$$

where $\epsilon_{ij}^{e,\text{upper}}$ denotes the idiosyncratic preference shock governing the upper-level decision of whether to out-migrate or to remain in the origin city.

By the properties of law of large numbers, the unconditional probability π_{ij}^e that a type e worker born in city i migrates to city j is given by:

$$\pi_{ij}^e = \left(\frac{V_{J,i}^e}{V_i^e}\right)^{\gamma_e} \cdot \left(\frac{V_{ij}^e}{V_{J,i}^e}\right)^{\eta_e} \tag{19}$$

where π_{ij}^e can be decomposed to two components. The first component $\left(\frac{V_{J,i}^e}{V_i^e}\right)^{\gamma_e}$ is the probability that a type e worker born in city i chooses to migrate rather than stay. The second component $\left(\frac{V_{ij}^e}{V_{J,i}^e}\right)^{\eta_e}$ is the probability of selecting destination j conditional on migrating from city i.

The term $V_{J,i}^e$ denotes the inclusive value for a type e worker migrating from city i, defined as the CES aggregator of indirect utilities across all potential destinations $k \neq i$:

$$V_{J,i}^e = \left[\sum_{k \in \mathcal{D}/i} \left(V_{ik}^e\right)^{\eta^e}\right]^{\frac{1}{\eta^e}} \tag{20}$$

where V_{ik}^e is the indirect utility of a type e worker from origin city i who chooses to migrate to destination city k. η^e governs the responsiveness to destination choice conditional on migration.

The term V_i^e denotes the inclusive value for a type e worker born in city i, aggregating the indirect utility of staying (V_{ii}^e) and the inclusive value of migrating $(V_{J,i}^e)$:

$$V_i^e = \left[(V_{ii}^e)^{\gamma^e} + (V_{J,i}^e)^{\gamma^e}\right]^{\frac{1}{\gamma^e}} \tag{21}$$

The parameter γ^e governs the responsiveness of the migration decision in the upper nest (stay versus migrate).

The population in each destination city j is the sum of inflows from all origin cities i, weighted by the migration shares:

$$\begin{aligned} S_j &= \sum_{i \in O} S_{i,0} \pi_{ij}^s, \\ U_j &= \sum_{i \in O} U_{i,0} \pi_{ij}^u. \end{aligned} \tag{22}$$

where $S_{i,0}$ and $U_{i,0}$ denote the initial populations of skilled and unskilled workers in city i, respectively. π_{ij}^s and π_{ij}^u denote the migration shares of skilled and unskilled workers between city i and city j.

3.3 Housing Market

We model housing as a local non-tradable good, where housing prices are determined by the intersection of local demand and an upward-sloping supply curve (Glaeser et al., 2006; Saiz, 2010).

Housing Supply We assume a perfectly competitive housing supply market, where the equilibrium price equals the marginal cost of providing housing. The housing supply

function is given by:

$$HS_j = r_j^{\mu} \tag{23}$$

where HS_j denotes the housing supply in city j, r_j represents the housing price, and μ is the housing supply elasticity. A higher value of μ indicates a more elastic supply.

Housing Demand Housing demand is derived from the utility maximization problem of workers under Cobb-Douglas preferences. Total demand in each city is therefore equal to the aggregate housing expenditure of all residents divided by the local housing price:

$$HD_j = \left(\frac{\alpha_h^s w_j^s S_j + \alpha_h^u w_j^u U_j}{r_j} \right) \tag{24}$$

Equilibrium Housing Prices Market clearing requires that housing demand equals housing supply $HD_j = HS_j$. Substituting the demand expression into the supply equation yields the equilibrium housing price:

$$r_j = \left(\alpha_h^s w_j^s S_j + \alpha_h^u w_j^u U_j \right)^{\frac{1}{1+\mu}} \tag{25}$$

Taking logarithms, the equilibrium housing price becomes:

$$\ln r_j = \frac{1}{1+\mu} \ln \left(\alpha_h^s w_j^s S_j + \alpha_h^u w_j^u U_j \right) \tag{26}$$

4 Structural Estimation

In this section, we structurally estimate the key parameters of the theoretical model. Specifically, we estimate the elasticity of substitution between AI and heterogeneous labor, labor supply elasticities, model-implied migration costs. Additionally, we recover underlying economic fundamentals, such as city-specific productivity. We calibrate the remaining parameters using estimates from the existing literature.

4.1 Labor Demand

The objective of this section is to estimate two key elasticities in the production function: the elasticity of substitution between unskilled labor and the skilled labor–AI composite, given by $1/(1-\sigma)$, and the elasticity of substitution between skilled labor and AI software, given by $1/(1-\rho)$. To identify these parameters, we exploit exogenous variation in labor supply shocks, following the approach of Baum-Snow et al. (2018).

By solving the representative firm's first-order conditions under perfect competition in the final goods market, we derive a system of equations characterizing the relationships among the capital–output ratio, skilled wages, and the skill premium in Appendix C.3.

We leverage exogenous variation in the relative supply of skilled labor ($\Delta \ln (S_j/U_j)$) to identify the labor demand elasticities by jointly estimating the following system:

$$\begin{cases} \Delta_t \ln\left(\dfrac{K_j}{Y_j}\right) = \beta_1 + \dfrac{(\rho-\sigma)(1-\omega_j^k)(1-\omega_j^{ks})}{Z_j} \cdot \Delta_t \ln\left(\dfrac{S_j}{U_j}\right) + \epsilon_{jt}^1 \\ \Delta_t \ln\left(w_j^s\right) = \beta_2 + \dfrac{(1-\rho)(1-\sigma)(1-\omega_j^{ks})}{Z_j} \cdot \Delta_t \ln\left(\dfrac{S_j}{U_j}\right) + \epsilon_{jt}^2 \\ \Delta_t \ln\left(\dfrac{w_j^s}{w_j^u}\right) = \beta_3 + \left[(\sigma-1) + \dfrac{(\sigma-\rho)(1-\sigma)\omega_j^k(1-\omega_j^{ks})}{Z_j}\right] \cdot \Delta_t \ln\left(\dfrac{S_j}{U_j}\right) + \epsilon_{jt}^3 \end{cases} \tag{27}$$

where $Z_j = (\sigma-\rho)\omega_j^k(1-\omega_j^{ks}) - (1-\rho)(1-\omega_j^k\omega_j^{ks})$, ω_j^k denotes the share of AI software within the AI–skilled nest in city j, and ω_j^{ks} represents the share of expenditure on the AI–skilled composite relative to total cost in city j.

Given China's rapid acceleration in AI adoption after 2015, we use data from two periods: 2012–2013 (Period 1) and 2017–2018 (Period 2) to construct AI software K_j, skilled wages w_j^s, and unskilled wages w_j^u.[8] We measure the skill ratio using population census data from 2010 and 2015.[9]

Estimating the elasticities of substitution between ρ and σ requires an instrumental variable that provides exogenous variation in the relative labor supply without being confounded by unobserved technological demand shocks. In Appendix C.3, we fully derive the components of the labor demand system from (27). This decomposition reveals that local labor demand is driven not only by the skill ratio, but also by technological shifts, including factor-augmenting technical change and AI task displacement.

To isolate exogenous variation in changes to the skill ratio that is orthogonal to AI-related technological change, we construct two separate instrumental variables based on source pollution shocks and agricultural shocks. We instrument for skilled labor supply using thermal inversion shocks as in Khanna et al. (2025), exploiting exogenous variation in skilled out-migration that is orthogonal to firms' AI investment. For unskilled labor, we construct a shift-share instrument based on agricultural shocks from Imbert et al. (2022). Specifically, we interact historical migration shares between city pairs with agricultural shocks in the origin city to instrument for unskilled migration flows, exploiting income shocks to rural workers in origin locations. The instrument is defined as follows:

$$\Delta_t \ln\widehat{(U_j)} = \ln\left(\sum_{i \in O \setminus \{j\}} \lambda_{ij}\, \Delta z_i\right) \tag{28}$$

[8] To recover wage levels, we combine skill premia estimated from Mincer regressions in Appendix B.3, using the CMDS dataset, with total wage bill data at the city level from the China Urban Construction Statistical Yearbook.

[9] Although the census data precede the estimation window, they represent the most recent period with comprehensive microdata, as population censuses in China are conducted every five years.

where Δz_i is the agricultural shock in origin city i, and λ_{ij} denotes the historical stock of migrants with agricultural hukou from i residing in j prior to 2000.

Combining the two sources of exogenous variation, the first-stage equation for the change in the relative skill supply is:

$$\Delta_t \ln\left(\frac{S_j}{U_j}\right) = \Delta_t \ln\widehat{(S_j)} + \Delta_t \ln\widehat{(U_j)} + \xi_j \tag{29}$$

where $\Delta_t \ln\widehat{(S_j)}$ and $\Delta_t \ln\widehat{(U_j)}$ denote the log changes in predicted skilled and unskilled labor inflows, respectively, and ξ_j denotes the error term. Table A1 presents the first-stage estimation results, using both instruments for the changes in the skill ratio.

Table 4: Labor Demand Elasticity Estimates

Parameter	Coefficient				
	(1) Inv. dist.	(2) Decay $d = 100$	(3) Decay $d = 200$	(4) Decay $d = 500$	(5) Global
$\widehat{\rho}$	0.846	0.843	0.833	0.816	0.774
	(0.061)	(0.064)	(0.068)	(0.079)	(0.116)
$\widehat{\sigma}$	0.393	0.428	0.402	0.334	0.304
	(0.252)	(0.218)	(0.216)	(0.247)	(0.297)
$\widehat{\rho} - \widehat{\sigma}$	0.454	0.416	0.431	0.481	0.471
	(0.217)	(0.173)	(0.169)	(0.201)	(0.258)
Method	GMM	GMM	GMM	GMM	GMM
Observations	123	123	123	123	123

Notes: Standard errors in parentheses. The regressions are weighted by the 2000 population of each city.

Table 4 presents our estimates of labor-demand elasticities. Column (1) uses our baseline AI software measure, which combines distance weighting with technological similarity. We find that $\widehat{\rho} = 0.846$ and $\widehat{\sigma} = 0.393$, and a formal test shows that $\widehat{\rho}$ is significantly greater than $\widehat{\sigma}$ ($p < 0.05$). This implies that AI is a relative complement to unskilled labor and a relative substitute for skilled labor. Columns (2)–(5) report estimates using alternative exponential distance-decay specifications and a specification that drops distance decay.

Appendix Table A2 reports results using a leave-one-out AI-software measure that excludes grants to universities in the firm's city, as described in Section 2.2.3. The corresponding estimates are reported there. The relative magnitudes of the estimated elasticities of substitution remain unchanged.

Our estimates indicate that AI technology differs fundamentally from standard capital-skill complementarity (Krusell et al., 2000). We show that even before the emergence of Large Language Models, AI possessed the unique capacity to automate both routine and

non-routine cognitive tasks traditionally performed by skilled labor, acting as a relative substitute for skilled labor at the macroeconomic level.

4.2 Labor Supply

Workers are born with type $e \in \{s, u\}$ in origin city i, and make a nested location choice. In the upper nest, they choose whether to stay in city i or migrate to another city. This decision is governed by the upper nest elasticity parameters $\{\gamma^s, \gamma^u\}$. Conditional on choosing to migrate, workers then select a destination city j in the lower nest, with this choice governed by the lower nest elasticity parameters $\{\eta^s, \eta^u\}$.

Lower Nest We take the logarithm of the second component of the equation (19):

$$\ln\left(\pi^e_{ij}|\text{ migrate from } i\right) = -\eta^e \ln V^e_{J,i} + \eta^e \left[\ln w^e_j - \alpha^e_h \ln r_j - \tau^e_{ij}\right] \tag{30}$$

where π^e_{ij} denotes the share of type e workers who migrate from origin city i to destination city j, conditional on choosing to migrate from i. w^e_j and α^e_h denote the nominal wage and housing price for type e workers in city j, respectively. The term $V^e_{J,i} = \left[\sum_{k \in \mathcal{D} \setminus i} \left(V^e_{ik}\right)^{\eta^e}\right]^{\frac{1}{\eta^e}}$ represents the inclusive value of migration for a type e worker in city i, capturing the aggregate attractiveness of all potential destination cities. Additionally, τ^e_{ij} denotes bilateral migration costs and other unobserved factors that affect migration decisions.

To estimate labor supply elasticities, we control for origin-city fixed effects and estimate a Poisson regression model in (30). This approach has the advantage of rationalizing the high frequency of zero migration flows between cities.

A key challenge in identifying labor supply elasticities is the presence of unobserved factors that jointly determine wages and migration decisions, potentially leading to biased estimates. To address this endogeneity, we exploit China's WTO accession as an exogenous trade shock, following Facchini et al. (2019a). We also interact this trade shock with the city-level share of undevelopable land to capture spatial heterogeneity in housing supply constraints. The intuition is straightforward: even if two cities experience an identical trade-induced shock to nominal wages, their real wage responses will differ. In cities with a more inelastic housing supply, nominal wage gains are largely offset by rising housing costs.

Table 5: Estimates of Lower Nest Parameters

	$\ln \pi^e_{ij}$			$\ln w^e_j - \alpha^e_h \ln r_j$
	Poisson		Poisson-IV	First stage-OLS
	(1)	(2)	(3)	(4)
Panel A: Skilled Labor				
$\ln w^s_j - \alpha^s_h \ln r_j$	2.171	2.389	4.558	
	(0.166)	(0.175)	(1.417)	
NTR gap$_j$				0.890
				(0.006)
NTR gap$_j \times x^{\text{unav}}_j$				-0.304
				(0.015)
Observations	23,560	23,560	23,560	23,560
Panel B: Unskilled Labor				
$\ln w^u_j - \alpha^u_h \ln r_j$	1.929	1.990	4.339	
	(0.269)	(0.286)	(1.297)	
NTR gap$_j$				0.906
				(0.005)
NTR gap$_j \times x^{\text{unav}}_j$				-0.491
				(0.011)
Observations	23,560	23,560	23,560	23,560
Controls	N	Y	Y	Y

Notes: This table reports estimates of labor supply elasticities. Columns 1–3 present Poisson and Poisson-IV estimates of the migration share equations, while Column 4 reports first-stage OLS estimates for the real wage instruments. All specifications include origin-city fixed effects. Standard errors are reported between parentheses and clustered at the level of destination cities. The regressions are weighted by the 2000 population of each city. Workers are classified as skilled if they hold a junior college degree or higher, and unskilled otherwise (high school or below). Migration shares are calculated using data from the 2015 population mini-census; wage data are drawn from the CMDS; city-level wage bill data from the China Urban Construction Statistical Yearbook (CUCSY); housing price data from Xitai (2015); and housing expenditure shares from Chen et al. (2023). The real wages of skilled and unskilled labor are instrumented using a trade shock interacted with local land unavailability. The trade shock is constructed following the methodology of Facchini et al. (2019a). Land unavailability is measured using Moderate Resolution Imaging Spectroradiometer (MODIS) data for 2001 (Wu and You, 2025), classifying land use types such as forest, wetland, water, and snow or ice as undevelopable.

Table 5 reports our estimates of the lower-nest labor supply elasticities. We find that $\eta^s > \eta^u$, indicating that skilled workers are more responsive to cross-city differences in real wages when making migration decisions.

Upper Nest Taking the logarithm of the first component of equation (19) yields:

$$\ln\left(\frac{1-\pi^e_{ii}}{\pi^e_{ii}}\right) = \gamma^e \ln\left(\frac{V^e_{J,i}}{V^e_{ii}}\right) \tag{31}$$

where π^e_{ii} denotes the probability that a worker of type e stays in their origin city i. The indirect utility of staying, V^e_{ii}, is given by $\left(\ln w_i - \alpha^e_h \ln r_i\right)$. The indirect utility of migrating, $V^e_{J,i}$, can be recovered from the origin-city fixed effects in equation (30).

To identify the upper nest labor supply elasticities, we exploit exogenous variation in labor demand driven by firm level productivity shocks. Using annual manufacturing survey data, we construct a balanced panel of firms over the period 2011 to 2015 and estimate total factor productivity at the two digit industry level following Levinsohn and Petrin (2003). We then construct a city-level productivity instrument by aggregating these firm-specific TFP measures, weighted by total firm sales within each city. The intuition behind this identification strategy is that city-specific TFP shocks are orthogonal to unobserved factors that influence local workers' out-migration decisions, yet they affect these decisions through their impact on local real wages V^e_{ii}.

Similarly, we construct an instrument for the inclusive value $V^e_{J,i}$ for migrants originating from city i. This construction follows a gravity-based form to remain consistent with the theoretical structure of the inclusive value. Specifically, we interact the destination-level TFP shock with the local share of undevelopable land, weighting by the inverse of geographical distance. The instrument is given by:

$$Z^J_i = \sum_{j \in \mathcal{D}} \left(\frac{1}{d_{ij}} \right) \times \left(\text{TFP}_j \times x^{\text{unav}}_j \right) \tag{32}$$

where d_{ij} denotes the geographic distance between cities i and j, and x^{unav}_j measures land unavailability in city j.

Table 6: Estimates of Upper Nest Parameters

	$\ln\left(\frac{1-\pi^e_{ii}}{\pi^e_{ii}}\right)$		$\ln\left(\frac{V^e_{J,i}}{V^e_{ii}}\right)$
	OLS	OLS-IV	First-stage OLS
	(1)	(2)	(3)
Panel A: Skilled Labor			
$\ln\left(\frac{V^e_{J,i}}{V^e_{ii}}\right)$	0.786	1.202	
	(0.105)	(0.330)	
TFP_i			-0.175
			(0.057)
$\sum_{j\in\mathcal{D}}\left(\frac{1}{d_{ij}}\right)\times\left(\text{TFP}_j\times x_j^{\text{unav}}\right)$			0.306
			(0.386)
First-stage F Stat			6.376
Observations	156	156	156
Panel B: Unskilled Labor			
$\ln\left(\frac{V^e_{J,i}}{V^e_{ii}}\right)$	0.468	1.477	
	(0.111)	(0.496)	
TFP_i			-0.172
			(0.046)
$\sum_{j\in\mathcal{D}}\left(\frac{1}{d_{ij}}\right)\times\left(\text{TFP}_j\times x_j^{\text{unav}}\right)$			0.147
			(0.286)
First-stage F Stat			8.772
Observations	156	156	156

Notes: Columns (1)–(2) report Poisson and Poisson-IV estimates, and column (3) reports first-stage OLS estimates for the instrument. Standard errors are reported in parentheses and clustered at the origin-city level. Regressions are weighted by city population in 2000. The Kleibergen–Paap F statistic is reported for IV estimates. The instruments are constructed by aggregating firm-level total factor productivity (TFP), estimated via the Levinsohn–Petrin method, using firm output within each city as weights.

Table 6 reports our estimates of the upper-nest labor supply elasticities. We find that $\gamma^s < \gamma^u$, indicating that skilled workers are less responsive to real wage differentials across cities and therefore less likely to outmigrate compared to unskilled workers.

4.3 Other Parameters

Migration Cost τ^e_{ij} Migration costs are chosen to rationalize observed migration shares and real income differentials across cities. The derivations are provided in Appendix C.4.

Hukou Strigency $\beta^e_{\textbf{hukou}}$ We parameterize the stringency of the hukou system as part of the migration cost function:

$$\tau_{ij} = \beta_0 + \beta^e_{\text{hukou}} \times \text{HukouStringency}^e_j + \beta^e_{Dist} \times \ln \text{Distance}_{ij} + \epsilon_{ij} \tag{33}$$

where β^e_{hukou} denotes the elasticity of migration costs with respect to hukou stringency, and β^e_{Dist} represents the elasticity of migration costs with respect to geographical distance. $\ln \text{Distance}_{ij}$ denotes denotes the log geographical distance between origin city i and destination city j, and $\text{HukouStringency}^e_j$ is the measure of hukou strictness for city j. Details on the institutional context of China's hukou system and the estimation of β^e_{hukou} are provided in Appendix D.2.

Housing Supply Elasticities μ In Appendix C.2, we derive the relationship between the housing supply elasticity and the land share in housing production. Using the estimate of 0.32 for the average land share reported by Chen et al. (2023), we obtain a housing supply elasticity of $\mu = 2.125$.

Agglomeration $\lambda^s, \lambda^u, \phi^s, \phi^u$ We calibrate the agglomeration force parameters using estimates from Combes et al. (2020), obtaining $\{\lambda^s, \lambda^u, \phi^s, \phi^u\} = \{0.0835, 0.0566, 0.144, 0.0728\}$.

Fundamentals ξ^s_j, ξ^u_j, A_j After calibrating the agglomeration parameters, we recover the fundamental productivity terms by inverting the model.

We first back out the city-specific distribution parameters using the first-order conditions of the representative firm:

$$\eta^s_j = \frac{\left(\frac{\text{Cost on AI-Skilled Sector}}{\text{Cost on Unskilled Labor Sector}}\right)^{1-\sigma}}{\left(\frac{w^u_j}{P^{es}_j}\right)^{\sigma} + \left(\frac{\text{Cost on AI-Skilled Labor Sector}}{\text{Cost on Unskilled Labor Sector}}\right)^{1-\sigma}}$$

$$\eta^u_j = \frac{\left(\frac{w^u_j}{P^{es}_j}\right)^{\sigma}}{\left(\frac{w^u_j}{P^{es}_j}\right)^{\sigma} + \left(\frac{\text{Cost on AI-Skilled Labor Sector}}{\text{Cost on Unskilled Labor Sector}}\right)^{1-\sigma}}$$

where P^{es}_j denotes the unit cost of the AI–skilled labor composite. Note that these expressions result from the normalization $\eta_{sj} + \eta_{uj} = 1$.

The corresponding unit cost for AI–skilled labor is given by:

$$P^{es}_j = \left(\widetilde{\Gamma}^k_j \left(\frac{v}{A^k}\right)^{\frac{\rho}{\rho-1}} + \left(1 - \widetilde{\beta}^k_j\right)\left(\frac{w^s_j}{A^s}\right)^{\frac{\rho}{\rho-1}}\right)^{\frac{\rho-1}{\rho}} \tag{34}$$

Finally, the fundamental productivity term A_j is recovered as the residual that rationalizes observed total costs given the production function.

Factor-specific Productivity A^s, A^k, θ To ensure that the aggregate task share of AI software is strictly positive, we require $\theta > \frac{\rho}{1-\rho}$. Given the estimated value $\hat{\rho} = 0.797$, this condition implies $\frac{\rho}{1-\rho} \approx 3.93$. We therefore set the parameter governing task specific productivity to $\theta = 4$.

We normalize the productivity of AI software to $A^k = 1$ and set the price of AI software to $v = 1$. We set the skill biased productivity parameter to $A^s = 10$ to ensure that the condition $\frac{A^k w_j^s}{A^s v} < 1$ holds in equilibrium. This restriction guarantees an interior solution in which not all tasks are automated by AI.

Task share of AI $\widetilde{\beta}_j^k$ We introduce a functional form for task-specific productivity $\psi(i)$ that allows the automation frontier to be endogenously determined by local skilled wages and technological conditions.

Table 7: Summary of Model Parameters

Variable	Description	Value	Source
α_h^s	Skilled housing share	0.24	(Chen et al., 2023)
α_h^u	Unskilled housing share	0.26	(Chen et al., 2023)
μ	Housing Supply Elasticity	2.125	(Chen et al., 2023)
λ^s	Agglomeration force of skilled labor	0.0835	(Combes et al., 2020)
λ^u	Agglomeration force of unskilled labor	0.0566	(Combes et al., 2020)
ϕ^s	Agglomeration force of skilled labor	0.144	(Combes et al., 2020)
ϕ^u	Agglomeration force of unskilled labor	0.0728	(Combes et al., 2020)
v	The unit cost of AI software	1	Calibration
A^k	The productivity of AI software	1	Calibration
A^s	Skill-biased productivity	10	Calibration
θ	Parameter of task-specific productivity	4	Calibration
A_j	Fundamental Productivity	$[0.333, 2.576]$	Estimation
ξ_j^s	Skilled fundamental productivity	$[0.041, 0.128]$	Estimation
ξ_j^u	Unskilled fundamental productivity	$[0.382, 0.750]$	Estimation
ρ	Elas. of sub. between K and S	0.846	Estimation
σ	Elas. of sub. between K or S and U	0.393	Estimation
$\widetilde{\beta}_j^k$	Task share of AI software	$[0, 0.444]$	Estimation
η^s	Skilled lower nest parameter	4.558	Estimation
η^u	Unskilled lower nest parameter	4.339	Estimation
γ^s	Skilled upper nest parameter	1.202	Estimation
γ^u	Unskilled upper nest parameter	1.477	Estimation
τ_{ij}^s	Skilled Migration Cost	$[0.748, +\infty]$	Estimation
τ_{ij}^u	Unskilled Migration Cost	$[0.964, +\infty]$	Estimation
β_{hukou}^s	Skilled hukou Elas. of Mig. costs	1.155	Estimation
β_{hukou}^u	Unskilled hukou Elas. of Mig. costs	1.633	Estimation

Notes: The table reports the summary of calibrated and estimated parameters.

Table 7 presents the summary of model parameters.

4.4 Model Fit

Our spatial equilibrium model consists of 110 cities. We initialize the economy using skill shares and population data from the 2010 Census. We then validate the model fit against the 2015 economy using data on GDP, AI patents, skill premiums, skill ratios, rents, and

population.[10] As shown in Figure A1, although the fit is not perfectly linear, the model moments are positively correlated with the data.

4.5 Robustness

In Appendix E, we assess the robustness of our model, measurement, and estimation. Appendix E.1 reports robustness checks to alternative migrant definitions, and Appendix E.2 presents additional checks that address potential unobserved heterogeneity by pooling unskilled workers across urban-hukou and agricultural-hukou statuses. Appendix E.3 verifies internal consistency across components of the structural estimation by examining whether assumptions imposed in one part of the procedure conflict with those used elsewhere. Finally, Appendix E.4 extends the model to include a non-tradable consumer-service sector and non-homothetic preferences over services.

5 Counterfactual Analysis

In this section, we conduct a series of counterfactual experiments within our spatial equilibrium framework to quantify the aggregate and distributional impacts of AI. Following the task-based framework of Acemoglu and Restrepo (2018, 2022), we decompose the AI into following technological shocks:

1. **The Extensive Margin of AI Automation:** represented as an expansion of the set of tasks performed by AI software, denoted by an increase in the automation frontier. Through this channel, AI takes over tasks previously performed by labor, effectively displacing workers to reduce production costs.

2. **The Intensive Margin of AI Automation:** represented as a uniform increase in the productivity of AI software in tasks that are already automated. This indicates existing AI becomes more efficient (eg. through improved algorithms or lower compute costs) without expanding the range of automated tasks.

By simulating these technological changes, we aim to: (1) understand how AI automation affects the spatial allocation of heterogeneous labor; (2) quantify the aggregate implications of this AI-induced spatial sorting; (3) disentangle the distinct mechanisms through which AI affects spatial sorting and aggregate productivity; and (4) evaluate the optimal design of the hukou system in the era of AI.

5.1 Measuring AI Automation

We measure the growth in AI software productivity between 2013 and 2018 by comparing AI performance relative to human benchmarks on three core tasks: image recognition,

[10] The skill premium is constructed using Mincerian regressions (CMDS 2015), and rent data is from Xitai.

speech recognition, and handwriting recognition. Task-level performance data are drawn from Kiela et al. (2023).

Table 8: Task Performance in 2013 and 2018

Task	2018	2013
Image recognition	111.69	55.84
Speech recognition	101.66	47.3
Handwriting recognition	102.67	98.67
Mean	105.34	67.27

Notes: The table reports benchmark performance scores for each AI task in 2013 and 2018. The bottom row presents the average performance across tasks.

Table 8 reports the relative task performance of AI compared with human benchmarks. Based on this comparison, we infer that AI software productivity (A_k) increased by approximately 56.6% over the period.

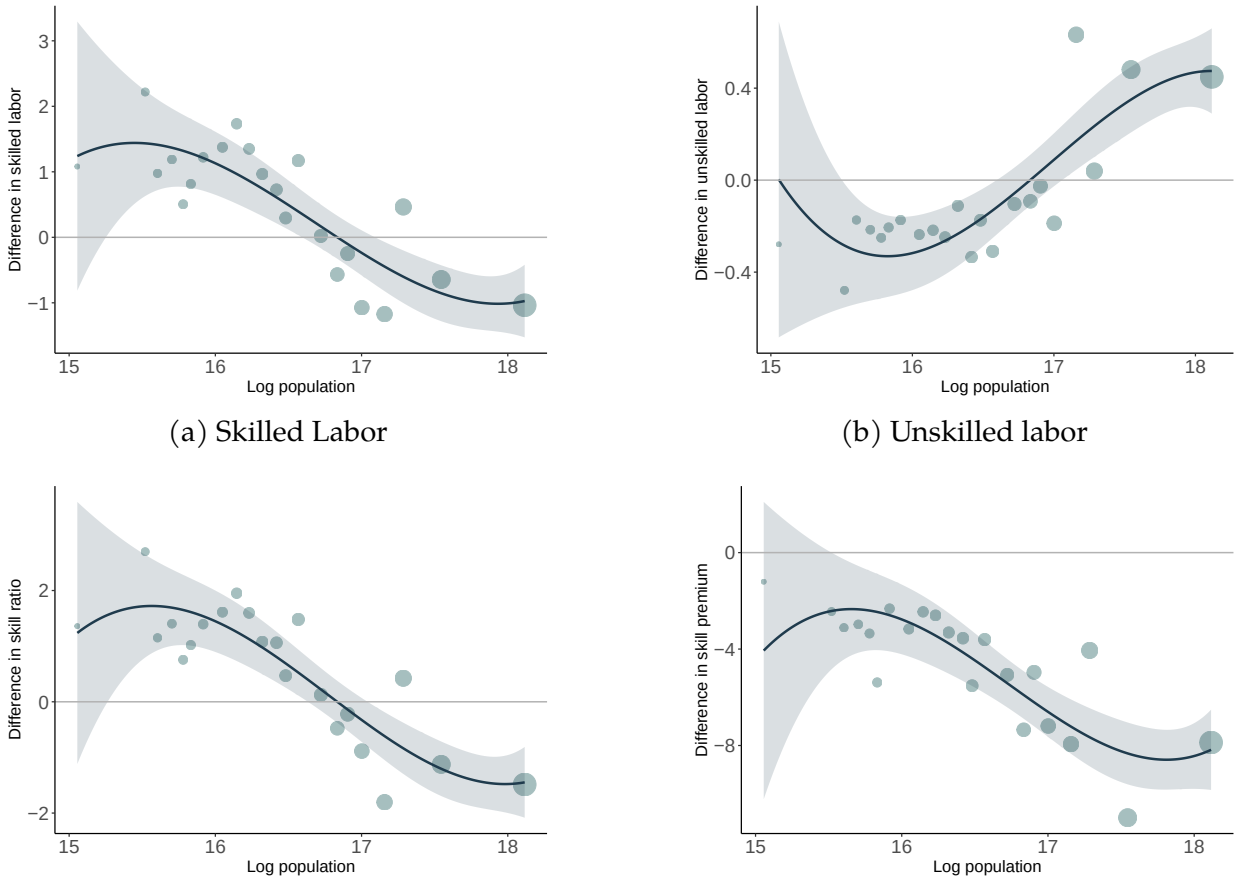
The growth in AI software productivity can generate two distinct effects. First, higher AI productivity increases the profitability of automation, thereby raising the equilibrium automation threshold $\widetilde{\beta}_j^k$. This corresponds to the extensive margin of automation, as a broader set of tasks becomes automated. Second, AI software may become more productive without expanding the range of automated tasks, representing the intensive margin of automation.

5.2 The Distributional Effects of AI

We begin by analyzing the distributional consequences of AI automation. The main objectives of this analysis are to: (1) document the patterns of AI-induced spatial sorting; and (2) quantify the unequal welfare gains resulting from this sorting. Understanding these distributional effects is crucial for identifying the mechanisms through which AI automation affects aggregate productivity.

5.2.1 Spatial Sorting

In this section, we analyze how AI automation affects the spatial allocation of labor and the skill premium across cities. We conduct a counterfactual simulation in which AI software productivity increases by 56.6% between 2013 and 2018, generating both an extensive and an intensive margin of AI automation.



(a) Skilled Labor

(b) Unskilled labor

(c) Skill Ratio

(d) Skill Premium

Figure 3: The Spatial Sorting Effects of AI Automation

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A polynomial regression is then fitted to the binned data.

Figure 3 presents the distributional effects of AI automation. We find that automation reverses the spatial sorting of skilled labor while driving unskilled labor into large cities. Consequently, AI automation lowers the skill ratio and the skill premium most significantly in these large urban centers.

We then decompose the effects of AI automation into extensive and intensive margins. To isolate the extensive margin when running the counterfactual increase in AI software productivity, A^k, we only allow the automation threshold $(\widetilde{\beta}_j^k)$ to adjust, reflecting the increased profitability of adopting AI for new tasks. Conversely, to evaluate the intensive margin, we hold the automation threshold $(\widetilde{\beta}_j^k)$ fixed at its baseline level and solely allow the productivity of existing AI software (A^k) to increase.

Figure A2 presents the distributional effects of the extensive margin of AI automation. Contrary to the overall results in Figure 3, we observe an out-migration of both skilled and unskilled labor from large cities. Furthermore, the extensive margin exhibits a much

smaller magnitude of effects, with a significantly narrower gap between large and small cities. This demonstrates that the overall impact of the extensive margin of AI automation is relatively small. In addition, Figure A3 presents the distributional effects of the intensive margin of AI automation, which are consistent with the result in Figure 3 in most cases.

A prominent distinction between the two types of AI automation is that the extensive margin tends to drive unskilled labor away from the largest cities, whereas the intensive margin incentivizes unskilled labor to resort to them. This difference is driven by the interplay between substitution and productivity effects of AI automation. At the extensive margin, as skilled labor is displaced by AI in the largest cities and returns to medium-sized and small cities, the relative agglomeration premium declines, lowering the marginal product of unskilled labor in the largest cities. Since AI at the extensive margin fails to generate sufficient compensating productivity gains, the substitution effect dominates, causing unskilled labor to also resort to smaller cities. Conversely, at the intensive margin, the productivity effects of AI automation dominate, attracting unskilled labor to the largest cities.

5.2.2 AI automation versus to RBTC (Eeckhout et al., 2026)

One may interpret our results through the lens of extreme-skill complementarity (Eeckhout et al., 2014) and IT-induced urban polarization (Eeckhout et al., 2026): as AI displaces routine cognitive tasks traditionally performed by middle-skill workers, unskilled workers increasingly co-locate with AI-skilled labor in large urban centers. These AI-skilled workers perform non-routine cognitive tasks, consistent with the rising share of non-routine cognitive work in large cities documented by Eeckhout et al. (2014, 2026). However, our results deliver novel implications for technology-induced spatial sorting and yield distinct predictions for aggregate and distributional outcomes.

Table A3 summarizes the core differences between our results and Eeckhout et al. (2026). First, we show that AI fundamentally reverses the canonical spatial sorting of skilled labor at the aggregate level. AI displaces cognitive workers and induces their outmigration from large cities, while unskilled labor that is relatively complementary to AI co-locates with AI software.[11] In contrast, Eeckhout et al. (2026) show that IT generates urban polarization: the share of routine cognitive labor declines, while non-routine cognitive and non-routine manual labor increase. This pattern reinforces canonical sorting, as high-skilled non-routine cognitive worker continues to agglomerate in large cities. One potential explanation for these contrasting findings is differences in baseline task composition between China and the United States.[12] Developing countries typically exhibit a larger initial share of routine cognitive employment, whereas developed economies have

[11] In the early stages of AI adoption, firms may require a complementary bundle of AI-skilled labor and AI-related investment, consistent with the Leontief structure in Acemoglu et al. (2022).

[12] Eeckhout et al. (2026) use evidence from the United States, whereas our paper uses evidence from China.

a much larger mass of non-routine cognitive workers. Consequently, even if AI primarily displaces routine cognitive tasks during our sample period, we can still observe a reversal of spatial sorting in equilibrium.

Second, their framework highlights that IT substitutes for routine work but is relatively complementary to non-routine work. In contrast, our results imply that AI software substitutes for cognitive labor without AI skills across both routine cognitive and non-routine cognitive tasks (Acemoglu et al., 2022; Copestake et al., 2023). Although the share of non-routine cognitive workers is small in our sample period, it varies across cities and contributes to our aggregate results on AI-induced reversal of spatial sorting. Some anecdotal evidence of AI adoption that substitutes for non-routine cognitive tasks can be found in China. For example, fintech firms have widely adopted machine-learning algorithms to detect fraud and assess creditworthiness, and some hospitals increasingly deploy computer-vision models to interpret medical images. More recently, advances in large language models may further expand the automation frontier into non-routine cognitive tasks. Accordingly, our results remain relevant under continued progress in AI automation.

Third, our paper investigates AI-induced spatial sorting using evidence from a developing country experiencing ongoing urbanization, whereas Eeckhout et al. (2026) use evidence from a developed economy with a steady-state spatial distribution of workers. This distinction has important implications for the long-run spatial equilibrium of developing economies. As developing countries urbanize toward a steady state under AI adoption, the resulting skill gradient across city sizes may differ markedly from that observed in developed economies. In particular, AI automation may permanently depress the skill ratio and skill-biased agglomeration in large cities.

5.2.3 Unequal Gains from AI

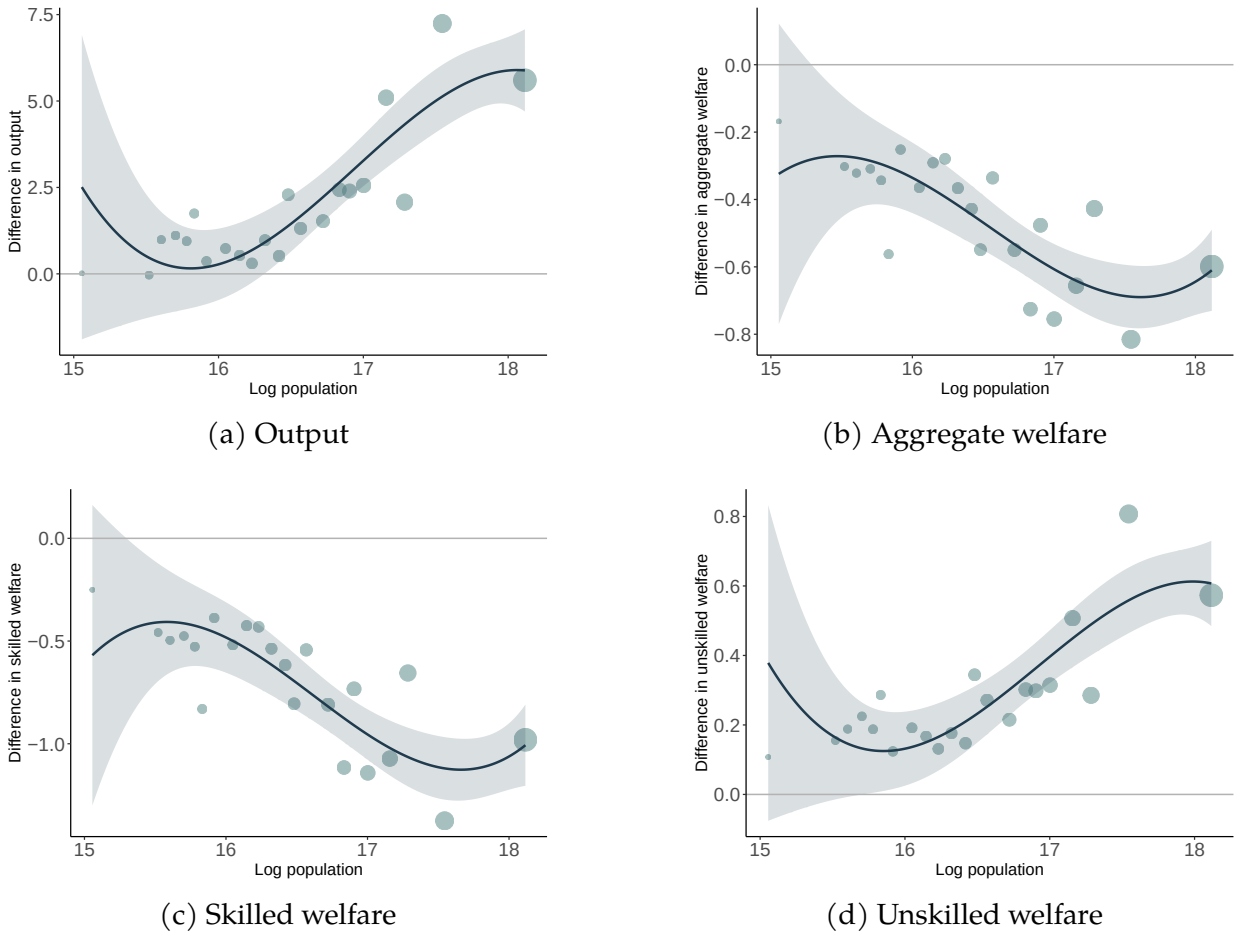
We then investigate the aggregate productivity and welfare gains across cities. Figure A7 presents the aggregate gains from AI automation. We find that while large cities experience the most substantial aggregate productivity gains, these benefits come at the expense of skilled worker welfare.

Combining this result with the distributional effects of AI automation, we show that aggregate productivity growth is primarily driven by the interaction between AI-driven relocation of unskilled labor toward large cities and local agglomeration forces. Although out-migration of skilled labor dampens agglomeration economies in large cities, these urban centers remain highly productive. Consequently, inflows of unskilled labor raise aggregate output. This mechanism is particularly relevant for developing countries with intermediate human capital, abundant unskilled labor, and a limited supply of skilled labor. In such settings, unskilled labor can remain a major contributor to growth, and co-location of AI and unskilled workers in large cities may be an important channel for

economic development in the AI era.

Another key implication is that the output gains from AI automation are non-linear across cities, where second-tier large cities experience the largest increase in output. This occurs because the migration costs to these second-tier cities are significantly lower than those to the largest cities. When workers make spatial resorting decisions, skilled labor is less likely to out-migrate from these cities, and unskilled labor is more likely to in-migrate, together contributing to aggregate productivity.

Figure A4 presents the gains from the extensive margin of AI automation across cities. It shows much smaller effects and relatively flat gains across city sizes. However, as shown in Figure A5, the aggregate gains from the intensive margin of AI automation yield implications similar to our overall effects of AI automation in Figure A7.



(a) Output

(b) Aggregate welfare

(c) Skilled welfare

(d) Unskilled welfare

Figure 4: The Unequal Gains from AI Automation

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A polynomial regression is then fitted to the binned data.

Linear Fit In Appendix A.1, we present counterfactual simulations using linear rather than polynomial fits. Polynomial regressions produce large standard errors in the sparse

tails of the covariate distribution. We therefore report linear fits as a robustness check. The qualitative conclusions are unchanged, which leaves our qualitative conclusions unchanged.

5.3 The Aggregate Implications of AI

Table 9 presents the aggregate implications of AI automation in spatial equilibrium. We again simulate an increase in AI productivity, decomposing the extensive and intensive margins by selectively shutting down each channel.

We find that AI automation increases aggregate productivity, reduces skilled welfare, and lowers the skill premium. We also show that both the extensive and intensive margins operate in the same direction. However, the aggregate effects are driven almost entirely by the intensive margin, while the extensive margin's contribution is quantitatively negligible.

Table 9: The Aggregate Implications of AI

	Y	$\mathcal{U}$	$\mathcal{U}_s$	$\mathcal{U}_u$	w_s	w_u	$\ln\left(\frac{w_s}{w_u}\right)$
Panel A. Baseline Estimates							
AI Aut	4.607%	−0.158%	−5.345%	2.336%	−5.547%	2.581%	−7.924%
AI Aut (Extensive)	0.013%	−0.051%	−0.168%	0.006%	−0.185%	0.002%	−0.187%
AI Aut (Intensive)	4.490%	−0.133%	−5.146%	2.278%	−5.337%	2.519%	−7.664%
Panel B. No Agglomeration							
AI Aut	0.013%	−0.003%	−0.036%	0.007%	−0.038%	0.008%	−0.046%
AI Aut (Extensive)	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
AI Aut (Intensive)	0.012%	−0.003%	−0.035%	0.007%	−0.037%	0.007%	−0.044%
Panel C. Migration Autarky							
AI Aut	4.486%	0.030%	−4.915%	2.415%	−5.010%	2.665%	−7.475%
AI Aut (Extensive)	0.021%	−0.038%	−0.145%	0.013%	−0.155%	0.013%	−0.167%
AI Aut (Intensive)	4.286%	−0.224%	−5.547%	2.343%	−5.706%	2.577%	−8.075%

To disentangle the underlying mechanisms, we conduct counterfactual simulations under two scenarios. First, we shut down endogenous agglomeration forces by setting $\{\lambda_s, \lambda_u, \phi_s, \phi_u\}$ to zero. Panel B of Table 9 shows that AI automation have only negligible aggregate effects. This result indicates that knowledge spillovers play a vital role in amplifying the aggregate productivity effects of AI-induced spatial sorting.

Second, we simulate a scenario of migration autarky by setting migration costs $\tau_{ij} = \infty$ for all city pairs. Panel C of Table 9 shows that the aggregate productivity effects of AI automation are smaller under migration autarky.

In addition, we find that the skill premium declines less under migration autarky in

response to AI automation than in the baseline model. This result contrasts with the standard intuition that preventing factor mobility exacerbates welfare losses from negative local shocks. The underlying mechanism is driven by spatial ripple effects. As AI displaces skilled workers in urban cognitive hubs, these workers resort into smaller, less productive cities. This dispersion undermines agglomeration economies in large cities while simultaneously generating a labor supply shock that depresses skilled wages in medium-sized cities. Together, the loss of agglomeration benefits and the spatial transmission of the wage shock cause a greater decline in the aggregate skill premium in the baseline model relative to migration autarky.

However, if we exclude the extensive margin of automation, the skill premium declines less in the baseline than under migration autarky. The intensive margin of AI automation raises productivity within the already-automated task set, making the productivity effect dominate. In this case, allowing migration helps reallocate skilled workers to arbitrage across locations where their marginal product is relatively higher, thereby dampening the aggregate decline in skilled wages.

5.4 The Effects of Reducing Migration Cost

Having documented the pattern of AI-induced spatial sorting, we then examine the role of migration costs and evaluate how hukou policy can be adjusted in response. Specifically, we simulate two counterfactual policy reforms: (1) a 50% reduction in hukou stringency for skilled labor in the 16 largest cities, and (2) a 50% reduction in hukou stringency for unskilled labor in the same cities. Details on the hukou stringency index and the elasticity between migration costs and hukou stringency are provided in Appendix D.2.

Following these policy changes, we recalculate migration costs using the following equation:

$$\tau_{ij}^{e'} = \tau_{ij}^{e} + \beta_{\text{hukou}}^{e} \times \Delta\text{HukouStringency}_{j}^{e} \tag{35}$$

We then re-run the counterfactual simulations to assess the aggregate implications of AI automation.

Table 10: The Aggregate Implications of Reducing Migration Cost

	Y	$\mathcal{W}$	$\mathcal{W}_s$	$\mathcal{W}_u$	w_s	w_u	$\ln\left(\frac{w_s}{w_u}\right)$
Panel A. Baseline Estimates							
AI Aut	4.607%	−0.158%	−5.345%	2.336%	−5.547%	2.581%	−7.924%
AI Aut (Extensive)	0.013%	−0.051%	−0.168%	0.006%	−0.185%	0.002%	−0.187%
AI Aut (Intensive)	4.490%	−0.133%	−5.146%	2.278%	−5.337%	2.519%	−7.664%
Panel B: Relax Skilled Hukou Restrictions							
AI Aut	4.892%	−0.134%	−5.546%	2.553%	−5.689%	2.823%	−8.278%
AI Aut (Extensive)	0.014%	−0.056%	−0.181%	0.005%	−0.197%	0.001%	−0.198%
AI Aut (Intensive)	4.767%	−0.107%	−5.338%	2.491%	−5.472%	2.756%	−8.007%
Panel C: Relax Unskilled Hukou Restrictions							
AI Aut	5.045%	−0.113%	−5.681%	2.599%	−5.852%	2.855%	−8.465%
AI Aut (Extensive)	0.014%	−0.058%	−0.187%	0.004%	−0.204%	0.000%	−0.204%
AI Aut (Intensive)	4.917%	−0.086%	−5.469%	2.536%	−5.631%	2.788%	−8.190%

Table 10 presents the aggregate effects of reducing hukou restrictions for skilled and unskilled labor. This analysis yields two key implications. First, because AI reverses traditional patterns of spatial sorting, relaxing hukou restrictions for unskilled labor in large cities is more effective than relaxing them for skilled labor. Second, under both types of policy implementation, economic surplus is reallocated from skilled to unskilled workers. In particular, relaxing hukou constraints reinforces the inequality-reducing effects of AI in terms of both overall welfare and wages.

6 Conclusion

In this paper, we document AI-induced spatial sorting and quantify its aggregate and distributional effects in China. We develop a theoretical framework by integrating a task-based model of AI into a spatial equilibrium model. We first estimate the elasticity of substitution between AI and heterogeneous labor, finding that AI is a relative substitute for skilled labor and a relative complement for unskilled labor. Building on these estimates, our paper provides a new perspective on spatial sorting and agglomeration. While the existing literature focuses on skill-biased agglomeration driving the concentration of skilled labor in large cities, we identify a novel channel through which AI reverses this pattern and reduces the skill premium most significantly in large cities. We also find that AI automation raises aggregate productivity through the interaction between agglomeration forces and AI-induced resorting of unskilled labor toward large cities.

We demonstrate that AI software and unskilled workers increasingly co-locate in large cities, while cognitive workers without AI skills are displaced and relocate toward smaller

cities.[13] This mechanism differs from the IT-driven displacement of routine cognitive workers in Eeckhout et al. (2026). Although our estimates primarily reflect variation from routine cognitive automation and only preliminary displacement effects in non-routine cognitive work, continued expansion of the AI automation frontier from routine to non-routine cognitive tasks may further strengthen the reversal of spatial sorting. Accordingly, our results remain applicable as large language models and other AI technologies extend automation more intensively into non-routine cognitive tasks.

Our results have two broader implications for developing countries. First, we highlight a potentially welfare-improving development strategy with AI-induced spatial sorting. In developing economies with abundant unskilled labor and intermediate skilled labor, the interaction between local agglomeration forces and AI-driven relocation of unskilled labor toward large cities can raise aggregate productivity. This mechanism also applies to emerging economies with intermediate human capital such as India and Brazil, where rapid growth in university-educated cohorts is expanding the supply of skilled labor entering the labor market.[14] Second, AI-induced spatial sorting may lower the skill ratio in the steady-state urban equilibrium, yielding large cities that are markedly less skill-intensive than those in developed economies. This perspective offers a new lens on urbanization in developing countries.

A limitation of our study is that we do not directly observe task-level substitution elasticities between AI and labor, nor do we differentiate between different occupations and AI. Future research using granular microdata could estimate AI–labor substitution elasticities at the occupation or task level, shedding light on the agglomeration mechanisms underlying AI-induced spatial sorting.

[13] As discussed above, we interpret AI software as a complementary bundle of AI-skilled labor and AI-related investment.

[14] Higher-education expansion has been rapid in both India and Brazil. In India, total enrollment in higher education rose from 27.5 million in 2010–11 to 41.4 million in 2020–21 (Ministry of Human Resource Development, Government of India, 2013; Ministry of Education, Government of India, 2023), while the annual number of students completing degrees reached 10.7 million by 2021–22 (Ministry of Education, Government of India, 2024). In Brazil, undergraduate enrollment increased from approximately 6.4 million in 2010 (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2011) to over 8.6 million in 2020 (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2022)

References

Acemoglu, Daron and Pascual Restrepo, "The race between man and machine: Implications of technology for growth, factor shares, and employment," *American economic review*, 2018, *108* (6), 1488–1542.

_ and _ , "Tasks, automation, and the rise in US wage inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

_ , David Autor, Jonathon Hazell, and Pascual Restrepo, "Artificial intelligence and jobs: Evidence from online vacancies," *Journal of Labor Economics*, 2022, *40* (S1), S293–S340.

Adamopoulos, Tasso, Loren Brandt, Chaoran Chen, Diego Restuccia, and Xiaoyun Wei, "Land security and mobility frictions," *The Quarterly Journal of Economics*, 2024, *139* (3), 1941–1987.

Autor, David H and David Dorn, "The growth of low-skill service jobs and the polarization of the US labor market," *American economic review*, 2013, *103* (5), 1553–1597.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan, "Why has urban inequality increased?," *American Economic Journal: Applied Economics*, 2018, *10* (4), 1–42.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying technology spillovers and product market rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

Brynjolfsson, Erik, Danielle Li, and Lindsey R Raymond, "Generative AI at work," Technical Report, National Bureau of Economic Research 2023.

Chen, Ziyang, Pierre-Philippe Combes, Sylvie Démurger, and Xiuyan Liu, "Housing Costs and Real Income Differences across Chinese Cities," 2023.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial wage disparities: Sorting matters!," *Journal of urban economics*, 2008, *63* (2), 723–742.

_ , _ , _ , Diego Puga, and Sébastien Roux, "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *Econometrica*, 2012, *80* (6), 2543–2594.

_ , **Sylvie Démurger, Shi Li, and Jianguo Wang**, "Unequal migration and urbanisation gains in China," *Journal of Development Economics*, 2020, *142*, 102328.

Copestake, Alexander, Max Marczinek, Ashley Pople, and Katherine Stapleton, "AI and services-led growth: Evidence from Indian job adverts," *World Bank Policy Research Working Paper*, 2023.

Davis, Donald R and Jonathan I Dingel, "A spatial knowledge economy," *American Economic Review*, 2019, *109* (1), 153–170.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, 2016, *106* (3), 479–524.

_ **and Cecile Gaubert**, "Spatial sorting and inequality," *Annual Review of Economics*, 2022, *14* (1), 795–819.

Dyevre, Arnaud, "Public R&D spillovers and productivity growth," Technical Report, Working paper. Eaton, J., & Kortum, S.(1999). International technology ···2023.

Eeckhout, Jan, Christoph Hedtrich, and Roberto Pinheiro, "IT and urban polarization," *American Economic Journal: Macroeconomics*, 2026, *18* (1), 223–259.

_ , **Roberto Pinheiro, and Kurt Schmidheiny**, "Spatial sorting," *Journal of Political Economy*, 2014, *122* (3), 554–620.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs are GPTs: Labor market impact potential of LLMs," *Science*, 2024, *384* (6702), 1306–1308.

Facchini, Giovanni, Maggie Y Liu, Anna Maria Mayda, and Minghai Zhou, "China's "Great Migration" : The impact of the reduction in trade policy uncertainty," *Journal of International Economics*, 2019, *120*, 126–144.

_ , **Maggie Y. Liu, Anna Maria Mayda, and Minghai Zhou**, "China' s "Great Migration" : The impact of the reduction in trade policy uncertainty," *Journal of International Economics*, 2019, *120*, 126–144.

Fan, Jingting, "Internal geography, labor mobility, and the distributional impacts of trade," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–288.

Felten, Edward, Manav Raj, and Robert Seamans, "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses," *Strategic management journal*, 2021, *42* (12), 2195–2217.

Gaubert, Cecile, "Firm sorting and agglomeration," *American Economic Review*, 2018, *108* (11), 3117–3153.

Giannone, Elisa, "Skilled-biased technical change and regional convergence," *University of Chicago. Unpublished manuscript*, 2017.

Glaeser, Edward L, Joseph Gyourko, and Raven E Saks, "Urban growth and housing supply," *Journal of economic geography*, 2006, *6* (1), 71–89.

Imbert, Clément, Joan Monras, Marlon Seror, and Yanos Zylberberg, "Floating population: migration with (out) family and the spatial distribution of economic activity," Technical Report 2023.

_ , Marlon Seror, Yifan Zhang, and Yanos Zylberberg, "Migrants and firms: Evidence from china," *American Economic Review*, 2022, *112* (6), 1885–1914.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, "Censo da Educação Superior 2010," Technical Report, Ministério da Educação, Brazil 2011.

_ , "Censo da Educação Superior 2020," Technical Report, Ministério da Educação, Brazil 2022.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 1993, *108* (3), 577–598.

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song, "The productivity consequences of pollution-induced migration in China," *American Economic Journal: Applied Economics*, 2025, *17* (2), 184–224.

Kiela, Douwe, Max Bartolo, Yixin Nie, Divyansh Kaushik, Atticus Geiger, Zhengxuan Wu, Bertie Vidgen, Grusha Prasad, Amanpreet Singh, Pratik Ringshia et al., "Dynabench: Rethinking benchmarking in NLP," *arXiv preprint arXiv:2104.14337*, 2021.

_ , Tristan Thrush, Kawin Ethayarajh, and Amanpreet Singh, "Plotting Progress in AI," *Contextual AI Blog*, 2023. https://contextual.ai/blog/plotting-progress.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The review of economic studies*, 2003, *70* (2), 317–341.

Ministry of Education, Government of India, "All India Survey on Higher Education 2020–21," Technical Report, Planning, Monitoring and Statistics Bureau, Ministry of Education 2023.

_ , "All India Survey on Higher Education (AISHE) 2021–22," Technical Report, Ministry of Education, Government of India 2024.

Ministry of Human Resource Development, Government of India, "All India Survey on Higher Education 2010–11," Technical Report, Department of Higher Education, Ministry of Human Resource Development 2013.

Moretti, Enrico, *The new geography of jobs*, Houghton Mifflin Harcourt, 2012.

Morten, Melanie and Jaqueline Oliveira, "The effects of roads on trade and migration: Evidence from a planned capital city," *American Economic Journal: Applied Economics*, 2024, *16* (2), 389–421.

Munshi, Kaivan and Mark Rosenzweig, "Networks and misallocation: Insurance, migration, and the rural-urban wage gap," *American Economic Review*, 2016, *106* (01), 46–98.

Noy, Shakked and Whitney Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence," *Science*, 2023, *381* (6654), 187–192.

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, and Mert Demirer, "The impact of ai on developer productivity: Evidence from github copilot," *arXiv preprint arXiv:2302.06590*, 2023.

Restrepo, Pascual, "Automation: Theory, evidence, and outlook," *Annual review of economics*, 2024, *16* (1), 1–25.

Roback, Jennifer, "Wages, rents, and the quality of life," *Journal of political Economy*, 1982, *90* (6), 1257–1278.

Rosen, Harvey S, "Housing decisions and the US income tax: An econometric analysis," *Journal of Public Economics*, 1979, *11* (1), 1–23.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Song, Yang, "What should economists know about the current Chinese hukou system?," *China Economic Review*, 2014, *29*, 200–212.

Tombe, Trevor and Xiaodong Zhu, "Trade, migration, and productivity: A quantitative analysis of China," *American Economic Review*, 2019, *109* (5), 1843–1872.

Webb, Michael, "The impact of artificial intelligence on the labor market," *Available at SSRN 3482150*, 2019.

Wu, Wenbin and Wei You, "Should governments promote or restrain urbanization?," *Journal of International Economics*, 2025, *155*, 104084.

Zhang, Jipeng, Ru Wang, and Chong Lu, "A quantitative analysis of Hukou reform in Chinese cities: 2000–2016," *Growth and Change*, 2019, *50* (1), 201–221.

Appendices

A Figures and Tables

Table A1: First-Stage Estimates: Labor Demand

Dependent Variable	$\Delta_t \ln\left(\frac{S_j}{U_j}\right)$		
	(1)	(2)	(3)
$\Delta_t \ln\left(\widehat{S_j}\right)$	-0.119	-0.171	
	(0.062)	(0.078)	
$\Delta_t \ln\left(\widehat{U_j}\right)$	0.236		0.338
	(0.122)		(0.134)
Period FE	Yes	Yes	Yes
Observations	123	123	123
Fist-stage F Stat	4.019	4.737	6.299
R-squared	0.071	0.054	0.054

Notes: This table reports the first-stage results of the labor demand estimation. The instrument for skilled labor is constructed using changes in thermal inversion days, following Khanna et al. (2025), while the instrument for unskilled labor is based on migration flows induced by agricultural shocks, following Imbert et al. (2022). Standard errors, clustered at the destination city level, are reported in parentheses. The regressions are weighted by the 2000 population of each city.

Table A2: Robustness: Labor Demand Elasticity Estimates

Parameter	Coefficient				
	(1) Inv. dist.	(2) Decay $d = 100$	(3) Decay $d = 200$	(4) Decay $d = 500$	(5) Global
$\widehat{\rho}$	0.814	0.856	0.841	0.815	0.769
	(0.082)	(0.056)	(0.063)	(0.080)	(0.121)
$\widehat{\sigma}$	0.328	0.416	0.372	0.310	0.303
	(0.257)	(0.213)	(0.221)	(0.257)	(0.299)
$\widehat{\rho} - \widehat{\sigma}$	0.486	0.440	0.469	0.505	0.466
	(0.211)	(0.170)	(0.177)	(0.213)	(0.263)
Method	GMM	GMM	GMM	GMM	GMM
Observations	123	123	123	123	123

Notes: Standard errors in parentheses. The regressions are weighted by the 2000 population of each city.

Table A3: Spatial Sorting under IT versus AI Automation

Dimension	Information Technology (IT)	Artificial Intelligence (AI)
Technology	Routine-biased technical change	Cognitive-task-biased automation
Elasticity of Substitution	1.25 in non-routine manual jobs, 2.5 in routine jobs, and 0.87 in non-routine cognitive jobs. IT is complementary to labor in non-routine cognitive occupations, while it substitutes for labor in routine cognitive and manual jobs.	4.93 for skilled workers and 1.53 for unskilled workers. AI is relatively substitutable with skilled labor and relatively complementary with unskilled labor.
High-skill spatial sorting	The marginal product of high-skill labor is higher in larger cities; with strong skill complementarity, this attracts high-skill workers to sort into large cities.	AI reduces the marginal product of high-skill labor, inducing high-skill workers to reallocate away from large cities.
Low-skill spatial sorting	Low-skill workers sort into large cities via local demand spillovers generated by high-skill workers.	Low-skill workers co-locate in large cities due to production complementarity with AI.
Spatial sorting	IT reinforces canonical sorting: middle-skill out-migration and high- and low-skill in-migration to large cities.	AI reverses canonical sorting: high-skill out-migration from large cities alongside low-skill in-migration.
Aggregate Trend	Developed economy with a high share of high-skill labor; aggregate sorting continues toward large cities.	Developing economy with a larger share of routine-cognitive workers within the skilled workforce and a smaller share of high-skill workers; aggregate spatial sorting reverses.

Table A4: Migration Shares by Skill Group

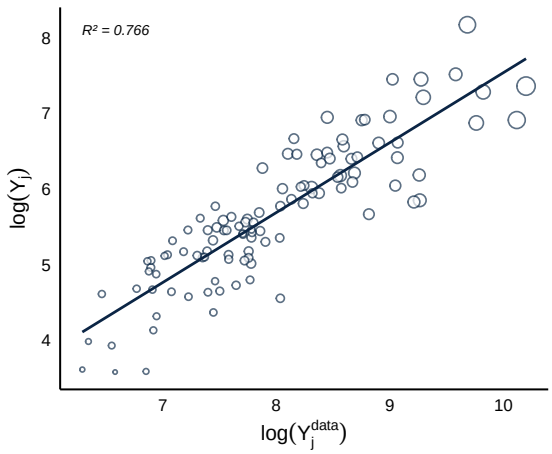
	Skilled		Unskilled	
Category	N	Share (%)	N	Share (%)
Panel A. Duration since hukou departure				
$\leq$ 1 year	2485	12.86	10461	15.62
1–5 years	7920	40.99	26121	39.01
5–10 years	4779	24.74	14003	20.91
$>$ 10 years	4136	21.41	16382	24.46
Panel B. Migration purpose				
Work/Job	12234	63.32	54594	81.52
Other	7086	36.68	12373	18.48
Panel C. Hukou origin				
Rural	4222	21.85	39096	58.38
Urban	15098	78.15	27871	41.62
Total	19320	100	66967	100

Notes: Migration types for skilled and unskilled workers are calculated using the 2015 Mini Census.

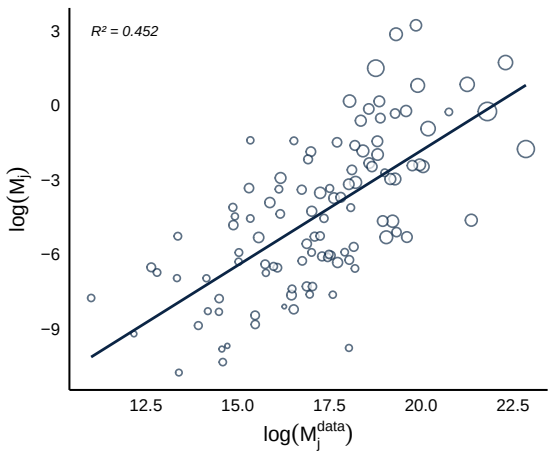
Table A5: Top 15 Industries for Migrants

Industry	Frequency	Share (%)
Wholesale and Retail Trade	15573	19.68
Accommodation and Catering	10568	13.35
Residential and Other Services	8010	10.12
Construction	6565	8.30
Miscellaneous Manufacturing	6518	8.24
Textiles and Garments	3882	4.91
Food Processing	2903	3.67
Transportation and Postal Services	2604	3.29
Computer and Electronics Mfg.	2442	3.09
Real Estate	2220	2.81
Electrical Machinery Mfg.	1767	2.23
Education	1639	2.07
Wood and Furniture	1393	1.76
IT and Software Services	1342	1.70
Agriculture, Forestry, and Fishery	1338	1.69
Total	79135	100

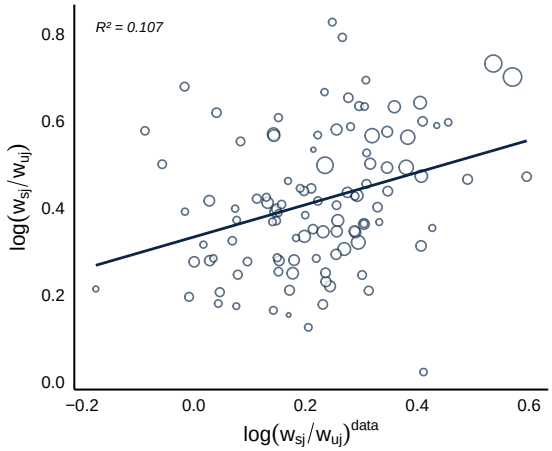
Note: This is calculated using data from the CMDS 2017. We retain all intercity migrants in the sample, restricting our analysis to those who migrated after 2010. The table presents the top 15 industries employing these migrants.



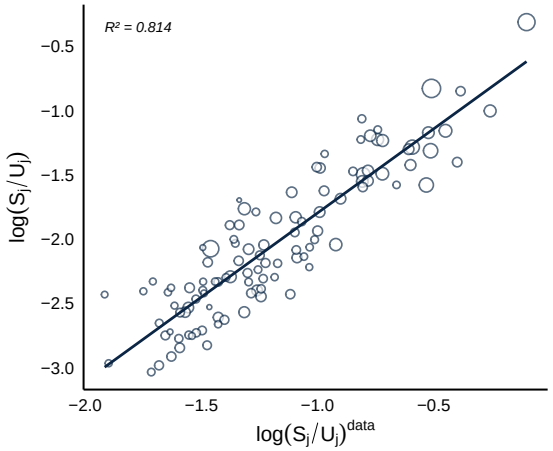
(a) Output



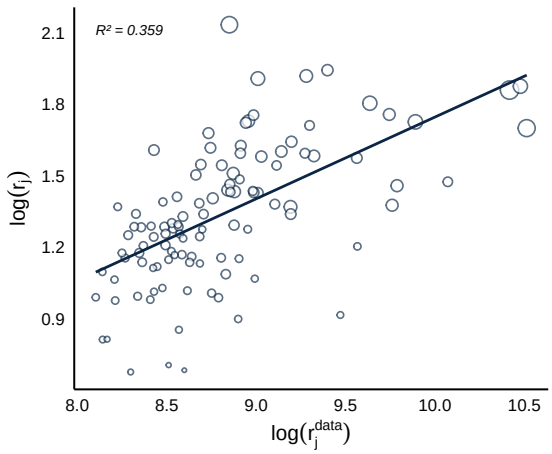
(b) AI software



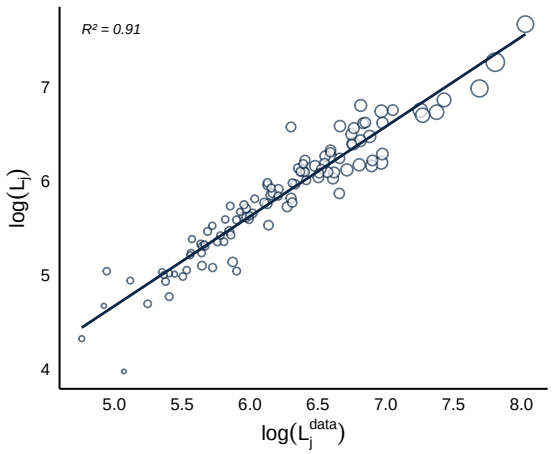
(c) Skill Premium



(d) Skill Ratio

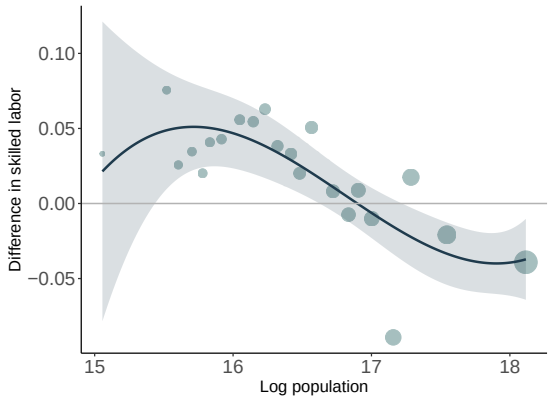


(e) Rent

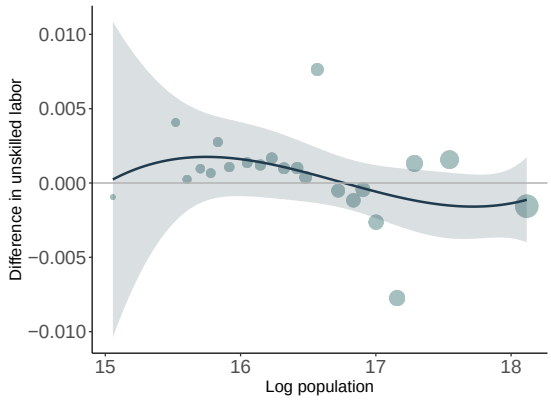


(f) Population

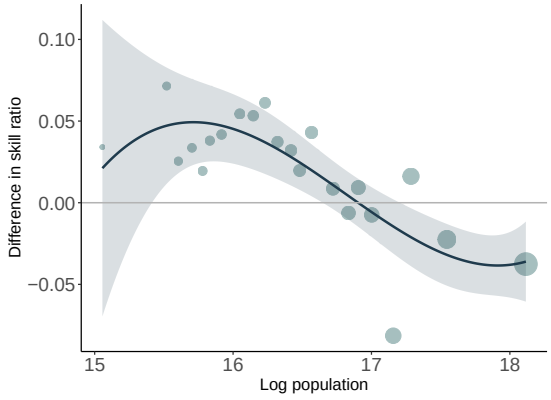
Figure A1: Model Fit



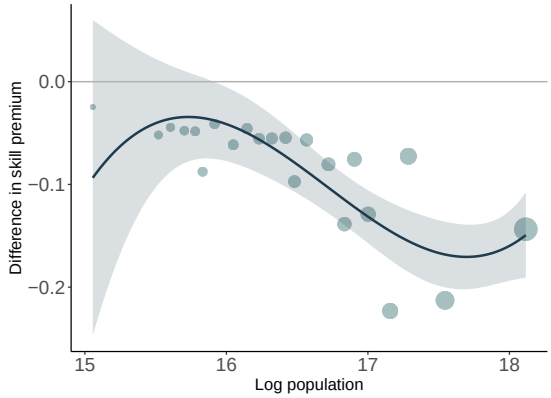
(a) Skilled Labor



(b) Unskilled labor



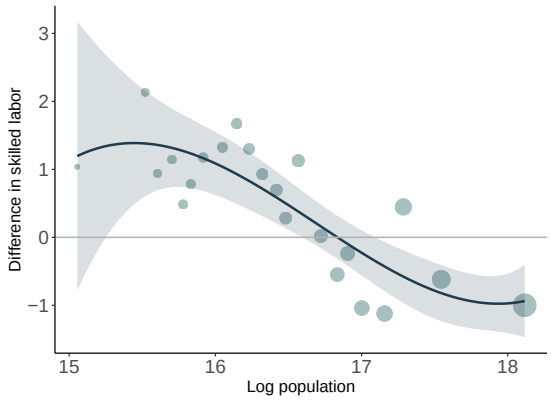
(c) Skill Ratio



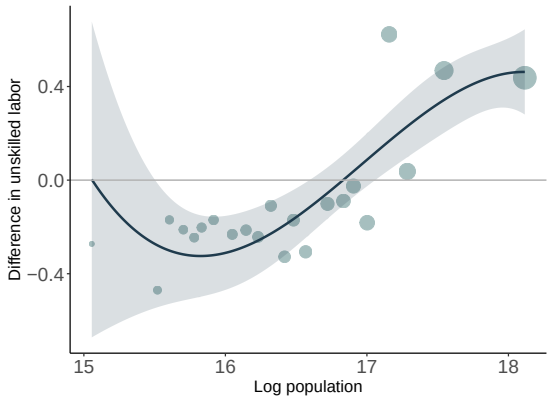
(d) Skill Premium

Figure A2: The Spatial Sorting Effects of the Extensive Margin of AI Automation

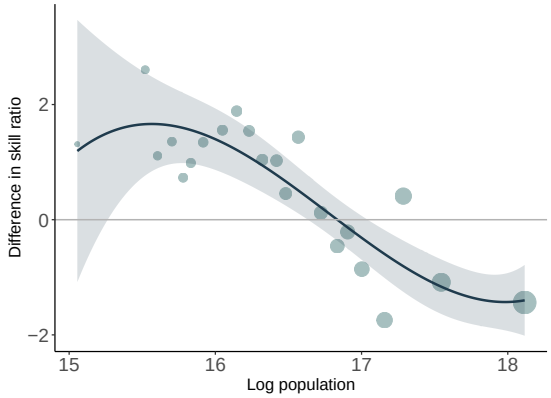
Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A polynomial regression is then fitted to the binned data.



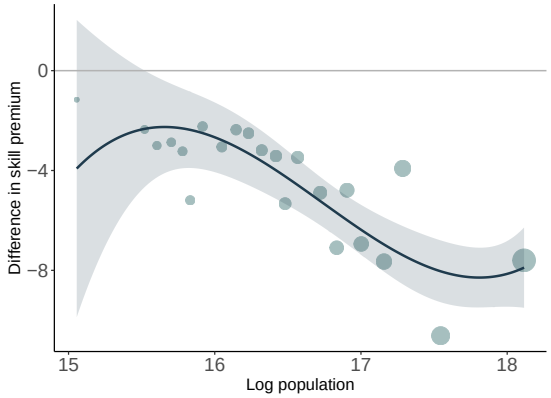
(a) Skilled Labor



(b) Unskilled labor



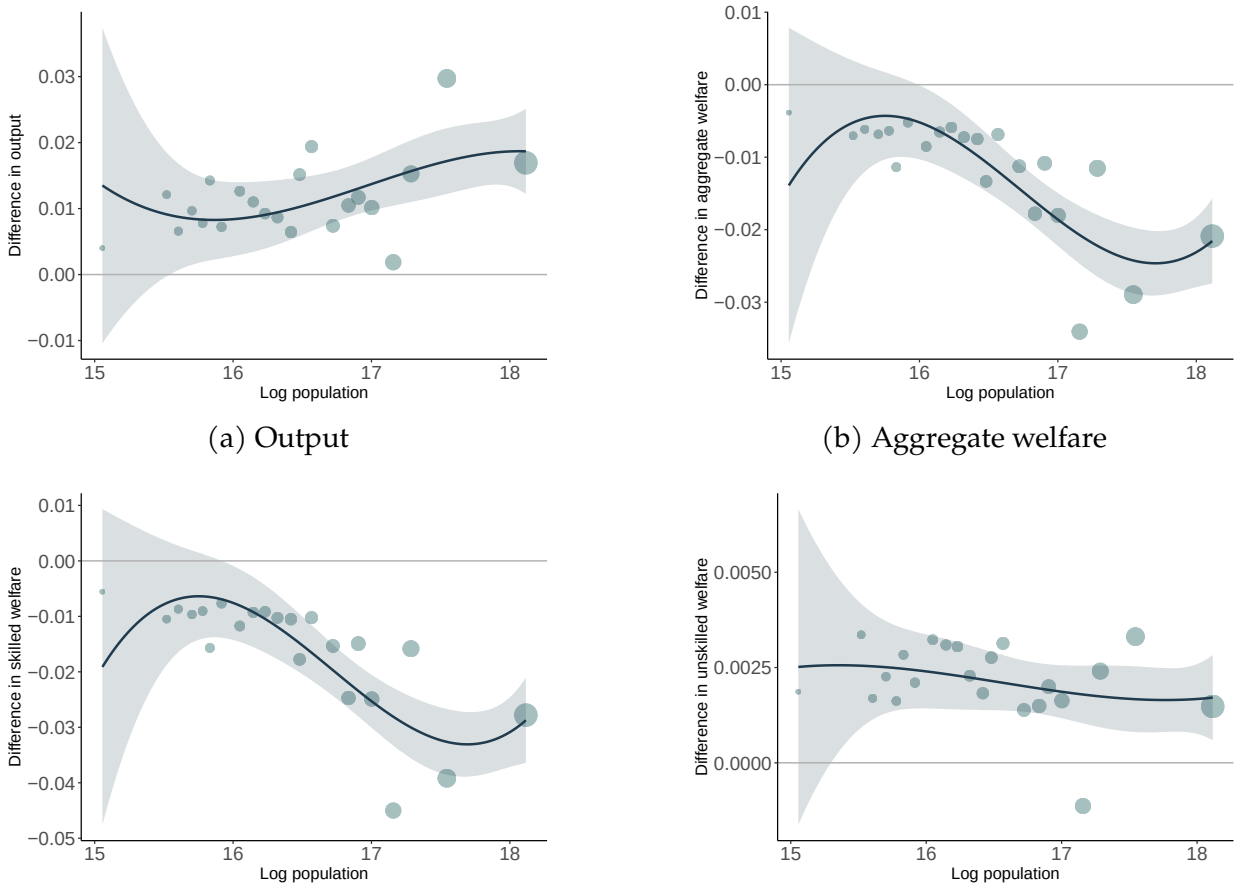
(c) Skill Ratio



(d) Skill Premium

Figure A3: The Spatial Sorting Effects of the Intensive Margin of AI Automation

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A polynomial regression is then fitted to the binned data.



(a) Output

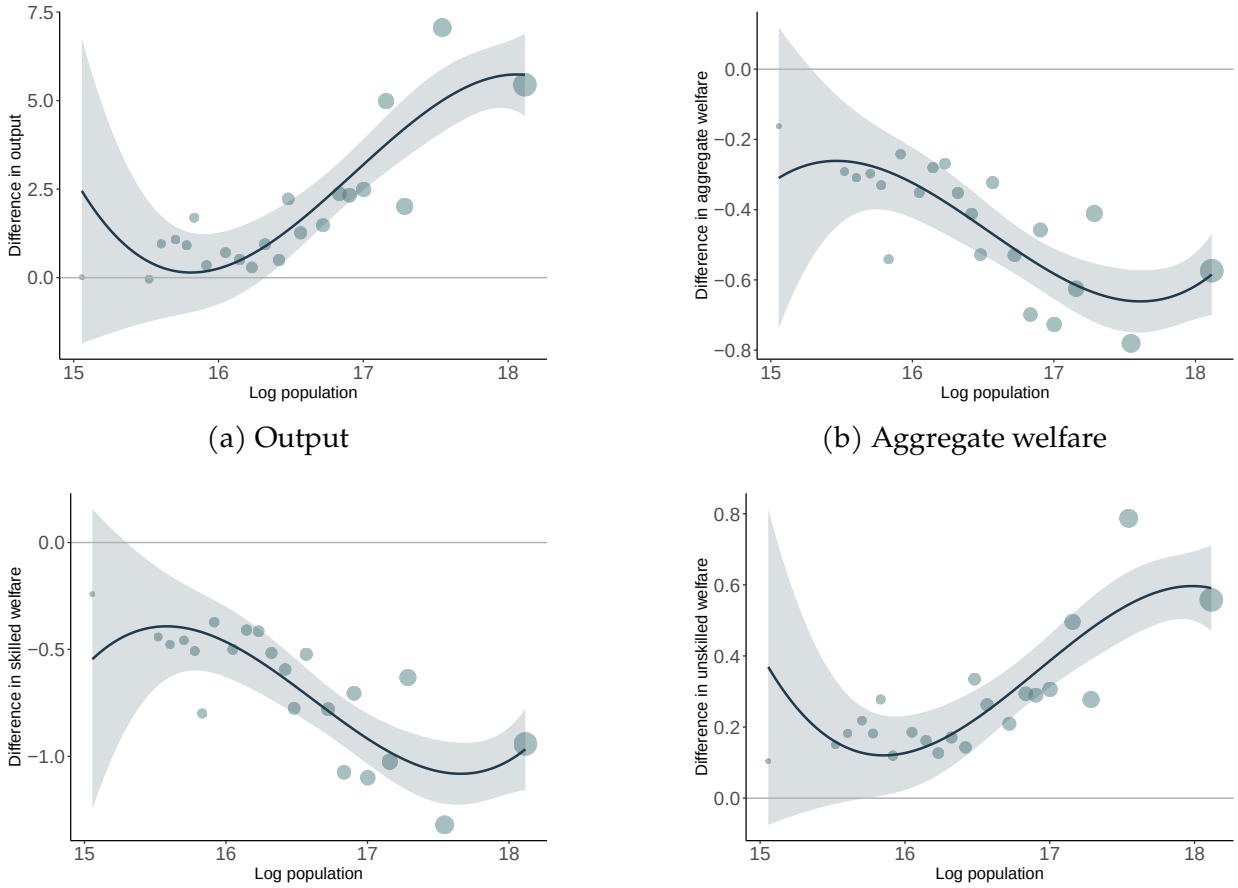
(b) Aggregate welfare

(c) Skilled welfare

(d) Unskilled welfare

Figure A4: The Unequal Gains from the Extensive Margin of AI Automation

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A polynomial regression is then fitted to the binned data.



(a) Output

(b) Aggregate welfare

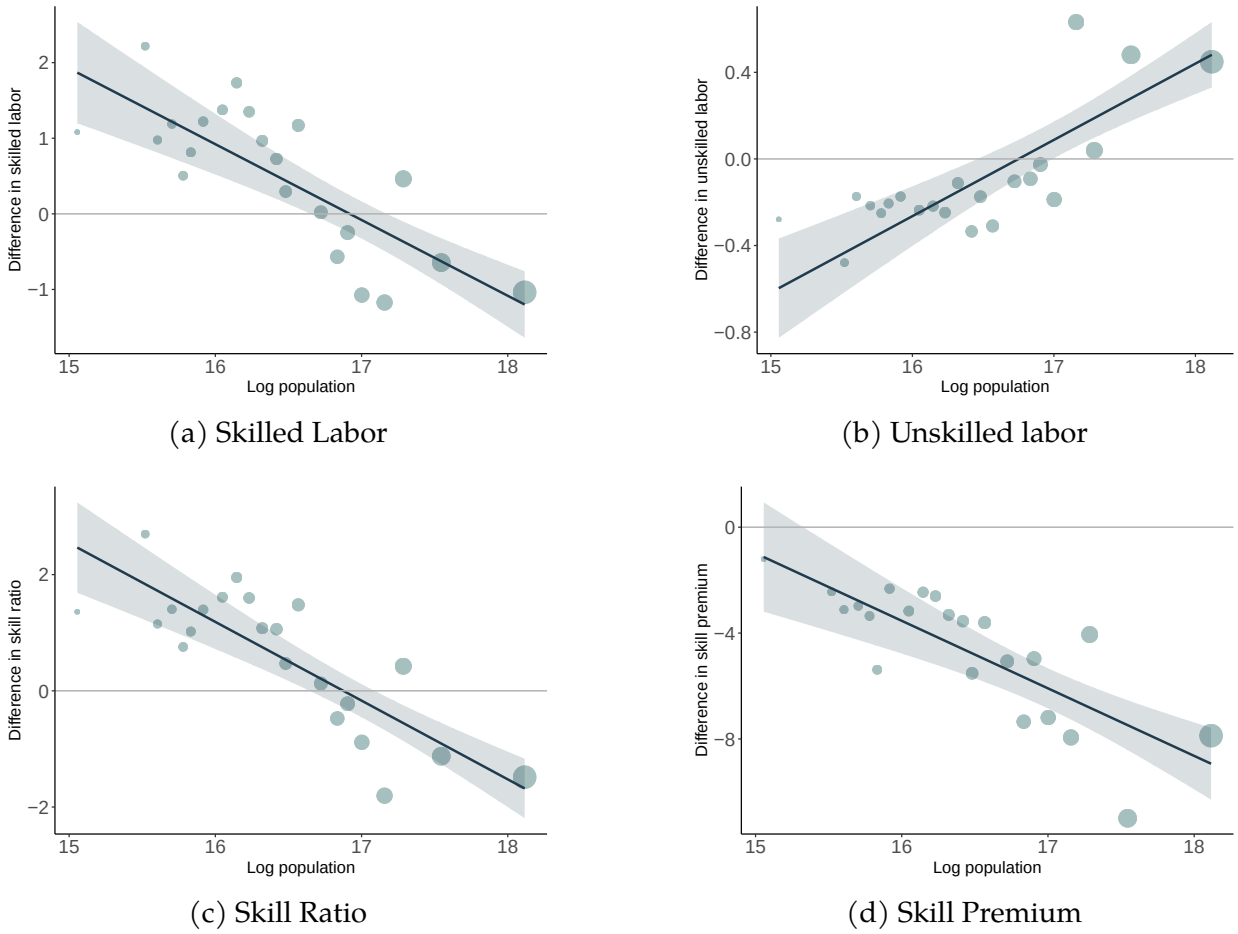
(c) Skilled welfare

(d) Unskilled welfare

Figure A5: The Unequal Gains from the Intensive Margin of AI Automation

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A polynomial regression is then fitted to the binned data.

A.1 Linear Fit of Counterfactual Simulations



(a) Skilled Labor

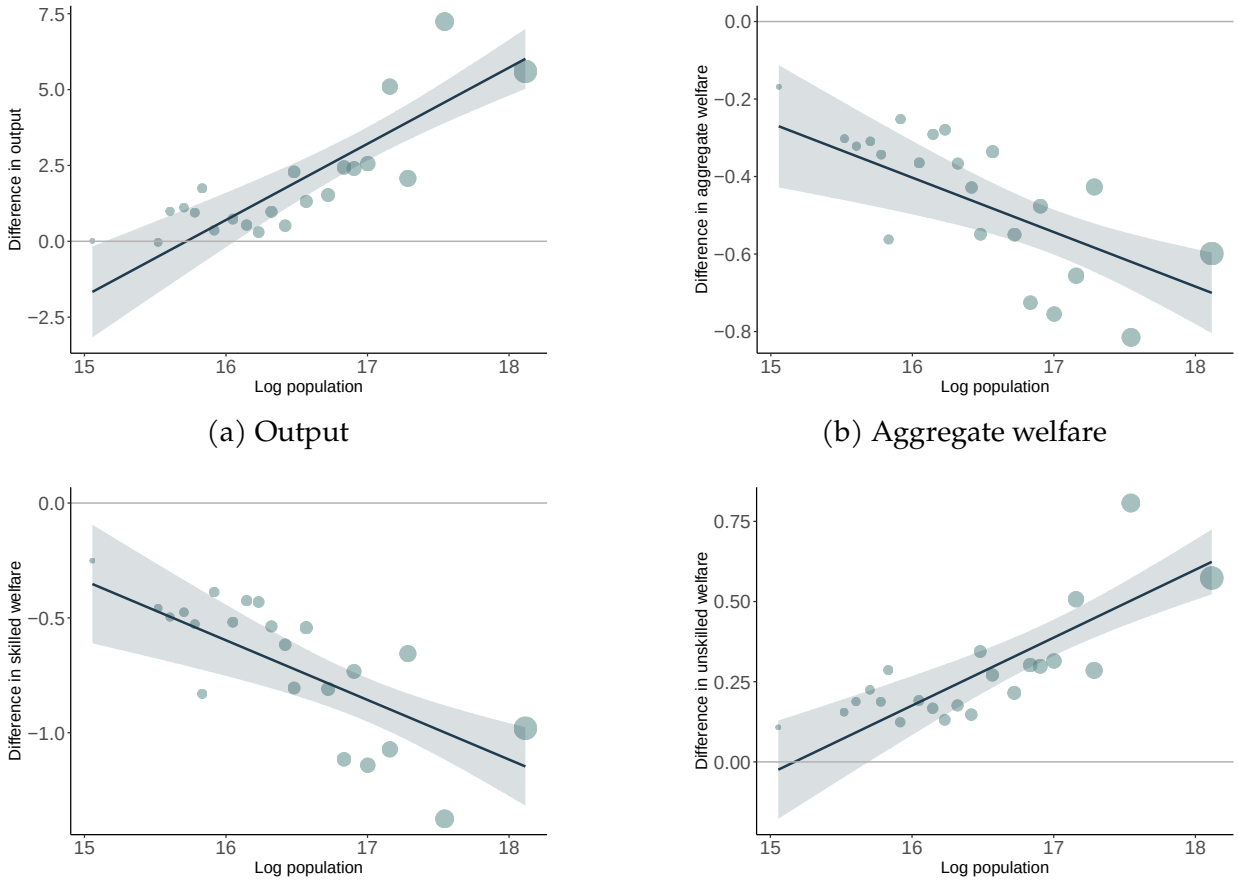
(b) Unskilled labor

(c) Skill Ratio

(d) Skill Premium

Figure A6: The Spatial Sorting Effects of AI Automation (Linear Fit)

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A linear regression is then fitted to the binned data.



(a) Output

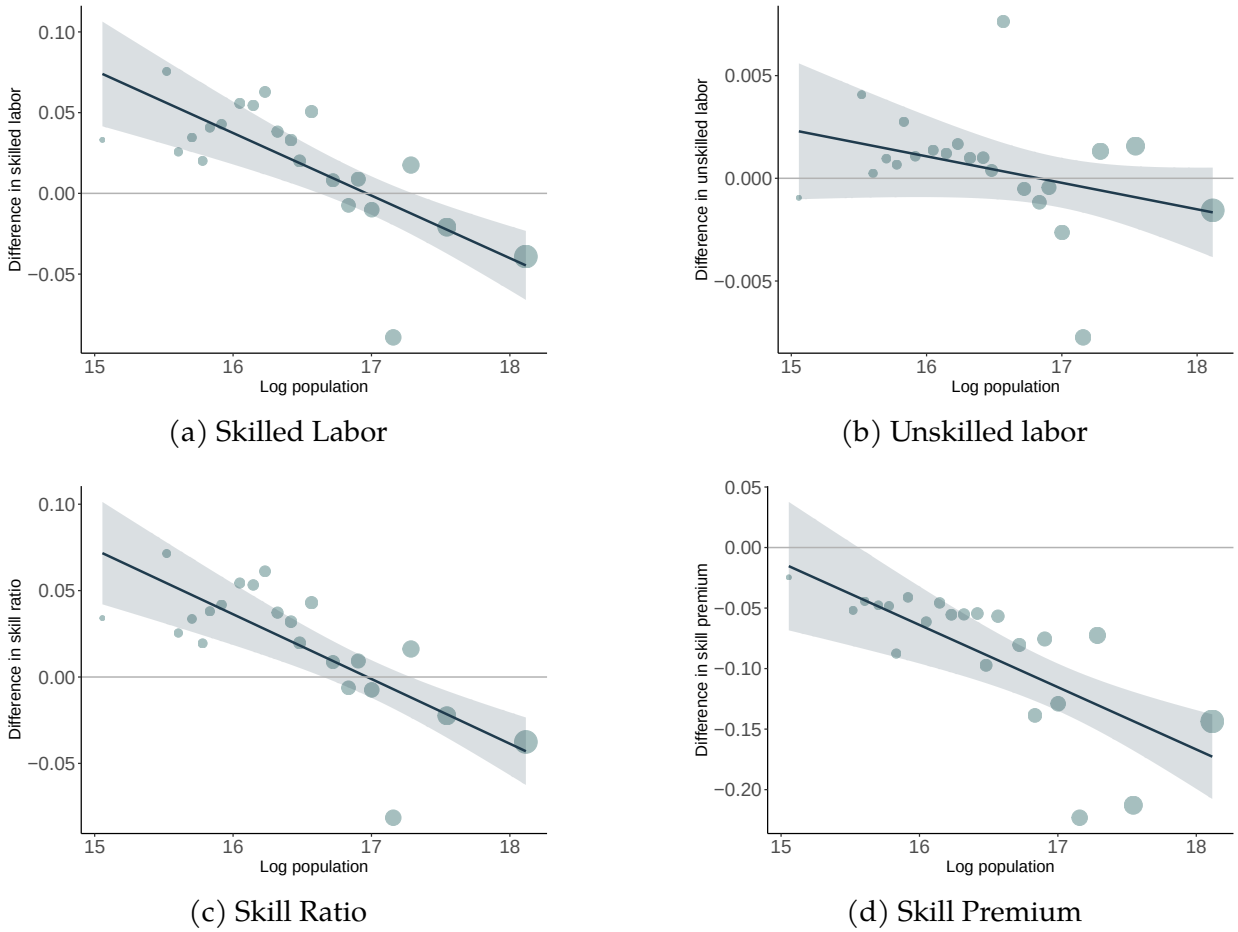
(b) Aggregate welfare

(c) Skilled welfare

(d) Unskilled welfare

Figure A7: The Unequal Gains from AI Automation (Linear Fit)

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A linear regression is then fitted to the binned data.



(a) Skilled Labor

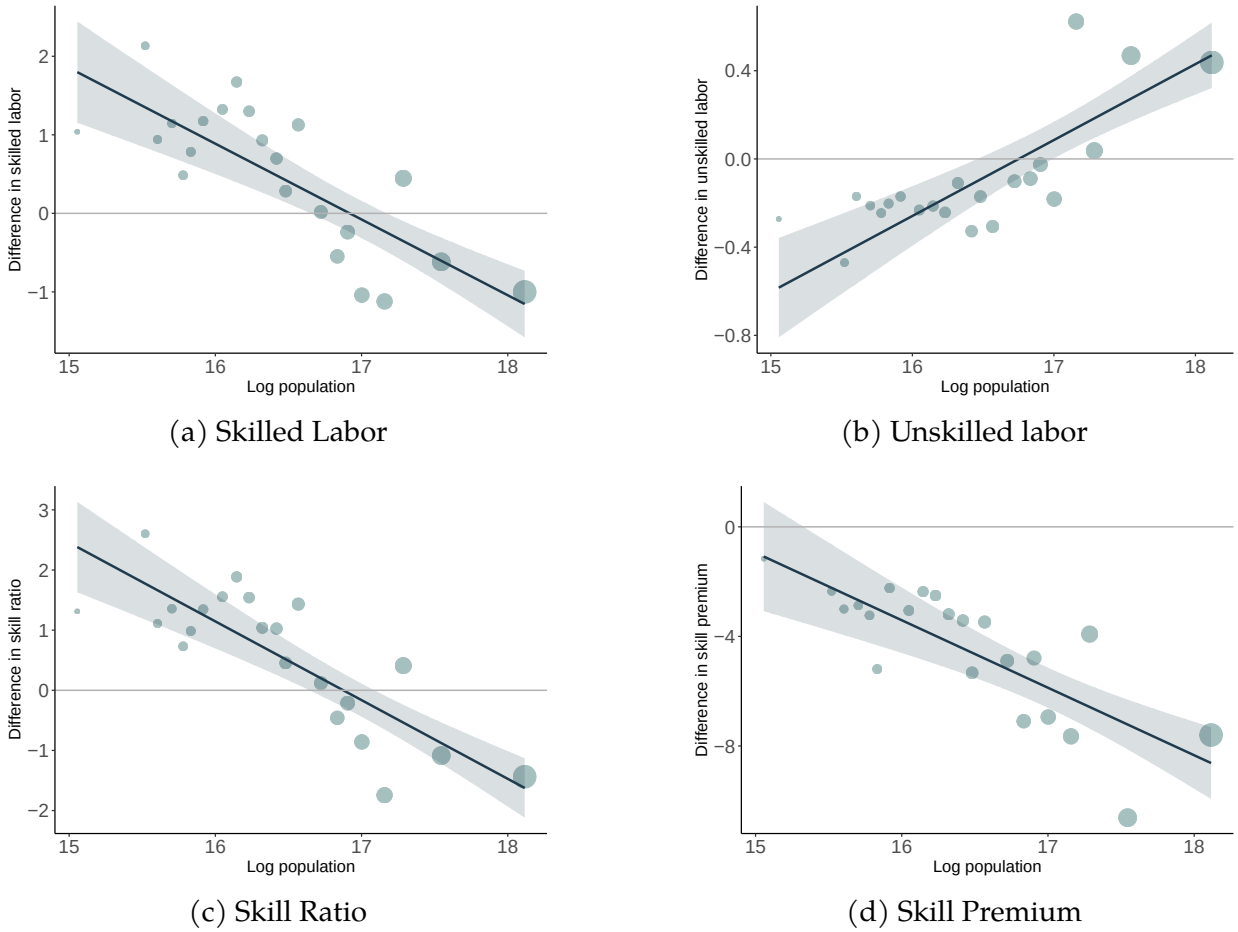
(b) Unskilled labor

(c) Skill Ratio

(d) Skill Premium

Figure A8: The Spatial Sorting Effects of the Extensive Margin of AI Automation (Linear Fit)

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A linear regression is then fitted to the binned data.



(a) Skilled Labor

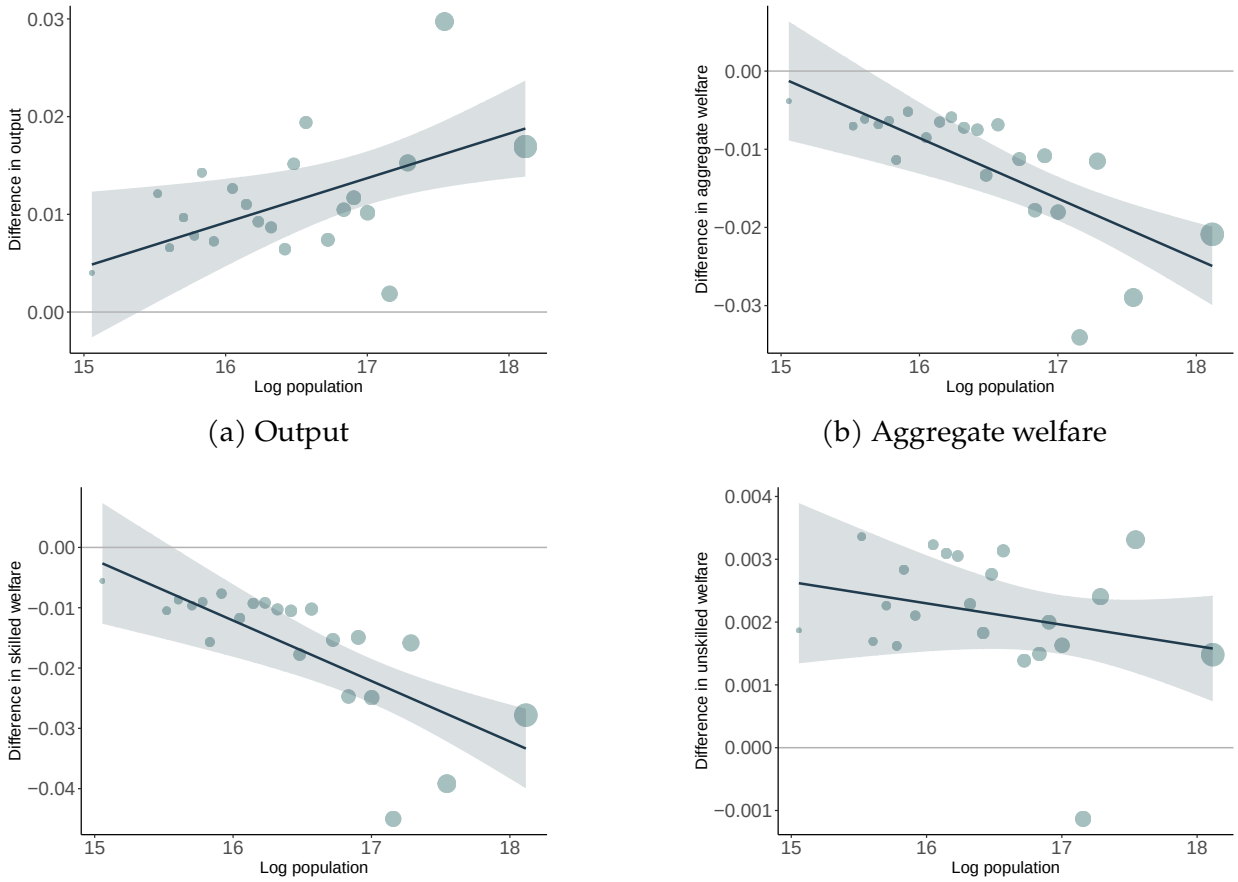
(b) Unskilled labor

(c) Skill Ratio

(d) Skill Premium

Figure A9: The Spatial Sorting Effects of the Intensive Margin of AI Automation (Linear Fit)

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A linear regression is then fitted to the binned data.



(a) Output

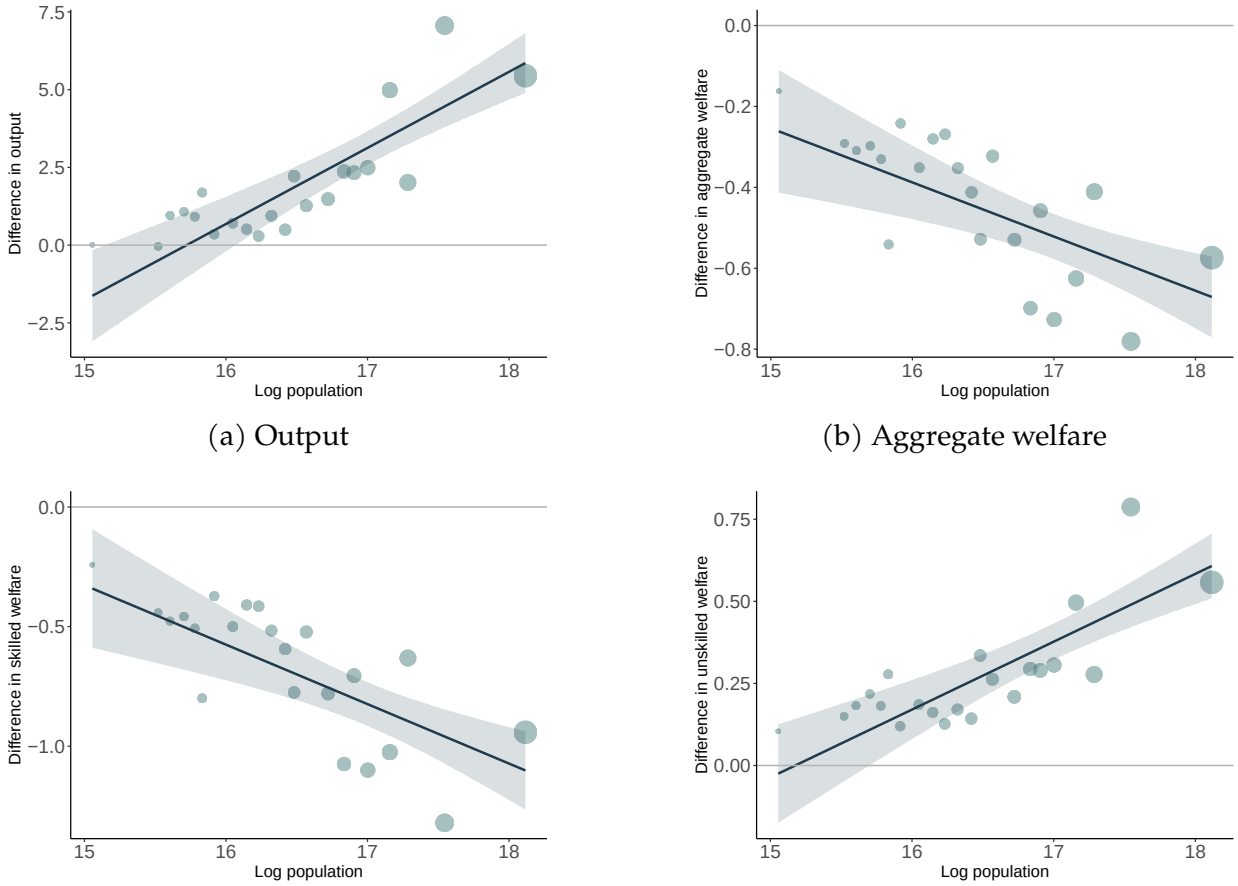
(b) Aggregate welfare

(c) Skilled welfare

(d) Unskilled welfare

Figure A10: The Unequal Gains from the Extensive Margin of AI Automation (Linear Fit)

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A linear regression is then fitted to the binned data.



(a) Output

(b) Aggregate welfare

(c) Skilled welfare

(d) Unskilled welfare

Figure A11: The Unequal Gains from the Intensive Margin of AI Automation (Linear Fit)

Notes: The 110 cities are sorted based on their year 2000 population and classified into 22 equal-sized bins, with each bin containing exactly five cities. A linear regression is then fitted to the binned data.

B Data and Measurement

B.1 Data

B.1.1 Census Data

We use data from the 2000, 2010, and 2015 population censuses in China. The 2000 Census data is sourced from IPUMS.[15] The 2010 Census and the 2015 "mini" Census are obtained from the National Bureau of Statistics (NBS). We identify migrants as individuals residing outside their hukou registration city and construct bilateral migration shares between city pairs using census data. We measure skill shares in the 2010 and 2015 censuses, defining skilled labor as individuals with a junior college degree or above and unskilled labor as those with a high school education or below.

B.1.2 Firm Data

Data on the software assets of publicly listed firms are sourced from CSMAR.[16] The Annual Surveys of Manufacturing Firms (2011–2015) are sourced from EPS Data.

B.1.3 Patent Data

Patent data are sourced from the China National Intellectual Property Administration (CNIPA), Google Patents, and Incopat. We match patents to publicly listed firms by exact string matching on applicant names.

B.1.4 Population and Wage Data

Data on city-level population and wages are obtained from the China Urban Construction Statistical Yearbooks (CUCSY, 2010–2020), maintained by the Ministry of Housing and Urban-Rural Development.

B.1.5 China Migrants Dynamic Survey (CMDS) Data

Data on the city-level skill premium are sourced from the China Migrants Dynamic Survey (CMDS) for the years 2012, 2013, 2015, 2017, and 2018. We employ Mincer regressions in equation (A1) to estimate skill premiums for each city.

B.1.6 Public Research Grants Data

We obtain public research grant data from the National Natural Science Foundation of China (NSFC), including details on grant timing, funding amounts, university affiliations, abstracts, and publications. We construct the pre-2010 stock of research grants for each university and compute a technological similarity index between each publicly listed firm

[15] This sample represents a 1% extract of the national population, comprising 11,804,344 individuals. We use the demographic information from this sample to calculate the occupational shares within each city.

[16] CSMAR is a research database that provides comprehensive data on publicly listed firms in China.

and university. We then measure firm-level AI software by isolating changes in software assets attributable to public AI R&D shocks.

B.1.7 Occupation Data

Because the Chinese Standard Classification of Occupations (CSCO) does not provide the task-level detail required for our analysis, we map Chinese occupations to the U.S. Standard Occupational Classification (SOC) system. This crosswalk allows us to leverage the task descriptors available in O*NET.

We construct the crosswalk in three steps. First, we use the mapping provided by the China Family Panel Studies (CFPS) to link CSCO codes to the International Standard Classification of Occupations (ISCO-88). Second, we convert ISCO-88 codes to ISCO-08 using the official ILO correspondence table. Third, we apply existing crosswalks to map ISCO-08 codes to SOC-2010 codes, which allows us to assign ONET task content measures to each Chinese occupation. Overall, this procedure maps 333 three-digit CSCO occupations to one or more ONET counterparts.

B.2 Descriptive Evidence on Grants and Firms

B.2.1 Public AI R&D Grants in China

In this section, we identify approximately 13,000 AI-related grants using a keyword-based search, accounting for about 10% of all public R&D grants, and document the spatial distribution and key characteristics of these grants across Chinese cities.

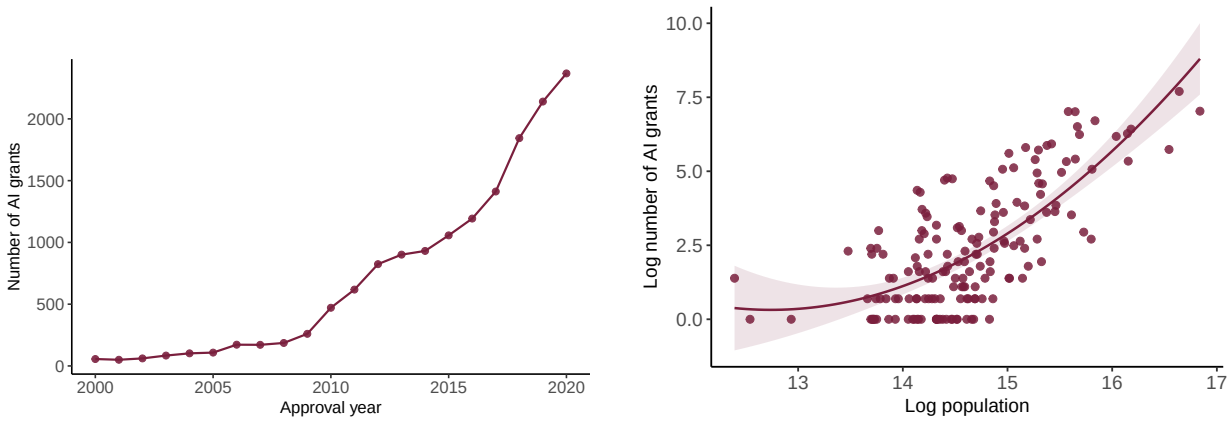
Types of Grants We focus on the General Program and the Junior Researcher Fund, which together account for more than 88 percent of AI-related grants. Table A6 shows that these programs represent 50.16 percent and 37.94 percent of total grants, respectively. Because both programs rely on a standardized peer-review process and support broad, investigator-initiated research, their allocation across topics is less likely to be directed by national industrial policy or place-based policy priorities than more specialized programs, such as the Regional Science Fund and Major Research Plan. This institutional feature motivates our use of these grants as a source of variation in the empirical analysis.

Table A6: Distribution of Public AI Grants by Type

Grant type	*N*	Share (%)	Mean funding	Funding share (%)
General Program	7,528	50.16	57.40	58.14
Junior Researcher Fund	5,695	37.94	23.19	17.77
Regional Science Fund	957	6.38	39.22	5.05
Joint Fund	307	2.05	149.75	6.19
Key Program	261	1.74	272.80	9.58
Major Research Plan	163	1.09	124.87	2.74
Special Fund Program	98	0.65	40.52	0.53

Notes: Mean funding is measured in 10,000 RMB.

Trends in AI Grants and Distribution Across Cities Figure A12 summarizes the time pattern of AI grant allocations and their spatial distribution across Chinese cities. Panel (a) shows that the number of approved public AI R&D grants increased sharply after 2010 and continued to grow rapidly over the subsequent decade. Panel (b) aggregates AI grants to the city level and shows that funding is disproportionately concentrated in larger cities.



(a) Temporal trends, 2010–2020

(b) Spatial distribution across cities

Figure A12: Temporal trends and spatial distribution of AI grants

Subjects and Universities Table A7 reports the distribution of AI grants across academic disciplines and shows that Computer Science, Electronics and Information Systems, and Artificial Intelligence account for the largest shares.

Table A7: Top Subjects by Number of AI Grants

Subject	Number of grants
Computer Science	3,221
Electronics and Information Systems	2,450
Artificial Intelligence	2,398
Automation	2,028
Mechanical Design and Manufacturing	1,218
Architecture and Civil Engineering	515
Semiconductor Science and Information Devices	376
Inorganic Nonmetallic Materials	324
Electrical Science and Engineering	324
Optics and Optoelectronics	323

Furthermore, Table A8 shows that top universities located in large cities receive a disproportionate share of public AI R&D grants.

Table A8: Top Universities by Number of AI Grants

University	City	Grants
Tsinghua University	Beijing	407
Harbin Institute of Technology	Harbin	348
Shanghai Jiao Tong University	Shanghai	325
Zhejiang University	Hangzhou	319
Huazhong University of Science and Technology	Wuhan	317
Xidian University	Xi'an	288
Tianjin University	Tianjin	254
Peking University	Beijing	252
Beihang University	Beijing	247
University of Electronic Science and Technology of China	Chengdu	244

B.2.2 Publicly Listed Firms

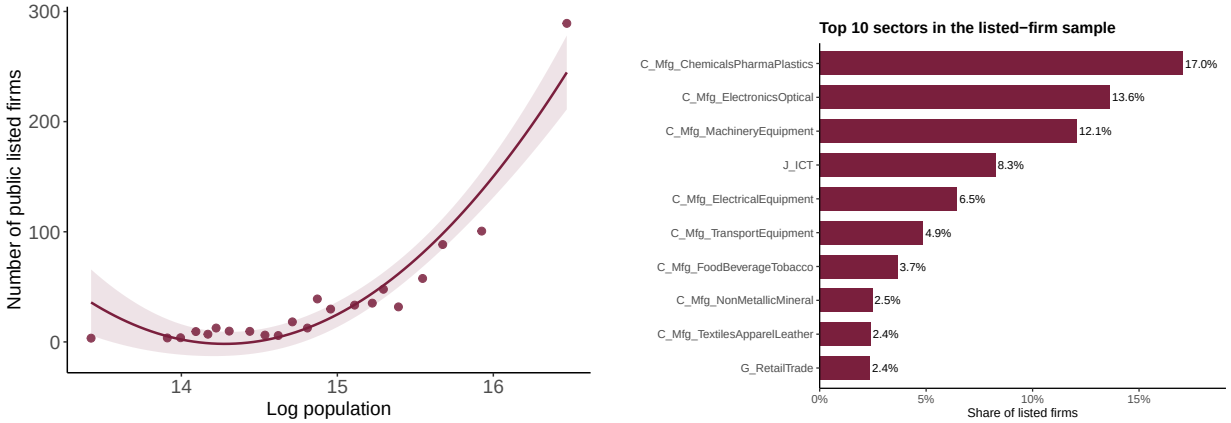
In this subsection, we describe software assets of publicly listed firms, document their sectoral and spatial distributions, and assess the representativeness of the listed-firm sample.

Software Assets We construct our firm-level panel using annual software-asset data for publicly listed firms from 2010 to 2018.[17] During this period, corporate software investment primarily took the form of outright purchases and custom development, both of which are capitalized as software assets on corporate balance sheets.

A potential measurement concern is that the rising prevalence of software subscriptions after 2016 might mechanically inflate reported software assets, reflecting higher recurring fees rather than technology adoption. However, under Chinese accounting standards, only purchased software or eligible customization costs are capitalized as assets, whereas subscriptions, maintenance, and other services are expensed. Therefore, increases in our software-asset measure reflect capitalized technology investment.

Industries and Cities As depicted in Figure A13, these firms disproportionately cluster within information and communication technology (ICT) and advanced manufacturing industries.

[17] Because listed firms are mandated to disclose these items annually, the CSMAR database provides longitudinal coverage of software capital.



(a) Geographical distribution of listed firms

(b) Sectoral distribution of listed firms

Figure A13: Spatial and sectoral concentration of listed firms

Representativeness To form the baseline city set for calibration of the spatial equilibrium model, we impose two sample restrictions. First, we keep cities that host at least one listed firm with positive software assets, ensuring that AI software is observed. Second, we require the city set to overlap with the coverage of our labor-supply estimates. The resulting estimation sample contains 4,274 firms in 110 cities. Compared with the full universe of 4,766 listed firms in 253 cities, the restricted sample retains 89.68 percent of listed firms.

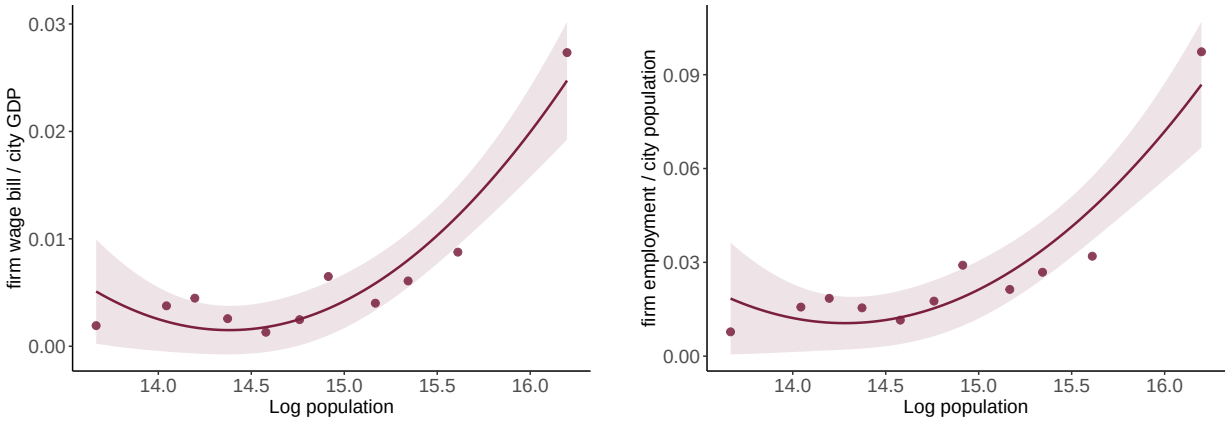
A potential concern is that listed firms do not operate in every city and are not well represented in smaller locations, which could introduce measurement error in city-level software. To adrress this, Table A9 summarizes listed-firm coverage across city population deciles. Even in smaller-city deciles, the median city hosts 5–10 listed firms. In addition, software investment levels are much lower in smaller cities, implying that our predicted AI-software investment measure is driven primarily by firms in medium-size and large cities, where listed-firm coverage is more informative.

Table A9: Distribution of Listed Firms by City Population Decile

Decile	Avg. log population	Total firms	Avg. firms	Median firms
1	13.69	37	3.4	2
2	14.07	72	6.5	5
3	14.23	126	11.5	8
4	14.45	87	7.9	7
5	14.66	125	11.4	8
6	14.84	281	25.5	16
7	15.07	367	33.4	32
8	15.30	450	40.9	29
9	15.57	732	66.5	36
10	16.16	1,997	181.5	112

One might also worry that publicly listed firms provide an incomplete view of local economic activity and labor markets. To assess representativeness, we compare listed-firm aggregates to city-level GDP and population. In our sample, the aggregate listed-firm wage bill accounts for an average of 0.54 percent of city GDP (median 0.27 percent). These firms employ an average of 2.50 percent of the local population (median 1.51 percent).

We next examine the representativeness of listed firms by city size. Figure A14 plots (i) the ratio of the total listed-firm wage bill to city GDP and (ii) the ratio of total listed-firm employment to city population against city size. Listed firms account for a disproportionately large share of economic activity and employment in larger cities.



(a) Listed-firm wage bill relative to local GDP

(b) Listed-firm employment relative to local population

Figure A14: Listed-firm contributions to local GDP and employment

Notes: The 110 cities are sorted by their population in 2000 and grouped into 11 equal-sized bins, with each bin containing ten cities. A polynomial regression is fitted to the binned data.

B.3 Other Measurement

Wage Estimation (Mincer Regression) To obtain skill premium at city level, we run the following Mincer regression:

$$\ln w_{ij}^{e} = \beta_0 + \delta_j^e + \beta_1 \,\mathrm{Age}_i + \beta_2 \,\mathrm{Age}_i^2 + \mathrm{X}_i\gamma' + \theta_d + \varepsilon_i^e. \tag{A1}$$

where $\ln w_{ij}^{e}$ denotes the log wage of individual in city j, while Age_i represents the age of individual i. δ_j^e denotes the city j skill type e fixed effects (with unskilled labor in Shanghai as the omitted group), and θ_d represents industry fixed effects. The vector of controls X_i includes hukou dummies (urban, rural), work type dummies (employer, employee, self-employed), and gender indicators. The estimated average wage fixed effect for labor of type e in city j is given by $\widehat{\beta_0} + \widehat{\delta_j^e}$. We then could recover the skill premium at city level.

C Model Derivation

C.1 Task-Specific Productivity

Lemma 1 (Automation Threshold). *Suppose $\psi(i) = i^{-1/\theta}$ with $\theta > 0$ and $A^k w_j^s/(A^s v) < 1$. The unique interior threshold satisfying cost-parity condition (13) is*

$$\widetilde{\beta}_j^k = \left(\frac{A^k w_j^s}{A^s v}\right)^{\theta} \in (0,1). \tag{A2}$$

The threshold is strictly increasing in the skilled wage w_j^s and AI productivity A^k, and strictly decreasing in the AI price v and skill-biased productivity A^s.

Proof. Substituting $\psi(i) = i^{-1/\theta}$ into the cost-parity condition (13) gives

$$\frac{w_j^s}{A^s} = \frac{v}{A^k}\left(\widetilde{\beta}_j^k\right)^{1/\theta}, \tag{A3}$$

which rearranges directly to (A2). Since $A^k w_j^s/(A^s v) < 1$, raising a number in $(0,1)$ to the power $\theta > 0$ yields a value in $(0,1)$, ensuring that $\widetilde{\beta}_j^k$ is interior solution. The monotonicity claims follow immediately from (A2). ∎

Remark 1. The condition $A^k w_j^s/(A^s v) < 1$ has a natural economic interpretation. At the most complex task $i = 1$, where $\psi(1) = 1$, it implies that the unit cost of skilled labor is strictly below the unit cost of AI software. Hence, skilled labor retains a strict comparative advantage at the upper end of the task distribution. Therefore, full automation is never optimal and the threshold $\widetilde{\beta}_j^k$ satisfies $\widetilde{\beta}_j^k < 1$.

Lemma 2 (Aggregate Task Share of AI software). *Let $\psi(i) = i^{-1/\theta}$ for $\theta > 0$ and suppose Lemma 1 holds. For $\rho \in (0,1)$, the aggregate task share of AI software,*

$$\widetilde{\Gamma}_j^k \equiv \int_0^{\widetilde{\beta}_j^k} \psi(i)^{\frac{\rho}{1-\rho}}\, di, \tag{A4}$$

is finite if and only if

$$\theta(1-\rho) > \rho. \tag{A5}$$

Under (A5), $\widetilde{\Gamma}_j^k$ *admits the closed form*

$$\widetilde{\Gamma}_j^k = \frac{\left(\widetilde{\beta}_j^k\right)^{1-\frac{\rho}{\theta(1-\rho)}}}{1-\frac{\rho}{\theta(1-\rho)}}, \tag{A6}$$

and is strictly increasing and strictly concave in $\widetilde{\beta}_j^k$.

Proof. Substituting $\psi(i) = i^{-1/\theta}$ into (A4):

$$\widetilde{\Gamma}_j^k = \int_0^{\widetilde{\beta}_j^k} i^{-\frac{\rho}{\theta(1-\rho)}}\, di. \tag{A7}$$

Since $\rho > 0$, the exponent $-\rho/[\theta(1-\rho)]$ is strictly negative, so the integrand diverges as $i \to 0^+$. The standard criterion for power integrals requires the exponent to exceed -1, which gives:

$$-\frac{\rho}{\theta(1-\rho)} > -1 \iff \theta(1-\rho) > \rho, \tag{A8}$$

establishing that (A5) is necessary and sufficient for finiteness. Under (A5), the exponent satisfies $\kappa \equiv 1 - \rho/[\theta(1-\rho)] \in (0,1)$, and applying the power rule to (A7) delivers (A6), using $i^\kappa \to 0$ as $i \to 0^+$. Differentiating (A6):

$$\frac{\partial \widetilde{\Gamma}_j^k}{\partial \widetilde{\beta}_j^k} = \left(\widetilde{\beta}_j^k\right)^{\kappa-1} > 0, \qquad \frac{\partial^2 \widetilde{\Gamma}_j^k}{\partial (\widetilde{\beta}_j^k)^2} = (\kappa-1)\left(\widetilde{\beta}_j^k\right)^{\kappa-2} < 0, \tag{A9}$$

since $\kappa \in (0,1)$, confirming strict monotonicity and strict concavity. ∎

Remark 2. Condition (A5) jointly disciplines the task-level AI productivity decay parameter θ and the elasticity of substitution between AI software and skilled labor, $\frac{1}{1-\rho}$. Intuitively, as $i \to 0^+$, the most productive tasks have $\psi(i) \to \infty$, causing the integrand $i^{-\rho/[\theta(1-\rho)]}$ in (A7) to diverge. Condition (A5) bounds this divergence so that aggregate cost savings over the set of automated tasks remain integrable and $\widetilde{\Gamma}_j^k$ is well-defined and finite. In addition, the concavity of $\widetilde{\Gamma}_j^k$ in $\widetilde{\beta}_j^k$ implies diminishing marginal increases in the AI task share as automation expands: cities with higher initial thresholds experience smaller incremental increases in $\widetilde{\Gamma}_j^k$ when $\widetilde{\beta}_j^k$ rises.

C.2 Housing Supply

Housing is produced by combining land and capital under the assumption of a perfectly competitive supply market. We specify the production function using a Cobb-Douglas technology:

$$HS_j = L_j^{\alpha} K_j^{1-\alpha} \tag{A10}$$

where α represents the land cost share, while L_j and K_j denote the quantities of land and capital used in city j, respectively.

Each city is characterized by a fixed, exogenous land supply $\bar{L}_j$, while capital is perfectly mobile across the economy. Under the assumption of perfect competition, the housing price r_j of city j equals the unit cost of production:

$$r_j = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)} p_j^{\alpha} v^{1-\alpha} \tag{A11}$$

where p_j is the local land price and v denotes the price of capital, normalized to the numeraire.

Land market clearing The equilibrium in the land market is characterized by the equality of land demand and the fixed endowment $\bar{L}_j$. Since the expenditure on land is a constant fraction α of total housing revenue, we have $p_j L_j = \alpha r_j HS_j$. Rearranging this yields the land market clearing condition:

$$L_j = HS_j \cdot \frac{\alpha r_j}{p_j} = \bar{L}_j \tag{A12}$$

Housing Supply Elasticity From the unit cost function (A11), holding the price of capital v fixed, the relationship between changes in the house price and the land price is given by:

$$d \ln r_j = \alpha \cdot d \ln p_j \tag{A13}$$

Taking logs in (A12) and differentiating yields:

$$d \ln HS_j + d \ln r_j - d \ln p_j = 0. \tag{A14}$$

By substituting $d \ln p_j = \frac{1}{\alpha} d \ln r_j$ into (A14), we obtain:

$$\ln HS_j = \left(\frac{1}{\alpha} - 1 \right) d \ln r_j \tag{A15}$$

By combining the unit cost of housing production function with the land market clearing condition, we show that the housing supply elasticity μ is determined by the land cost

share α. This relationship is given by:

$$\mu = \frac{d \ln HS_j}{d \ln r_j} = \frac{1-\alpha}{\alpha} \tag{A16}$$

C.3 Labor Demand

By solving the representative firm's cost minimization problem in each city and normalizing the price of AI software as the numeraire, we derive the first order conditions for input demand. We then totally differentiate these conditions, which yields a system of three structural equations characterizing labor demand. In particular, the AI software share, the skill premium, and skilled wages can each be expressed as functions of the local skill ratio and exogenous technological change.

$$
\begin{aligned}
d\ln\left(\frac{K_j}{Y_j}\right) = {} & \frac{(1-\omega_j^{ks})(1-\omega_j^{k})(\rho-\sigma)}{Z_j} d\ln\left(\frac{S_j}{U_j}\right) \\
& + \frac{1-\omega_j^{ks}\omega_j^{k}}{Z_j} d\ln v - \frac{\rho(1-\omega_j^{k}) - \sigma(1-\omega_j^{ks}\omega_j^{k})}{Z_j}\left(d\ln A^{k} + \frac{1-\rho}{\rho} d\ln \widetilde{\Gamma}_j^{k}\right) \\
& + (1-\omega_j^{k})\left[\frac{(1-\omega_j^{s}\omega_j^{k})(\rho-\omega_j^{ks})}{Z_j} - \omega_j^{ks}\right]\left(d\ln A^{s} + \frac{1-\rho}{\rho} d\ln(1-\widetilde{\beta}_j^{k})\right) - d\ln A_j
\end{aligned}
$$

$$
\begin{aligned}
d\ln\left(\frac{w_j^{s}}{w_j^{u}}\right) = {} & \left[\sigma - 1 + \frac{\omega_j^{k}(\sigma-\rho)(1-\sigma)(1-\omega_j^{ks})}{Z_j}\right] d\ln\left(\frac{S_j}{U_j}\right) \\
& + \frac{\omega_j^{k}(\sigma-\rho)}{Z_j} d\ln v - \frac{\omega_j^{k}(\sigma-\rho)}{Z_j}\left(d\ln A^{k} + \frac{1-\rho}{\rho} d\ln \widetilde{\Gamma}_j^{k}\right) \\
& + \omega_j^{k}(\sigma-\rho)\left[\frac{(1-\omega_j^{k})(\omega_j^{ks}-\rho)}{Z_j} + 1\right]\left(d\ln A^{s} + \frac{1-\rho}{\rho} d\ln(1-\widetilde{\beta}_j^{k})\right)
\end{aligned}
$$

$$
\begin{aligned}
d\ln\left(w_j^{s}\right) = {} & \frac{(1-\rho)(1-\sigma)(1-\omega_j^{ks})}{Z_j} d\ln\left(\frac{S_j}{U_j}\right) \\
& + \frac{\omega_j^{k}\left[\omega_j^{ks}(1-\sigma)+\sigma-\rho\right]}{Z_j} d\ln v - \frac{\omega_j^{k}\left[\omega_j^{ks}(1-\sigma)+\sigma-\rho\right]}{Z_j}\left(d\ln A^{k} + \frac{1-\rho}{\rho} d\ln \widetilde{\Gamma}_j^{k}\right) \\
& + \frac{\rho Z_j - (1-\rho)\left[\omega_j^{ks}(1-\sigma)+\sigma-\rho\right](1-\omega_j^{k})}{Z_j}\left(d\ln A^{s} + \frac{1-\rho}{\rho} d\ln(1-\widetilde{\beta}_j^{k})\right)
\end{aligned}
$$

where $Z_j = (\sigma-\rho)\omega_j^{k}(1-\omega_j^{ks}) - (1-\rho)(1-\omega_j^{k}\omega_j^{ks})$, ω_j^{k} denotes the share of AI software within the AI–skilled nest in city j, and ω_j^{ks} represents the share of expenditure on the AI–skilled composite relative to total cost in city j.

The share of AI software within the AI–skilled composite is given by:

$$\omega_j^k = \frac{\left(\left(\widetilde{\Gamma}_j^k\right)^{\frac{1-\rho}{\rho}} A^k K_j\right)^{\rho}}{\left(\left(\widetilde{\Gamma}_j^k\right)^{\frac{1-\rho}{\rho}} A^k K_j\right)^{\rho} + \left((1-\widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j\right)^{\rho}} \tag{A17}$$

The share of skilled labor within the AI-skilled labor composite is defined as:

$$\omega_j^s = \frac{\left((1-\widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j\right)^{\rho}}{\left(\left(\widetilde{\Gamma}_j^k\right)^{\frac{1-\rho}{\rho}} A^k K_j\right)^{\rho} + \left((1-\widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j\right)^{\rho}} \tag{A18}$$

The share of the AI-skilled labor composite in production is defined as:

$$\omega_j^{ks} = \frac{\eta_j^s \left[\left(\left(\widetilde{\Gamma}_j^k\right)^{\frac{1-\rho}{\rho}} A^k K_j\right)^{\rho} + \left((1-\widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j\right)^{\rho}\right]^{\frac{\sigma}{\rho}}}{\eta_j^u (U)^{\sigma} + \eta_j^s \left[\left(\left(\widetilde{\Gamma}_j^k\right)^{\frac{1-\rho}{\rho}} A^k K_j\right)^{\rho} + \left((1-\widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j\right)^{\rho}\right]^{\frac{\sigma}{\rho}}} \tag{A19}$$

The unskilled labor share in total production is given by:

$$\omega_j^u = \frac{\eta_j^u (U)^{\sigma}}{\eta_j^u (U)^{\sigma} + \eta_j^s \left[\left(\left(\widetilde{\Gamma}_j^k\right)^{\frac{1-\rho}{\rho}} A^k K_j\right)^{\rho} + \left((1-\widetilde{\beta}_j^k)^{\frac{1-\rho}{\rho}} A^s S_j\right)^{\rho}\right]^{\frac{\sigma}{\rho}}} \tag{A20}$$

C.4 Migration cost

To back out the model-implied migration cost, we first compute the relative migration cost between cities, conditional on selecting a reference destination city k. The ratio of migration probabilities is given by:

$$\frac{\pi_{ij}^e}{\pi_{ik}^e} = \left[\frac{\exp(-\tau_{ij}^e)V_{ij}^e}{\exp(-\tau_{ik}^e)V_{ik}^e}\right]^{\eta_e} \tag{A21}$$

Rearranging the expression, we obtain the relative migration cost:

$$\frac{\exp(-\tau_{ij}^e)}{\exp(-\tau_{ik}^e)} = \left(\frac{\pi_{ij}^e}{\pi_{ik}^e}\right)^{\frac{1}{\eta_e}} \left(\frac{V_{ik}^e}{V_{ij}^e}\right) \tag{A22}$$

We then substitute this expression for relative migration cost into the upper nest equation:

$$\begin{aligned}
\ln\left(\frac{1-\pi_{ii}^e}{\pi_{ii}^e}\right) &= \gamma_e \ln\left(\frac{V_{J,i}^e}{V_{ii}^e}\right) \\
&= \frac{\gamma_e}{\eta_e} \ln\left[\left(\exp\left(-\tau_{ik}^e\right)V_{ik}^e\right)^{\eta_e} + \sum_{j\in J\setminus\{k\}}\left(\left(\frac{\pi_{ij}^e}{\pi_{ik}^e}\right)^{\frac{1}{\eta_e}}\left(\frac{V_{ik}^e}{V_{ij}^e}\right)V_{ij}^e\right)^{\eta_e}\right] - \gamma_e \ln V_{ii}^e
\end{aligned}$$

Finally, the migration cost between city i and city k for a worker of type e can be expressed as:

$$\begin{aligned}
\tau_{ij}^e &= \ln V_{ij}^e - \ln V_{ii}^e + \frac{1}{\eta_e}\ln\left(\frac{1-\pi_{ii}^e}{\pi_{ij}^e}\right) - \frac{1}{\gamma_e}\ln\left(\frac{1-\pi_{ii}^e}{\pi_{ii}^e}\right) \\
&= \ln\left(\frac{w_j^e}{w_i^e}\right) - \alpha_h^e \ln\left(\frac{r_j}{r_i}\right) + \frac{1}{\eta_e}\ln\left(\frac{1-\pi_{ii}^e}{\pi_{ij}^e}\right) - \frac{1}{\gamma_e}\ln\left(\frac{1-\pi_{ii}^e}{\pi_{ii}^e}\right)
\end{aligned} \tag{A23}$$

D Other Estimation

D.1 Numerical Simulation of Production Function

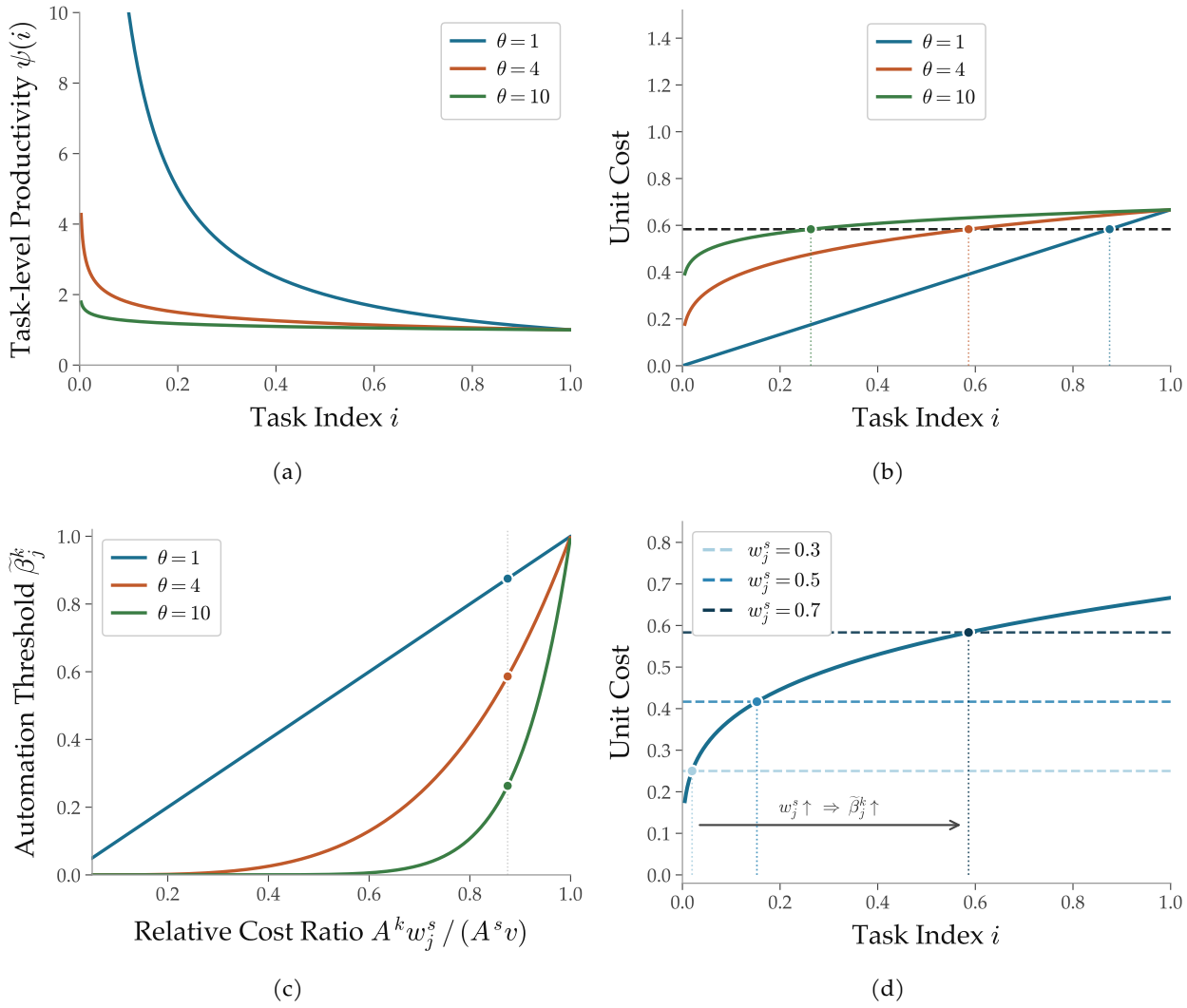


Figure A15: Numerical Simulation of Production Function

Notes: Panel (a) plots task-specific AI productivity $\psi(i) = i^{-1/\theta}$ for $\theta \in 1, 4, 10$. The function is strictly decreasing and convex in the task index i: low-index tasks are highly amenable to automation, whereas high-index tasks confer little productive advantage to AI software. Panel (b) compares the AI unit cost $v/(A^k \psi(i))$ with the skilled-labor unit cost w_j^s/A^s (dashed line); filled circles mark the automation threshold $\widetilde{\beta}_j^k$ at which the two unit costs are equal. Panel (c) plots the closed-form threshold $\widetilde{\beta}_j^k = (A^k w_j^s / A^s v)^\theta$ as a function of the relative cost ratio $A^k w^s j/(A^s v)$. Panel (d) illustrates the comparative static $\partial \widetilde{\beta}_j^k / \partial w_j^s > 0$: a higher skilled wage shifts the labor unit-cost line upward and expands the set of tasks reallocated to AI software. All panels use $A^s = 1.2$, $A^k = 1.5$, and $v = 1.0$.

D.2 Migration Cost and Hukou System

In this section, we recover model-implied migration costs and estimate their elasticity with respect to Hukou stringency across worker types.

We first derive migration costs as a function of observables and estimated parameters in Appendix C.4. Having estimated the labor supply elasticities, we recover the bilateral migration costs relying solely on the normalization that stayers face zero migration costs ($\tau_{ii} = 0$).[18] We next estimate the elasticity of these costs with respect to hukou stringency, since the hukou system is a primary determinant of destination-specific migration barriers in China. The remainder of this section summarizes the institutional background on hukou and describes our estimation procedure and results.

Hukou System China's hukou system restricts migrants' access to local public goods and social services. Each individual is assigned a place-based hukou at birth, determined by their parents' registration status and geographic location. (Song, 2014) This population management policy aims to regulate urban growth by limiting internal migration, particularly to large cities, to mitigate congestion externalities and pressure on local resources. Consequently, hukou restrictions tend to be most stringent in the largest metropolitan areas.

The elasticity between migration cost and hukou stringency is estimated using the following equation:

$$\tau_{ij} = \beta_0 + \beta^e_{\text{hukou}} \times \text{HukouStringency}^e_j + \beta^e_{Dist} \times \ln \text{Distance}_{ij} + \epsilon_{ij} \tag{A24}$$

where $\ln \text{Distance}_{ij}$ denotes denotes the log geographical distance between origin city i and destination city j, $\text{HukouStringency}^e_j$ denotes the hukou stringency index for workers of type e in destination city j, as constructed by Zhang et al. (2019) using textual analysis of local hukou policy documents. We estimate this equation by applying an exponential transformation to the dependent variable and using a Poisson regression model.

However, municipal hukou stringency is inherently endogenous, as local governments may adjust residency restrictions in response to anticipated migration flows. To address this concern, we follow Imbert et al. (2023) and instrument for hukou stringency using historical grain output per developable urban area in 2000. The intuition is that the hukou system was historically used to manage urban food security. Consequently, cities facing tighter food supply constraints adopted more stringent hukou policies.

[18] Compared with the approaches in Fan (2019) and Morten and Oliveira (2024), our method does not require imposing additional functional form assumptions on migration costs (Imbert et al., 2023).

The instrumental variable for hukou stringency of city j is constructed as follows:

$$Z_j^{\text{hukou}} = \ln\left(\frac{\text{Grain Output}_{j,2000}}{\text{Urban Area}_{j,2000}^{\text{Developable}}}\right)$$

where $\text{Grain Output}_{j,2000}$ represents the total grain output of city j in 2000, and $\text{Urban Area}_{j,2000}^{\text{Developable}}$ denotes the total developable urban area of the city in 2000.

Table A10: Migration Cost and Hukou Stringency

	τ_{ij}^e		$\text{HukouStringency}_j^e$
	Poisson (1)	Poisson-IV (2)	First-stage OLS (3)
Panel A: Skilled Labor			
$\text{HukouStringency}_j^s$	1.675	1.155	
	(0.443)	(0.470)	
Z_j^{hukou}			-0.061
			(0.014)
$\ln(\text{Distance}_{ij})$	0.091	0.123	0.025
	(0.048)	(0.052)	(0.008)
First-stage F Stat			20.070
Observations	1,539	1,539	1,539
Panel B: Unskilled Labor			
$\text{HukouStringency}_j^u$	1.079	1.633	
	(0.171)	(0.425)	
Z_j^{hukou}			-0.043
			(0.018)
$\ln(\text{Distance}_{ij})$	0.116	0.096	0.019
	(0.036)	(0.041)	(0.008)
First-stage F Stat			5.674
Observations	3,229	3,229	3,229

Notes: Columns 1–2 report Poisson and Poisson-IV estimates of the migration-cost equation, while Column 3 reports first-stage OLS estimates for the hukou stringency instruments. Standard errors are reported in parentheses and clustered at the destination-city level. All regressions are weighted by each city's 2000 population. We construct the instrument using grain output per unit of urban area at the city level. Urban area is measured following Wu and You (2025).

Table A10 reports the estimates of migration cost elasticities with respect to hukou stringency. We find that $\beta_{\text{hukou}}^u > \beta_{\text{hukou}}^s$, indicating that hukou stringency contributes more to the migration costs faced by unskilled labor. This aligns with the fact that skilled workers generally face lower migration barriers and can obtain local hukou more easily, as many Chinese cities actively implement policies to attract skilled migrants.

E Robustness

E.1 The Definition of Migrants

In our main analysis, we classify individuals as migrants if their current residence differs from their hukou (household registration) city. We exclude long-term migrants who left their hukou location more than 10 years ago, as their initial migration decision may no longer reflect current economic conditions. In addition, we include all migrants regardless of their stated motive (work vs. other reasons), and we pool rural and urban hukou holders into a single category.[19]

A potential concern with this measurement is that sample selection may systematically vary by skill group. We examine this in Table A4 by reporting migration duration, stated motives, and hukou type within each skill group.

First, Panel A of Table A4 reveals no systematic heterogeneity in migration duration across skill types, with approximately 20–25 percent of all migrants having left their hukou city more than a decade ago. Second, Panel B demonstrates that the share of unskilled workers migrating for employment is 20 percentage points higher than that of skilled workers. This difference stems mainly from migration for education, improved living conditions, and marriage. To assess the sensitivity of our labor supply elasticity estimates, we restrict the sample to include only those migrating for work, ensuring consistency with our model where migration is driven by wages. We find a quantitatively similar labor supply elasticity.

[19] Since the 2015 Census does not explicitly record agricultural hukou status, we define an individual as a rural migrant if they hold land in a rural area.

E.2 Rural and Urban Unskilled Workers

A further consideration is whether aggregating rural and urban migrants masks important heterogeneity, potentially biasing our estimates.

Panel C of Table A4 reveals that 58 percent of unskilled migrants hold a rural hukou, compared to only 22 percent of skilled migrants. A potential concern is that these individuals might still be engaged in agriculture, rendering them incompatible with an urban production framework. To address this, Table A11 demonstrates that less than 5 percent of unskilled laborers are actually employed in the agricultural sector. Furthermore, when disaggregating by hukou type, we find no systematic difference between rural and urban unskilled labor: both groups exhibit an agricultural employment share of approximately 4 percent.

Table A11: Distribution of Unskilled Employment by Industry

	Full Sample		Rural		Urban	
Industry Type	Freq.	Share (%)	Freq.	Share (%)	Freq.	Share (%)
Non-Agricultural	64,093	95.71	37,366	95.57	26,727	95.90
Agricultural	2,874	4.29	1,730	4.43	1,144	4.10
Total	66,967	100	39,096	100	27,871	100

Notes: The full sample is restricted to unskilled workers.

Another concern with pooling unskilled workers by rural versus urban hukou status is that these groups may participate in distinct labor and housing markets. For example, rural unskilled migrants may sort into different occupations or sectors (e.g., disproportionately into construction) and may exhibit different housing consumption patterns (e.g., living in dormitories) than urban unskilled workers.

We then verify the industrial shares of rural unskilled and urban unskilled labor. Figure A16 shows that the industrial distributions of rural unskilled and urban unskilled labor are highly correlated, with an R^2 of 0.945. This evidence alleviates concerns that the two groups participate in fundamentally different labor markets.

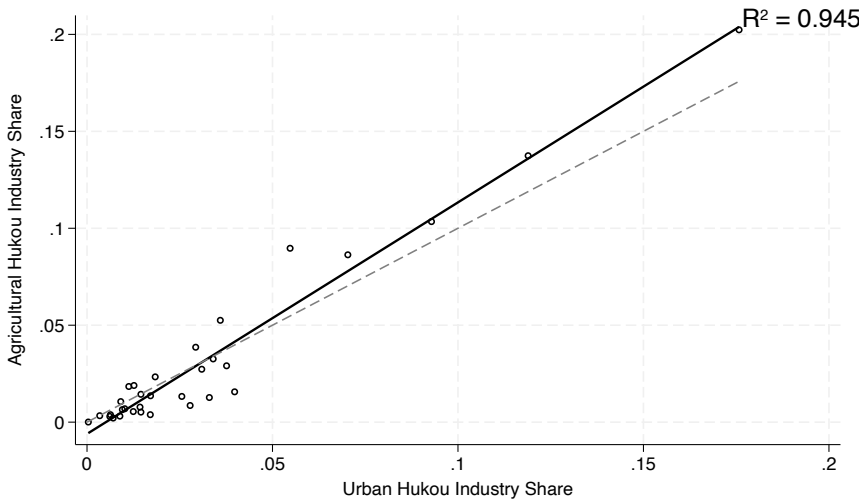


Figure A16: Sectoral Sorting of Labor by Hukou Status

Notes: This figure plots the industrial distribution of agricultural and urban hukou unskilled labor, using data from the 2017 CMDS. The sample is restricted to migrants who moved after 2010.

Regarding housing expenditure, we calibrate the housing share in household expenditures using the estimates reported in Table E.2 of Chen et al. (2023). Within the unskilled category, the difference in housing expenditure shares between rural migrants and urban unskilled workers is negligible.

We also find that both types of unskilled workers exhibit a slightly higher average housing expenditure share than skilled workers. Accordingly, we calibrate the housing budget shares for skilled and unskilled labor to 0.24 and 0.26, respectively, following Chen et al. (2023). Although these values are quantitatively similar, the modest difference between skill groups is sufficiently captured by our skill-specific parameters.

Table A12: Share of Housing Expenditure

Population	**Small city** (0.5 million)	**Median city** (1 million)	**Average city** (2 million)	**Mega-city** (Shanghai)
Panel A: Share of housing in households' expenditure				
High-skilled				
Not allowing for city area expansion	0.182	0.205	0.228	0.310
Allowing for city area expansion	0.182	0.205	0.228	0.310
Low-skilled				
Not allowing for city area expansion	0.199	0.223	0.247	0.334
Allowing for city area expansion	0.199	0.223	0.247	0.334
Rural migrants				
Not allowing for city area expansion	0.238	0.254	0.269	0.323
Allowing for city area expansion	0.238	0.254	0.269	0.323

Notes: This table is adapted from Table 4 in Chen et al. (2023). It reports the housing share of total household expenditure for skilled, unskilled, and rural migrant households.

E.3 Consistency between Model

Our empirical strategy separately estimates labor demand elasticities, labor supply elasticities, and the elasticity of migration costs with respect to hukou stringency. While each estimation block relies on standard within-block exclusion restrictions, integrating the resulting parameters into a spatial equilibrium framework additionally requires that the identifying assumptions in one block do not violate those of the others.

Table A13: Instruments across Estimation Blocks

Block	Dependent var	Endogenous var	Instruments
Labor demand	AI software–output ratio; skilled wage; skill premium	Change in local skill ratio	Thermal inversions; agricultural-shock-induced migration
Labor supply (Lower nest)	Bilateral migration share	Destination real wage	Trade shocks; trade shocks × share of undevelopable land
Labor supply (Upper nest)	Out-migration rate	Inclusive value minus local real wage	TFP shocks; tfp shocks × share of undevelopable land
Migration costs	Recovered migration cost	Hukou stringency	Grain output per developable urban area

First, we note that wages appear as both the dependent variable in the labor demand estimation and the independent variable in the labor supply estimation. Identifying the labor supply elasticity using real wage changes induced by trade and TFP shocks relies on a key exclusion restriction. Specifically, these shocks, and their interactions with the share of undevelopable land, must shift migration behavior only through the real wage channel, rather than through migration costs or other non-wage determinants of location choice.

In the labor demand estimation, we instrument changes in local skill ratios using thermal inversions and agricultural-shock-induced migration flows. As derived in Appendix C.3, the residuals in our labor demand equations consist solely of AI-induced technical change and local productivity. Therefore, our identification relies on the assumption that agricultural-shock-induced migration flows do not systematically correlate with local technical change, affecting wages only through changes in the skill ratio. Similarly, for thermal inversions, the identifying assumption is that pollution affects local amenities (and thus migration costs), which subsequently alters wages strictly through changes in the skill ratio, rather than through direct effects on local technical change. While thermal inversions could potentially violate the exclusion restriction by directly affecting labor productivity, our use of 2010–2015 differences mitigates this concern. Over this short five-year period, thermal inversions are unlikely to systematically shift underlying labor

productivity.

By summarizing these potential sources of variation, we find no evidence of cross-block inconsistency. This is primarily because we exploit entirely distinct sources of variation across our estimation steps: we rely solely on demand-channel instruments to estimate labor supply, and supply-channel instruments to estimate labor demand. Consequently, these orthogonal strategies ensure consistency and do not violate each other's identifying assumptions.

E.4 Model Extension

In our theoretical model, we assume that both skilled and unskilled workers are employed in the tradable sector. However, Table A5 shows that most migrants are employed in the consumer service sector. Motivated by this evidence, we extend the framework to incorporate a non-tradable sector. We assume this sector hires only local labor and produces goods that are consumed locally.

The production of non-tradable sectors is given by:

$$Y_j^{CS} = A_j^{CS}\left[(U_j)^{\zeta} + A^s(S_j)^{\zeta}\right]^{\frac{1}{\zeta}} \tag{A25}$$

Consumers consume tradable goods (the numeraire), local non-tradable services, and housing. We assume housing is a necessity. We assume consumers have Price-Independent Generalized Linearity (PIGL) preferences. The indirect utility of a consumer in city j with income w_j is defined as:

$$V(e_r, P_r) = \frac{1}{\varepsilon}\left[\frac{w_j}{r_j^{\psi}\left(p_j^{CS}\right)^{1-\psi}}\right]^{\varepsilon} - \nu \ln\left(\frac{r_j}{p_j^{CS}}\right) \tag{A26}$$

with $\varepsilon > 0$ and PIGL shifters $\nu > 0$. p_{jCS} and r_j denote the prices of consumer services and housing, respectively.

Applying Roy's identity, the expenditure shares on housing and consumer services are given by:

$$s_j^h = \psi + \nu\left[\frac{r_j^{\psi}\left(p_j^{CS}\right)^{1-\psi}}{w_j}\right]^{\varepsilon}, s_j^{CS} = (1-\psi) - \nu\left[\frac{r_j^{\psi}\left(p_j^{CS}\right)^{1-\psi}}{w_j}\right]^{\varepsilon} \tag{A27}$$

where $s_j^h + s_j^{CS} = 1$, w_j^e denotes the wage of a type e worker in city j.