ACI Research Paper #08-2025

Physical and Human Capital Accumulation in a Spatial Economy

Lin MA

Yunlong SONG

Yang TANG

July 2025

Please cite this article as:

Ma, Lin, Yunlong Song, and Yang Tang, "Physical and Human Capital Accumulation in a Spatial Economy", Research Paper #08-2025, *Asia Competitiveness Institute Research Paper Series (July 2025)*

Physical and Human Capital Accumulation in a Spatial Economy

Lin Ma Yunlong Song Yang Tang *

July 7, 2025

Abstract

We examine how production factor accumulations in space respond to economic shocks or events, and how these responses modify the aggregate and distributional impacts of these shocks. To explore their responses and implications, we develop a dynamic spatial model that incorporates capital accumulation and skill acquisition. Focusing on China's trade liberalization and infrastructure expansion in the early 2000s, we show that allowing for capital accumulation amplifies welfare gains and intensifies between-skill inequality, while skill acquisition attenuates impacts on skill premiums by balancing skill supply in response to shocks. Most importantly, our findings highlight the critical role of capital-skill interactions in shaping aggregate and spatial impacts, suggesting that both capital and skill adjustments are essential to understanding the full impacts of economic shocks/events on welfare and inequality.

Keywords: international trade; skill premium; economic geography; capital accumulation

JEL Classification: F12;O11;R12

*Respectively: School of Economics, Singapore Management University (linma@smu.edu.sg); Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy, National University of Singapore (ylsong@nus.edu.sg); School of Social Sciences, Nanyang Technological University (tangyang@ntu.edu.sg). A previous version of this paper circulated under the title "The Long and Short Run Spatial Impacts of Trade". We are grateful for the stimulating discussions with Davin Chor, B. Ravikumar, Dimitrijie Ruzic, Liugang Sheng, Daniel Xu, and various seminar and conference participants in Jinan Summer Trade Workshop (2023), AMES (2023), OASIS, and Shanghai Jiaotong University.

1 Introduction

A central question in spatial economics is how economic shocks or events propagate within a country. In the short run, the answers depend on geography, industrial composition, and factor mobility. In the long run, factor accumulations such as physical and human capital adjustments might play increasingly important roles. For example, a trade liberalization can reduce the cost of sourcing investment goods and hence increase the rate of capital accumulation, magnifying the welfare impact of the trade shock over time[1]. The same shock might encourage or discourage human capital accumulation, depending on whether the country has a comparative advantage in the skilled or unskilled sectors. Moreover, factor accumulations interact with each other: faster physical capital accumulation could affect the return to skill, leading to faster human capital accumulation and vice versa. Understanding the rich dynamics of both physical and human capital accumulation and how they respond to and propagate economic shocks is a question that requires careful modeling of both forces in general equilibrium. Such a framework is currently lacking in the literature.

To study spatial factor accumulations and their roles in determining the aggregate and distributional impacts of economic shocks, we develop a dynamic spatial model with endogenous physical and human capital accumulation. In particular, we incorporate heterogeneous workers and endogenous skill acquisition into a dynamic spatial framework with forward-looking migration and capital accumulation decisions (Caliendo et al., 2019; Kleinman et al., 2023). In the model, workers are born unskilled and can choose to become students and, subsequently, skilled workers for the rest of their lifetimes. Acquiring skills incurs two costs: a fixed cost that captures the financial and psychological costs of skill acquisition and an opportunity cost of foregone income during one's spell as a student. Upgrading skills brings in higher income and lowers migration costs, as low-skilled workers might face discrimination in destination locations due to policy barriers. On the supply side, we introduce capital-skill complementarity to anchor the interaction between physical and human capital accumulations. The accumulation of capital stock in each location is determined by the forward-looking investment decision of the landlord as in Kleinman et al. (2023). Each location's skill composition is determined by the inflow

[1] For instance, see Ravikumar et al. (2019) and Artuc et al. (2022) for discussions on how trade impacts depend on capital accumulation.

of skilled and unskilled workers as well as the skill upgrading decisions made by local unskilled workers.

The model highlights the critical roles played by the interaction of physical and human capital accumulation in propagating economic shocks across space. Consider a hypothetical productivity shock that uniformly increases the local productivity of all locations in a country. The productivity boom reduces investment costs and promotes physical capital accumulation. More capital stocks boost production and real income. Thus, allowing investment amplifies the total welfare gain compared to a traditional spatial model without capital accumulation. In addition, due to the faster rate of capital accumulation, the productivity shock increases the marginal return of skilled workers more than that of unskilled workers, therefore increasing the skill premium. More importantly, physical and human capital accumulation can interact positively with each other in response to the shock: skill acquisition strongly enlarges welfare gain when capital accumulation is present, but almost does not affect welfare change in the absence of capital accumulation. The reason is that, with endogenous capital adjustment, skill upgrading magnifies the shock-induced capital gain due to capital-skill complementarity, and the additional capital gain boosts welfare gain.

In the above example, incorporating production factor accumulations has rich implications for quantitative analysis. Some questions arise naturally given the previous example. How do factor adjustments change the impacts on other outcomes, such as skill premium? How do factor accumulations affect the impacts of other shocks, such as trade liberalization and infrastructure improvement? How do they change the shock's spatial impacts, such as shock-driven migration, the spatial distribution of the local impacts, and winners and losers in space? We aim to provide answers to these questions.

After developing the model, we quantify it in the context of China, a country that experienced drastic trade liberalization and massive infrastructure investment in the past two decades. We model four sectors that differ in trade costs and factor intensity, which, together with factor endowments, determine the comparative advantage of China. The geographical units are mapped to "prefectures" in China. We utilize various datasets to invert the model to recover locational fundamentals and structurally estimate the skill upgrading and type-specific migration costs. Migration costs consist of geographic travel costs and type-specific policy barriers. Consistent with previous literature, policy barriers in China for skilled and unskilled workers are significant, equivalent to 2.2 and 2.5 times respectively the geographic travel costs between an average city-pair within China, with unskilled workers facing higher policy barriers. The skill upgrading costs are also considerable, equivalent to 74 percent of an unskilled worker's lifetime utility.

To highlight the roles of factor accumulation in explaining the aggregate and spatial impacts of economic shocks, we focus on three key questions: 1) In the absence of skill acquisition, does an economic shock impact the economy equally with and without endogenous capital accumulation? 2) In the absence of physical capital accumulation, does the same shock impact the economy similarly with and without skill upgrading? and 3) Conditional on capital accumulation (upskilling), does adding skill upgrading (capital accumulation) affect the quantitative impacts of the shock? To answer these questions, we consider four model setups: the benchmark model with physical and human capital accumulation, the model with skill acquisition but no capital accumulation, the model with endogenous capital accumulation but no upskilling, and the bare-bone model without both factor accumulations. For each model setup, we compare the baseline economy that captures factual economic shock to a counterfactual economy without the shock. We measure the aggregate and spatial impacts of a shock on capital stocks, skill ratio, welfare, and skill premium, and explore how the impacts change across different model setups.

In quantitative analysis, we focus on two historical events happening in early 2000s China: China's accession to the WTO in 2001 and its large-scale investment in domestic infrastructure during 2000-2015. Back in the early 2000s, the two facts from the data that 1) China is a capital-scarce country relative to the rest of the world (ROW), and 2) the skilled sector is less capital-intensive than the unskilled one, imply that China had a comparative advantage in the skilled sector. Thus, in a conventional quantitative trade model, trade liberalization favors skilled workers in China due to the Stolper-Samuelson theorem. In the counterfactual experiment using the bare-bone model, we find a similar result that the average unskilled real wage increases by 1.2 percent, but the skilled real wage increases by 1.7 percent in steady states. Endogenizing factor accumulations significantly amplifies welfare gain, particularly for skilled workers: the impacts on unskilled and skilled real wages become 1.6 and 2.7 percent, respectively. Such an amplification effect is mainly due to accelerated capital formation fueled by lower investment costs following trade liberalization. Ignoring capital adjustment will shut down this channel and underestimate the welfare impact of trade.

More interestingly, capital accumulation and skill acquisition interact negatively with each other in response to the trade liberalization: endogenous capital accumulation diminishes the trade-driven skill upgrading, as capital gains from trade weaken China's comparative advantage in the skilled sector; allowing skill adjustment also dampens the trade-induced capital growth, because skill upgrading strengthenes China's comparative advantage in skilled sector, which utilizes less capital input than unskilled sector. Although capital-skill complementarity implies a positive interaction between capital and skill formation, it turns out that the patterns of comparative advantage are more important in determining the factor adjustments interactions when evaluating the impacts of trade.

Regarding spatial impacts of trade, factor accumulations impose strong heterogeneous spatial impacts and substantially differentiate the winners and losers within China generated by the trade shock. Due to the coastal cities' geographic closeness to the ROW, coastal cities in China reap the most gains from trade liberalization. They attract more migrants, produce more, and provide higher real wages than inland cities. Allowing capital and skill adjustment reinforces their locational advantages: the gap between the top winner and the last winner in terms of the unskilled real wage increase widens by 3 times, and the gap for skilled real wage increase widens even larger, by over 4 times. Moreover, factor adjustments enlarge domestic migration responsiveness to trade liberalization. For example, one of the top coastal winners, Dalian, is 7 times more effective in attracting migrants than it would be if there were no factor accumulation, while one of the inland losers in workforce, Chengdu, loses 5 times more workers [2].

Factor accumulation plays a similar role in explaining the impacts of infrastructure improvement in China. It magnifies welfare gains and enlarges the spatial heterogeneity in local impacts. Moreover, the impacts of infrastructure expansion on skill premium critically depend on factor adjustments. The expansion is close to skill-neutral in the bare-bone model simulation, as it only slightly increases the average skill premium in China by 0.2 percent. However, it becomes strongly skill-biased conditional on capital formation,

[2] More specifically, in the bare-bone model simulation and the benchmark model simulation, the impacts of the local population in Dalian are 1.3 and 9.9 percent increase respectively, while these in Chengdu are −1.1 to −5.1 percent respectively.

raising skill premium by 1.3 percent, as the infrastructure expansion drives capital growth, which in turn increases the return of skill due to capital-skill complementarity. The shock becomes slightly unskill-biased under the simulation with skill acquisition but no capital accumulation: skill premium reduces by 0.4 percent because of the infrastructure-induced larger supply of skill. Finally, incorporating both keeps the shock skill-biased as skill premium increases by 1.2 percent — the positive force from capital accumulation dominates the negative one from upskilling.

This paper mainly speaks to the literature on quantitative spatial models (Allen and Arkolakis, 2014; Ahlfeldt et al., 2015; Caliendo et al., 2019; Allen and Arkolakis, 2022; Kleinman et al., 2023). The closest to this paper is Kleinman et al. (2023), relative to which we introduce endogenous skill acquisition and show rich interactions between factor accumulations across space. Our model is well-suited for studying skill premiums, as it incorporates multiple sectors, multiple production factors, and capital-skill complementarity into this strand of models.

It is also related to a broad literature investigating the spatial impact of trade (Feenstra and Hanson, 1996; Goldberg and Pavcnik, 2007; Helpman et al., 2010; Parro, 2013; Autor et al., 2013, 2021), infrastructure investment (Faber, 2014; Donaldson and Hornbeck, 2016; Banerjee et al., 2020), and other economic shocks such as internal migration liberalization (Bryan and Morten, 2019). Many papers in this literature assume fixed endowments of production factors and abstract away from the spatial and intertemporal dimensions. A strand of papers particularly related to our research is the trade models that endogenize factor endowments at the aggregate level, such as Findlay and Kierzkowski (1983), Borsook (1987), Falvey et al. (2010), and Blanchard and Willmann (2016). Our paper introduces the space and time dimension to this literature. We show that factor endowments respond to economic shocks differently across space within a country, and their response to shocks is conditional on other factors' response.

Finally, we also contribute to the literature studying China's spatial economy (Fan, 2019; Tombe and Zhu, 2019; Ma and Tang, 2020, 2024; Cai et al., 2022). Our work is closest to Fan (2019), which considers the spatial impacts of shocks in a static model with exogenously determined capital stock and skill supply. Relative to Fan (2019), we endogenize capital accumulation and skill acquisition in a dynamic framework. Our results show that the endogenous response of factor accumulation could lead to drastically different quantification results.

The rest of the paper is structured as follows. Section 2 describes our dynamic spatial framework; Section 3 takes our model to China's economy and shows how to calibrate the model; Section 4 discusses our quantitative results, and Section 5 concludes.

2 Model

In this section, we introduce a dynamic spatial model that features physical and human capital accumulation across space. Time is discrete and indexed by $t = 0, 1, 2, \cdots, \infty$. The economy has N geographically segmented locations indexed by i or n. Throughout the paper, we use i to index the origin and n to index the destination. The economy also contains $j = 1, \cdots, J$ sectors that differ in capital and skill intensity. Locations can trade with one another, and workers can migrate between locations subject to migration costs.

2.1 Workers

Workers derive flow utility u from the following consumption bundle each period:

$$u = \ln \left[\prod_{j=1}^{J} \left(\frac{c^j}{\gamma^j} \right)^{\gamma^j} \right], \tag{1}$$

where γ^j is the expenditure share on goods produced by sector j, with the assumption $\sum_j \gamma^j = 1$. The industry-level composite good, c^j, is a constant elasticity of substitution (CES) aggregator over N location-specific varieties in sector j, following Armington (1969):

$$c^j = \left[\sum_{n=1}^{N} \left(c_i^j \right)^{\frac{\theta}{\theta+1}} \right]^{\frac{\theta+1}{\theta}}, \ \theta > 0,$$

where $\theta - 1$ is the elasticity of substitution across varieties.

The model includes three types of individuals: unskilled workers (l), skilled workers (h), and students (s). Unskilled and skilled workers are employed full-time and inelastically supply one unit of their respective labor type each period. Students, whom we can also call trainees, work part-time. They contribute only a fraction of unskilled labor in

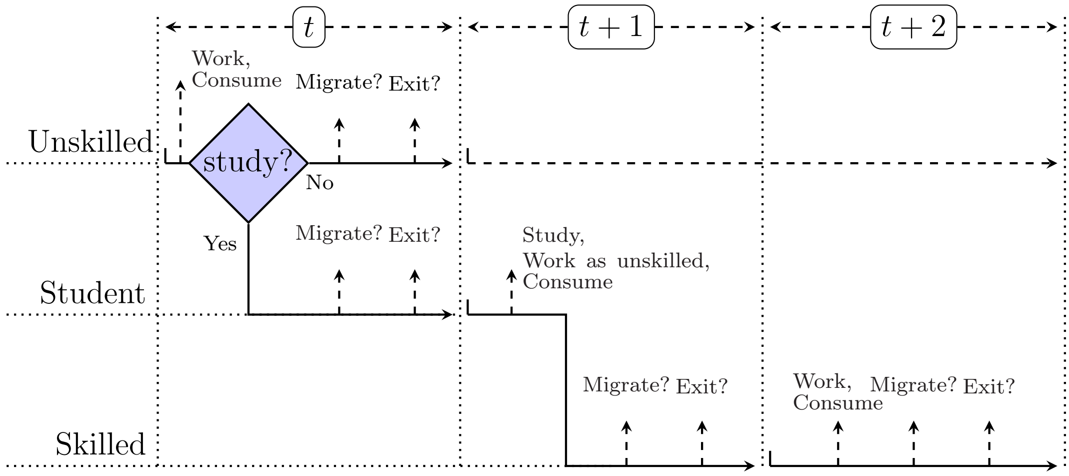


Figure 1: Individual's Timeline

the local labor market, while allocating the rest of their time to acquiring human capital. All individuals are hand-to-mouth consumers and do not save. At the end of each period, all individuals—including students—decide where to migrate for the next period, facing type-specific migration costs.

Figure 1 summarizes the timeline of individual decisions. All individuals are born as unskilled workers. In each period t, unskilled workers decide whether to enroll as students in the next period before the migration decision. If an unskilled worker decides to enroll, they migrate as a "future student" at the end of period t, spend period $t+1$ as a student, and become a skilled worker starting in period $t+2$. If the worker decides to stay unskilled, she starts period $t+1$ as unskilled again, and the process repeats. Lastly, at the end of each period, all individuals face an exogenous exit shock. Exiting workers are replaced by new-born unskilled workers in the following period so that the total population in the model stays constant. In the rest of the section, we provide more details on the individuals' decision problems.

Skilled Workers We start with the simplest recursive problem of a skilled worker at location i. Skilled workers only make migration but not education decisions, and therefore, their recursive problems adopt the familiar format as seen in Caliendo et al. (2019):

$$v_{it}^{h} = \ln\left(b_{it}\frac{w_{it}^{h}}{p_{it}}\right) + \max_{\{n\in N\}}\left\{\xi\beta\mathbb{E}\left[v_{nt+1}^{h}\right] - \kappa_{ni,t}^{h} + \rho\varepsilon_{nt}\right\}. \tag{2}$$

In the expression above, v_{it}^h is the value of a skilled worker at location i, period t. The first term on the right-hand side of the equation is the indirect current utility, in which b_{it} is the amenity, w_{it}^h is the wage rate for skilled workers, and $p_{it} = \prod_{j=1}^{J} \left(p_{it}^j\right)^{\gamma^j}$ is ideal price index at location i.

The second term captures the option value of living in location i as a skilled worker: it is the expected value of the next period conditional on the worker's migration decision. The worker chooses a future location n, considering the expected value of living in location n in period $t+1$, as captured in the term $\mathbb{E}\left[v_{nt+1}^h\right]$, against the cost of moving from i to n captured in the term $\kappa_{ni,t}^h$.[3] As is standard in the dynamic discrete choice literature, we also introduce an idiosyncratic preference shock $\{\varepsilon_{nt}\}_{n=1}^N$ that is i.i.d across workers, locations, and time. The shock follows a Gumbel distribution with the cumulative distribution function (CDF): $F(\varepsilon) = e^{e^{(-\varepsilon-\bar{\gamma})}}$, where $\bar{\gamma}$ is the Euler-Mascheroni constant. The parameter ρ captures the dispersion of the Gumbel shock, and could be interpreted as the inverse of the migration elasticity. Lastly, $\beta \in (0,1)$ is the discount rate against the future, and $\xi \in (0,1]$ is the probability of surviving into the next period. A non-surviving worker in city i is replaced by an unskilled new worker in $t+1$ at the same location.

Unskilled worker and student We now turn to the problem of the unskilled workers and describe their migration and education decisions. An unskilled worker living in location i at time t maximizes the recursive value function v_{it}^l by making education and migration decisions sequentially as defined by the following system of equations:

$$v_{it}^l = \ln\left(\frac{b_{it} w_{it}^l}{p_{it}}\right) + \max\left\{\tilde{v}_{it}^l + \psi\zeta^l, \tilde{v}_{it}^s - \omega + \psi\zeta^s\right\} \tag{3}$$

$$\tilde{v}_{it}^d = \max_{\{n \in N\}} \left\{\xi\beta\mathbb{E}\left[v_{nt+1}^d\right] - \kappa_{ni,t}^d + \rho\varepsilon_{nt}\right\}, \ d \in \{l, s\}, \tag{4}$$

$$v_{i,t+1}^s = \ln\left(\frac{b_{i,t+1} \times \iota \times w_{i,t+1}^l}{p_{i,t+1}}\right) + \max_{\{n \in N\}} \left\{\xi\beta\mathbb{E}\left[v_{n,t+2}^h\right] - \kappa_{ni,t+1}^h + \rho\varepsilon_{n,t+1}\right\}. \tag{5}$$

Similar to the problem of the skilled workers, the value of location i to the unskilled worker depends on the current wage rate w_{it}^l and the ideal price index, p_{it}. Different from the problem of the skilled workers, unskilled workers make education decisions as summarized in the second term of Equation (3). If the worker decides to forgo education, they will

[3] Standard properties on bilateral migration cost $\kappa_{ni,t}^d$ apply: (1) $\kappa_{ni,t}^d > 0$ for $n \neq i$, (2) $\kappa_{ii,t}^d = 0$, and (3) $\kappa_{ni,t}^d \leq \kappa_{nj,t}^d + \kappa_{ji,t}^d$ for any third location j.

enter the migration phase with the continuation value $\tilde{v}^l_{it}$ as described in equation (4), and proceed into the next period as an unskilled worker. If the worker decides to upgrade skills, they pay a utility cost denoted as ω, become students, and enter the migration phase with the continuation value $\tilde{v}^s_{it}$. The students' continuation value differs from that of the unskilled worker in the future value function as described in Equation (5): as students, they expect to devote $(1-\iota)$ fraction of their time to study, and therefore can only supply ι fraction of their (unskilled) labor to the market in period $t+1$. Moreover, the students will become skilled workers in period $t+2$, and therefore expect a different value function and migration costs $\kappa^h_{ni,t+1}$ in the future. The unskilled worker weighs the potential benefits of upgrading in terms of wage rates and migration costs as summarized in $\tilde{v}^s_{it}$ against the costs of education as summarized in ω and ι to make the education decision. We also introduce an idiosyncratic preference shock for skill in Equation (3), denoted as $\{\zeta^d\}_{d\in\{l,s\}}$, that is i.i.d. across individuals and time. The idiosyncratic shock follows the same standard Gumbel distribution as ε, and together with the scale parameter ψ, controls the elasticity of education choices with respect to the changes in the continuation value.[4]

2.1.1 Solution to the Individual Decision Problems

We provide the solution to the individual problem discussed in the previous section here, and refer the readers to Appendix A.2 for more details. Starting from the problem of the skilled workers, and denote the expected value of $v^h_{(\cdot)}$ as $V^h_{it} \equiv \mathbb{E}\left(v^h_{it}\right)$, we show that:

$$V^h_{it} = \ln\left(\frac{b_{it}w^h_{it}}{p_{it}}\right) + \rho \ln\left\{\sum_{n=1}^{N} \exp\left[(\xi\beta V^h_{n,t+1} - \kappa^h_{ni,t})/\rho\right]\right\}. \tag{6}$$

[4] As common in the literature, the idiosyncratic shock also smoothes the binary choice problem. Without the shock, all unskilled workers at a location will choose to upgrade or stay unskilled at the same time, which is clearly at odds with the data.

Similarly, denote $V^l_{it} \equiv \mathbb{E}\left(v^l_{it}\right)$ as the expected utility of unskilled workers at location i, time t, it is straightforward to show that:

$$V^l_{it} = \ln\left(\frac{b_{it} w^l_{it}}{p_{it}}\right) + \psi \ln\left[\exp\left(\frac{\widetilde{V}^l_{it}}{\psi}\right) + \exp\left(\frac{\widetilde{V}^s_{it} - \omega}{\psi}\right)\right] \tag{7}$$

$$\text{with } \widetilde{V}^d_{it} \equiv \mathbb{E}\left[\widetilde{v}^d_{it}\right] = \rho \ln \sum_{n=1}^{N} \exp\left[\left(\xi\beta V^d_{n,t+1} - \kappa^d_{ni,t}\right)/\rho\right], \ d \in \{l, s\}$$

And lastly, the students' expected value function, $V^s_{it} \equiv \mathbb{E}\left(v^s_{it}\right)$ is given by

$$V^s_{it} = \ln\left(\frac{b_{it} \times \iota w^l_{it}}{p_{it}}\right) + \rho \ln\left\{\sum_{n=1}^{N} \exp\left[\frac{\xi\beta V^h_{n,t+1} - \kappa^h_{ni,t}}{\rho}\right]\right\}, \tag{8}$$

Following the property of discrete choice models, the share of unskilled population who choose to upgrade their skill in location i period t is

$$D^{ls}_{it} = \frac{\exp\left[\left(V^s_{it+1} - \omega\right)/\psi\right]}{\exp\left[\left(V^l_{it+1}\right)/\psi\right] + \exp\left[\left(V^s_{it+1} - \omega\right)/\psi\right]}, \tag{9}$$

where $1/\psi$ captures the skill aquisition elasticity. The migration probability for unskilled workers (l), future students (s), and skilled workers (h) from i to n at time t is given by

$$D^d_{ni,t} = \frac{\exp\left[\left(\xi\beta V^d_{nt+1} - \kappa^d_{ni,t}\right)/\rho\right]}{\sum_{n'=1}^{N} \exp\left[\left(\xi\beta V^d_{n',t+1} - \kappa^d_{n'i,t}\right)/\rho\right]}, \ d \in \{l, s, h\}, \tag{10}$$

where $1/\rho$ captures the migration elasticity repectively.[5]

2.1.2 Local Labor Supply

At the end of each period, $\left(L^l_{it} + L^s_{it} + L^h_{it}\right)(1 - \xi)$ of the workers in location i exit the model and are replaced with the same number of unskilled workers. After unskilled workers making skill upgrading choices, the number of local unskilled workers that are ready to migrate is $L^l_{it} - D^{ls}_{it} L^l_{it}$, where $D^{ls}_{it} L^l_{it}$ is population of new students in location i.

Finally, the population of unskilled workers, students, and skilled workers in each

[5] We used $\kappa^d_{ni,t}$ to denote the type-specific migration costs for the ease of exposition, with the understanding that $\kappa^s_{ni,t} = \kappa^l_{ni,t}$, for all n, i.

location evolves as follows:

$$L_{nt+1}^{l} = \xi \sum_{i=1}^{N} D_{ni,t}^{l} \left(L_{it}^{l} - D_{it}^{ls} L_{it}^{l}\right) + \left(L_{nt}^{l} + L_{nt}^{s} + L_{nt}^{h}\right)(1-\xi) \tag{11}$$

$$L_{nt+1}^{s} = \xi \sum_{i=1}^{N} D_{ni,t}^{s} D_{it}^{ls} L_{it}^{l}, \tag{12}$$

and

$$L_{nt+1}^{h} = \xi \left(\sum_{i=1}^{N} D_{ni}^{h} \left(L_{it}^{h} + L_{it}^{s}\right) \right). \tag{13}$$

The supply of unskilled labor in location n then is $\tilde{L}_{nt+1}^{l} \equiv L_{nt+1}^{l}(1 - D_{nt+1}^{ls}) + \iota L_{nt+1}^{s}$. The supply of skilled labor L_{nt+1}^{h} includes inflows of skilled workers and fresh graduates as indicated in (13).

2.2 Landlords

We follow Kleinman et al. (2023) in modeling landlords. Landlords are immobile and have access to the financial market. At the beginning of period t with an initial endowment of capital stock, the landlords optimally choose the sequences of consumption and investments to maximize their lifetime utility. Similar to workers, at the end of each period, only a fraction ξ of landlords survive into the next period. New-born landlords replace the deceased ones and inherit their capital. The landlord's lifetime utility takes the form

$$v_{it}^{k} = \sum_{s=0}^{\infty} (\xi\beta)^{t+s} \ln c_{it+s}^{k},$$

where the superscript k denotes landlords and c_{it}^{k} is the composite consumption. The logarithm form of utility flow also implies that the intertemporal elasticity of substitution is one. Landlord's budget constraint is given by:

$$r_{it} k_{it} = p_{it} \left[c_{it}^{k} + k_{it+1} - (1-\delta) k_{it}\right],$$

where r_{it} is the rate of return on capital at time t and p_{it} is the aggregate price index defined before. Following Kleinman et al. (2023), the logarithm utility flow implies a constant saving rate $\xi\beta$. The law of motion of capital stock can thus be characterized as:

$$k_{i,t+1} = \xi\beta \left(1 - \delta + \frac{r_{it}}{p_{it}}\right) k_{it}. \tag{14}$$

2.3 Production

The production side of the model follows a multi-industry Armington setup (Armington, 1969). Firms at location i, industry j, produce a single variety and operate in a perfectly competitive market. Production requires unskilled labor (L_{it}^{lj}), skilled labor (L_{it}^{hj}), and capital (k_{it}^{j}) as inputs. The production function in location i and industry j at time t features a nested CES functional form as:

$$y_{it}^{j} = z_{it} \left[(\mu^{j})^{\frac{1}{\sigma}} (L_{it}^{lj})^{\frac{\sigma-1}{\sigma}} + (1-\mu^{j})^{\frac{1}{\sigma}} (L_{it}^{ej})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

where z_{it} is location-specific productivity and L_{it}^{ej} is the equipped skilled labor that embodies both skilled labor and capital:

$$L_{it}^{ej} = \left[(\lambda^{j})^{\frac{1}{\eta}} (k_{it}^{j})^{\frac{\eta-1}{\eta}} + (1-\lambda^{j})^{\frac{1}{\eta}} (L_{it}^{hj})^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}.$$

The parameters μ^{j} and λ^{j} govern the industry-specific weights of unskilled labor and capital, respectively. σ is the elasticity of substitution between unskilled and equipped skilled labor, and η is the elasticity of substitution between skilled labor and capital. To capture capital-skill complementarity for all industries, we follow the macroeconomics literature and assume $\sigma > \eta$ so that capital is more complementary with skilled workers than unskilled ones.[6]

Both unskilled and skilled workers are freely mobile across sectors within a location. The production structure implies that the unit cost of production in industry j and location i, denoted as c_{it}^{j}, is:

$$c_{it}^{j} = \frac{1}{z_{it}} \left\{ \mu^{j} (w_{it}^{l})^{1-\sigma} + (1-\mu^{j}) \left[\lambda^{j} (r_{it})^{1-\eta} + (1-\lambda^{j}) (w_{it}^{h})^{1-\eta} \right]^{\frac{1-\sigma}{1-\eta}} \right\}^{\frac{1}{1-\sigma}}. \tag{15}$$

Combining solutions from the profit maximization problem and zero profit condition, we can further obtain payment shares to unskilled labor (ϕ_{it}^{lj}), skilled labor (ϕ_{it}^{hj}), and

[6]For more details, see Duffy et al. (2004) and Krusell et al. (2000).

capital (ϕ_{it}^{kj}) for industry j respectively:

$$\phi_{it}^{lj} = \left[1 + \frac{1-\mu^j}{\mu^j}\left(\frac{w_{it}^l}{w_{it}^e}\right)^{\sigma-1}\right]^{-1}, \tag{16}$$

$$\phi_{it}^{hj} = \left[1 + \frac{\mu^j}{1-\mu^j}\left(\frac{w_{it}^e}{w_{it}^l}\right)^{\sigma-1}\right]^{-1}\left[1 + \frac{\lambda^j}{1-\lambda^j}\left(\frac{w_{it}^h}{r_{it}}\right)^{\eta-1}\right]^{-1}, \tag{17}$$

$$\phi_{it}^{kj} = \left[1 + \frac{\mu^j}{1-\mu^j}\left(\frac{w_{it}^e}{w_{it}^l}\right)^{\sigma-1}\right]^{-1}\left[1 + \frac{1-\lambda^j}{\lambda^j}\left(\frac{r_{it}}{w_{it}^h}\right)^{\eta-1}\right]^{-1}. \tag{18}$$

We assume standard iceberg trade costs between locations. In any industry j, the price of a variety in location n imported from location i is $p_{ni,t}^j = \tau_{ni,t}c_{it}^j$. Lastly, as shown in the Appendix A.1, the price index in location n and industry j, denoted as $p_{n,t}^j$, satisfies:

$$p_{nt}^j = \left[\sum_{i=1}^{N}\left(\tau_{ni,t}^j c_i^j\right)^{-\theta}\right]^{-\frac{1}{\theta}}. \tag{19}$$

2.4 Closing the Model

Agglomeration and Congestion We assume that location-specific amenities and productivity depend on the population to allow for potential agglomeration and congestion externality. Specifically, the amenity in city i is determined by an exogenous location fundamental amenity, $\bar{b}_{it}$, together with population size $L_{it}^l + L_{it}^s + L_{it}^h$:

$$b_{it} = \bar{b}_{it}(L_{it}^l + L_{it}^s + L_{it}^h)^{\alpha_b},$$

where α_b captures the population elasticity of amenity. We assume $\alpha_b < 0$ to capture the negative externality led by congestion. Similarly, the local productivity is given by

$$z_{it} = \bar{z}_{it}(L_{it}^l + L_{it}^s + L_{it}^h)^{\alpha_z},$$

where $\bar{z}_{it}$ is the exogenous component of productivity and α_z is the population elasticity of productivity. We assume $\alpha_z > 0$ to capture the agglomeration effects.

Dynamic Equilibrium. Given initial conditions $\{L_{i0}^l, L_{i0}^s, L_{i0}^h, k_{i0}\}$ in each location, the dynamic equilibrium contains a sequence of location-specific prices $\{w_{it}^l, w_{it}^h, r_{it}, p_{it}\}_{t=0}^{\infty}$, quantities $\{L^l_{it}, L^s_{it}, L^h_{it}, k_{it}\}_{t=0}^{\infty}$ and value functions $\{v^l_{it}, v^s_{it}, v^h_{it}\}_{t=0}^{\infty}$, such that the following conditions hold:

1. Workers maximize their lifetime utility by making migration and skill-upgrading decisions.

2. Landlords maximize their lifetime utility by making consumption and investment decisions.

3. The evolution of population and capital is characterized as in equations (11) - (13), and (14).

4. Labor markets for unskilled and skilled workers and capital market clear in each location.

$$w^l_{it} = \frac{\sum_{j=1}^{J} \phi^{lj}_{it} X^j_{it}}{\tilde{L}^l_{it}} \tag{20}$$

$$w^h_{it} = \frac{\sum_{j=1}^{J} \phi^{hj}_{it} X^j_{it}}{L^h_{it}} \tag{21}$$

$$r_{it} = \frac{\sum_{j=1}^{J} \phi^{kj}_{it} X^j_{it}}{k_{it}} \tag{22}$$

   where X^j_{it} denotes total revenue earned in location i and industry j at time t.

5. Trade balance condition holds in all locations:

$$X^j_{it} = \gamma_j \sum_{n=1}^{N} \pi_{ni,t} X_{nt} = \gamma_j \sum_{n=1}^{N} \pi_{ni,t} \sum_{s=1}^{J} X^s_{nt}, \tag{23}$$

   where $\pi_{ni,t}$ denotes the trade share between origin i and destination n at time t defined in equation (30) in the Appendix A.1.

The economy's steady state is a dynamic equilibrium when all the exogenous fundamentals of the economy and endogenous variables stay constant over time. We formally define the steady state of the economy as follows:

Steady State. A steady state of the economy is an equilibrium in which the endogenous variables are constant over time: $\{w^{l*}_i, w^{h*}_i, r^*_i, v^{l*}_i, v^{s*}_i, v^{h*}_i, L^{l*}_i, L^{s*}_i, L^{h*}_i, k^*_i\}$.

3 Quantification

Our baseline model aims to capture the salient features of the Chinese economy in the early 2000s. In particular, we solve the initial static equilibrium in year 2000, allow the trade and migration frictions to evolve between 2000 and 2015 as in the data, and then solve the transition path towards a long-run steady state defined by the geography in 2015. Similar to Caliendo et al. (2019) and Kleinman et al. (2023), we do not need to assume that the initial static equilibrium is in steady state; instead, we only need to assume that the initial equilibrium in 2000 is on a transition path towards a long-run steady state. Different from Caliendo et al. (2019), the lack of inter-prefecture trade data in China prevents us from applying the dynamic hat algebra commonly used in the literature (see Ma and Tang, 2024, for a detailed discussion). Instead, we solve the model in levels and therefore rely on the estimates of inter-prefecture trade and migration costs from Ma and Tang (2024), and discipline the other parameters in the context of our model. In this section, we start with the basic geographic information and then provide an outline for calibrating and estimating the model's key parameters.

Each period in the model corresponds to five years. We quantify the model to 196 prefecture-level cities in China plus one location representing the rest of the world (ROW). This sample of 196 prefectures is the largest balanced panel across all the data sources, as explained later. The prefectures in our sample are representative: for example, they account for 92.8 percent of total output and 83.5 percent of the total urban population in China in the year 2000.

We model four broad sectors: skill-intensive and unskill-intensive manufacturing, and skill-intensive and unskill-intensive services. To do so, we first estimate the skill-intensity of each industries observed in China's 2002 Industrial Classification for National Economic Activities, following Fan (2019).[7] The skill intensity of each industry is estimated as the income share of workers with education attainments of high-school or above among all workers, using the data from the *2005 One Percent Population Survey*. Lastly, we categorize the industries with above-median skill intensity as the skilled sectors and the rest as unskilled sectors. The categorization exercise is done separately for manufacturing and service industries. Table 13 in the Appendix B.4 provides the detailed mapping

[7] We use the 2-digit classification with 82 industries, out of which 29 are manufacturing, and 53 are service industries.

between industries and the four sectors in the paper. The end results are sensible. For example, the most skilled-intensive manufacturing industries are Petroleum and Telecommunication equipment, and the most skilled-intensive service industries are Education and Finance. On the other end of the spectrum, the most unskilled-intensive manufacturing industries are Clothing and Wood&Furniture, and the most unskilled-intensive service industries, Construction and Accomodation&Food. Lastly, we assume that only the products from the manufacturing sectors are tradable, and those from the service industries are non-tradable.

3.1 Initial Conditions in Year 2000

Population The initial distribution of the population across prefectures and skill type comes from the 2000 Census. We use the urban population and define a skilled worker as one with a high school diploma or above.

Capital Stock We use the perpetual inventory method to estimate prefecture-level initial capital stocks in the year 2000. Following Zhang et al. (2004), we use investment data in *China City Statistical Yearbooks* from 1994 to 2000 to construct a panel dataset of capital stocks at the prefecture level. Specifically, the capital stock in location i at time t is given by:

$$K_{it} = (1 - \delta)K_{it-1} + I_{it},$$

where I_{it} is the real investment observed in the data, and K_{it} is the sequence of capital stock inferred using the perpetual inventory method. We compute real investment as "Nominal Investment$_{it}$ $\times$ Investment Deflator$_{it}$", where the nominal investment is proxied using "Gross Fixed Capital Formation" from the *China City Statistical Yearbooks*; the investment deflators also come from the same source. To infer the initial capital stock, we adopt the standard approach as in Young (2003) and assume capital stock in 1994 is equal to real investment in that year divided by the depreciation rate.

Rest of the World The ROW is an aggregate of 32 OECD countries. Table 11 in the Appendix lists all the countries included in the ROW. For each country, we observe population size by skills in 2000 from OECD Statistics, capital stocks in 2000 from Penn

World Table.

3.2 Geography

Solving the model in level requires information on trade and migration costs across time. To this end, we modify the estimation methods in Ma and Tang (2020, 2024) to infer these costs in the year 2000, 2005, 2010, and 2015. Conditional on the other parameters discussed later, we then solve the steady state using the geography in 2015, and solve the transition path from year 2000 towards that steady state. Along the transition path, we use the estimated trade and migration costs in the years on and before 2015, and assume these costs stay at the 2015 level in the periods after the year 2015.

Trade Costs To infer the cross-prefecture trade costs τ_{ni}, we start from the freight transportation time by road and railroad between any two prefectures estimated in Ma and Tang (2024), and convert the transportation time to trade costs using the elasticity and costs parameters estimated in Ma and Tang (2020), with the following formula:

$$\tau_{ni} = \frac{1}{\theta_T}\Gamma\left(\frac{1}{\theta_T}\right)\left(\sum_m \exp\left(-a_m d_{ni,m} - b_m\right)\right)^{-\frac{1}{\theta_T}},$$

where $d_{ni,m}$ is the freight transportation time by mode m from i to n, θ_T is the transportation mode elasticity, and a_m and b_m are mode specific cost parameters. The equation above can be micro-founded using a discrete choice model across mode of transportation (road or rail), as discussed in detail in Allen and Arkolakis (2014). We repeat this exercise to compute trade costs between any Chinese prefectures for the five-year interval from 2000 to 2015.[8]

Different from within-China trade costs, the abundance of trade data between China and the ROW allows us to infer *sector-specific* trade costs between Chinese prefectures and the ROW using methods similar to Head and Ries (2001). To do so, we start with the 27 Chinese port cities identified in Ma and Tang (2024), and assume all the port cities face the same trade cost with the ROW in sector j, year t, denoted as "$\tau^j_{ROW,t}$". Conditional

[8] Ma and Tang (2024) also estimated trade cost elasticities with respect to transportation time. However, their estimation depends on a model structure that features a discrete choice across both modes and routes. Our model abstracts away from the route choices and is therefore closer to Ma and Tang (2020).

on $\tau^j_{ROW,t}$, the trade costs between a non-port prefecture i with the ROW is given by $\tau_{i,\text{port}_i} \times \tau^j_{ROW,t}$, where "port$_i$" is the nearest port to prefecture i determined by the τ matrix within China. The trade costs between China and ROW could be sector-specific as they depend on tariff rates that vary across sectors.

We then follow Head and Ries (2001) to back out the changes in trade costs between China's port cities with the ROW from the observed trade flows, $\widehat{\tau^j_{ROW,t}} \equiv \tau^j_{ROW,t}/\tau^j_{ROW,2000}$, relative to the levels in 2000. As shown in the Appendix, the changes in trade costs can be inferred as:

$$\widehat{\tau^j_{\text{ROW},t}} = \left(\frac{\widehat{S^j}_{(\text{CN,ROW}),t} \times \widehat{S^j}_{(\text{ROW,CN}),t}}{\widehat{S^j}_{(\text{ROW,ROW}),t} \times \widehat{S^j}_{(\text{CN,CN}),t}} \right)^{-\frac{1}{2\theta}}, \tag{24}$$

where $\widehat{S^j}_{(ni),t} = S^j_{(ni),t}/S^j_{(ni),2000}$ is the changes in exports from i to n in sector j between year t and the initial year 2000, as recorded in the World Input/Output Database (WIOD). With the trade elasticity parameter θ and the observed flows, we calculate the *changes* in trade costs for each sector in each 5-year interval relative to 2000 between 2000-2015. The results show around 10 percent drops in trade costs from 2000 to 2005, due to China's accession to WTO, but only insignificant changes from 2005 to 2015. We assume trade costs after 2015, $\tau^j_{ROW,t>2015}$, are stay at 2015's level. Finally, we determine the initial levels of trade costs in 2000, denoted as $\tau^j_{ROW,2000}$, by inverting the model in the initial spatial equilibrium and exactly matching the observed trade costs in that year.[9] With the estimated $\tau^j_{ROW,t}$, we have complete trade cost matrices across all locations in all sample years.

Migration Costs Individuals can migrate across prefectures within China subject to type-specific friction, and no international immigration is possible between China and the ROW. We discipline the migration frictions in China as follows.

We start with the *One Percent Population Survey in 2005*, from which one can observe the share of migrants in prefecture n with hukou from origin prefecture i for skill type d

[9] Head and Ries (2001) directly inferred the trade costs between countries each year. We cannot directly adopt their methods because our model features a rich internal geography inside China, while Head and Ries (2001) abstracted away from internal geography. As a result, the levels of $\{\tau^j_{ROW,t}\}$ inferred using Head and Ries (2001) do not exactly align the observed and the model-simulated trade shares at the aggregate level. To ensure consistency between the baseline model and the data, we only use their methods to infer the *changes* in trade costs across years and rely on inverting the model in the initial equilibrium to back out the initial levels of trade costs.

as a fraction of the population in the origin i, denoted as $\bar{D}^d_{ni,t}$. Interpreting this statistic through the lens of the model, however, is not straightforward. Note that $\bar{D}^d_{ni,t}$ in the data is the *stock* measure of migration from i to n observed at time t, which contains waves of movers in the past. To be specific, $\bar{D}^d_{ni,t}$ can be expressed as:

$$\bar{D}^d_{ni,t} = \frac{L^d_{it} D^d_{ni,t} + \sum_{\tau=1}^{\infty} L^d_{it-\tau} D^d_{ni,t-\tau} (D^d_{nn,t-\tau})^{\tau}}{L^d_{it}}, \ d \in \{l, s\},$$

In the expression above, the first term in the numerator is the migration flow from the origin i to n at period t. In addition to the concurrent movers, the stock of migrants also includes those who moved τ periods ago $L^d_{it-\tau} D^d_{ni,t-\tau}$, and stayed in n thereafter with a probability $(D^d_{nn,t-\tau})^{\tau}$. The second term in the numerator accounts for the earlier waves of migrants retrospectively from $\tau = 1$ to distance history. In steady state, the expression above simplifies to:

$$\bar{D}^d_{ni} = \frac{D^d_{ni}}{1 - D^d_{nn}}, \tag{25}$$

where D^d_{gi} is the model-predicted migration probability in steady state:

$$D^d_{ni} = \frac{\exp\left[(\beta v^d_{n,t+1} - \kappa^d_{ni})/\rho\right]}{\sum_{n'=1}^{N} \exp\left[(\beta v^d_{n't+1} - \kappa^d_{n'i})/\rho\right]}.$$

Double differencing the equation above leads to an expression of migration flows that only depends on migration costs, κ^d_{ni}:

$$\frac{\bar{D}^d_{ni}}{\bar{D}^d_{ii}} \frac{\bar{D}^d_{in}}{\bar{D}^d_{nn}} = \frac{D^d_{ni}}{D^d_{ii}} \frac{D^d_{in}}{D^d_{nn}} = \exp\left[-\frac{1}{\rho}(\kappa^d_{ni} + \kappa^d_{in})\right]. \tag{26}$$

The equation above is similar to that in Bryan and Morten (2019), based on which one can directly infer migration costs from the observed migration flows. However, doing so implicitly assumes a symmetric costs matrix ($\kappa^d_{ni} = \kappa^d_{in}, \forall i, n$), which is at odds with the institutional conext of migration barriers in China. As well documented in the literature (see, for example, Tombe and Zhu, 2019; Fan, 2019), the hukou system enacted by the destination prefectures discourages inbound migration and therefore acts as institutional migration barrier. Moreover, the hukou system is more stringent in larger and richer destinations, so that, for example, migrating from a small inland prefecture to Shanghai is significantly harder than the other way around. To account for the asymmetry in migration frictions, we further assume that the migration cost matrix consists of a symmetric component that relies on the transportation infrastructure, and an asymmetric component that is rooted in hukou and acts as an entry barrier. In particular, we assume:

$$\kappa_{ni}^d = \kappa_n^d + \bar{\kappa}_{ni},$$

where $\bar{\kappa}_{ni} = \bar{\kappa}_{in}$ is symmetric travel cost between location i and n that depends on the passenger travel infrastructure, and κ_n^d is type-specific entry barrier for entering location n. We allow the entry barrier as type-specific, because hukou barriers often discriminate against less educated workers.

We take the symmetric travel costs from Ma and Tang (2024), and estimate the entry barriers for all locations, using the rearranged equation (26):

$$\frac{\bar{D}_{gi}^d}{\bar{D}_{ii}^d}\frac{\bar{D}_{ig}^d}{\bar{D}_{gg}^d}\exp\left(\frac{2\bar{\kappa}_{gi}}{\rho}\right) = \exp\left[-\frac{1}{\rho}\left(\kappa_g^d + \kappa_i^d\right)\right]. \tag{27}$$

Note that the left hand side of the equation above is observed conditional on a migration elasticity ρ (we will discuss the estimate of ρ later in the next section). We then estimate the entry barriers κ_n^d for each location and skill type using Poisson regression based upon equation (27).

We find that the migration frictions are substantial and, on average, higher for unskilled workers than for skilled ones. Figure 2 presents the histogram of entry barriers for unskilled and skilled workers across locations. In the figure, we normalize the entry barriers by the average bilateral geographic travel cost in China. The average migration barriers are formidable, equivalent to 2.2 times of average travel costs for skilled workers or 2.5 times for unskilled workers. Equivalently, more than two-thirds of the migration costs in China are due to the location-specific entry barriers. The higher migration barriers for unskilled workers likely come from the discriminative hukou policy.

The estimated migration entry barriers are consistent with the policy barriers implied by the Hukou reform index in Fan (2019). Figure 3 compares the average estimated barriers and the inverse of the Hukou reform index at 2005 across four different city tiers. Both barrier measures show that the province-level municipalities, such as Beijing

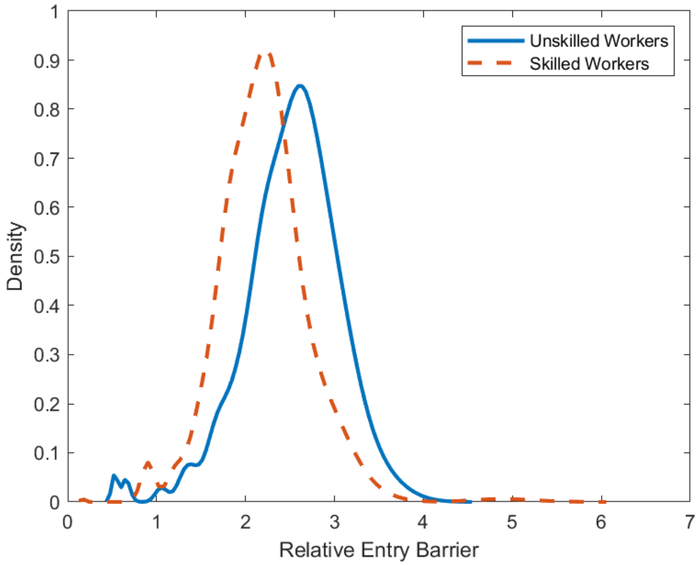


Figure 2: Distribution of Entry Barriers

Notes: This figure shows the histogram of the estimated entry barriers for unskilled and skilled workers. Entry barriers reported here are estimated using PPML and normalized by the average bilateral travel cost between any city pair in 2000 China.

and Shanghai, have the highest average entry barriers and are less open, and the sub-provincial cities, such as Shenzhen, have the lowest average barriers among the four city tiers. Finally, both measures also suggest that other provincial capital cities and smaller cities have modest policy barriers.

3.3 Parameterization

We discipline all the other parameters in the following three steps. In the first step, we externally determine some of the parameters based on the common estimates in the literature. We then infer the second set of parameters by inverting the model in the initial static equilibrium. Lastly, the we infer the parameters affecting population distribution by inverting the model along the transition path. In the rest of the section, we briefly discuss the quantification strategy of these parameters.

Pre-determined Parameters Table 1 summarizes the externally-determined parameters. The trade elasticity $\theta = 5$ comes from Costinot and Rodríguez-Clare (2014). We assume a five-year discount rate of $\beta = 0.97^5 \approx 0.86$, consistent with an annual interest rate of 3 percent. Similarly, we set the five-year depreciation rate at $\delta = 1 - 0.9^5 \approx 0.41$,

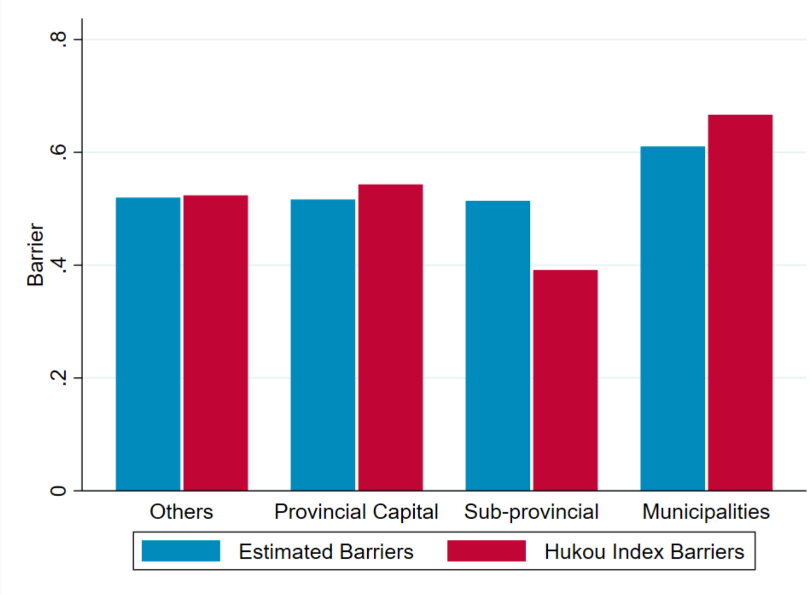


Figure 3: Estimated Barriers and Hukou Reform Index

Notes: This figure compares the estimated entry barriers by PPML and implied policy barriers from the Hukou reform index constructed in Fan (2019), across four city tiers. The "Hukou Index Barriers" is the inverse of the Hukou reform index in 2005, as a low reform index implies high barrier. The "Estimated Barriers" is the average barrier of both skill types and then normalized so that it has the same mean as the "Hukou Index Barriers". The "Municipalities" represent the provincial-level municipalities. They are the largest cities in China. One tier below are "Sub-provincial" cities. Further below are "Provincial Capital" cities.

following an annual discount rate of 10 percent from Zhang et al. (2004). The five-year migration elasticity is $\rho = 3\beta$, following Kleinman et al. (2023). From the urban literature, we assume an agglomeration elasticity of $\alpha_z = 0.1$ from Redding and Turner (2015) and a congestion elasticity of $\alpha_b = -0.3$ from Allen and Arkolakis (2022). The parameters that govern the complementarity between skill and capital stock come from the macroeconomic literature: we take the elasticities of substitution $\sigma = 2$ and $\eta = 1$ from Duffy et al. (2004). The estimation on education elasticity in the context of China is rare, and we follow Hu and Ma (n.d.) and set $\psi = 5.74$. The parameter ι, which indicates the fraction of time allocated to work during the period of being students, is assumed to be 0.2, as students usually spend 4 years to obtain a degree and each period in the model represents 5 years. Finally, according to the World Development Indicators from the World Bank[10], the average annual mortality rate in China during 2000-2020 is 0.7 percent, implying a five-year survival rate of $\xi = 0.993^5 \approx 0.965$.

[10]We use the Crude Death Rate per 1,000 people in China, 2022 revision.

Table 1: External Calibrated Parameters

Name	Value	Source	Description
α_z	0.1	Redding and Turner (2015)	Agglomeration elasticity
α_b	-0.3	Allen and Arkolakis (2022)	Congestion elasticity
β	0.86	-	Five-year discount factor
ι	0.2	-	Fraction of worktime as students
θ	5	Costinot and Rodríguez-Clare (2014)	Trade elasticity
ρ	3β	Kleinman et al. (2023)	Inverse of migration elasticity
ψ	5.74	Hu and Ma (n.d.)	Inverse of upskilling elasticity
σ	2	Duffy et al. (2004)	EoS between l and e
η	1	Duffy et al. (2004)	EoS between h and k
δ	0.41	Zhang et al. (2004)	Five-year capital depreciation rate from 0.1 annual rate
τ_{ni}	-	Ma and Tang (2024)	Bilateral trade cost
γ^j	-	China 2002 IO table	Sectoral consumption share
ξ	0.965	World Bank	Five-year mortality rate of $1-\xi$

Notes: This table reports the results of calibrated parameters in the model. These parameters either come from the literature or data.

Inverting the Initial Equilibrium In the second step, we infer the following set of parameters by inverting the model in the initial static equilibrium in the year 2000. These parameters are the location fundamental productivity $\{\bar{z}_i\}_{i=1}^N$, the sector-specific trade costs between China and ROW in 2000, $\{\tau^j_{ROW,2000}\}_{j=1}^J$, and the production function paramters $\{\mu^j, \lambda^j\}_{j=1}^J$. Similar to Kleinman et al. (2023), solving the initial static equilibrium is equivalent to finding vectors of factor prices $\{w_i^d, r_i\}_{i=1}^N$ that clear the goods and factor markets. Solving the initial static equilibrium only requires information on the initial population and capital distribution, and does not require one to solve the forward-looking migration decisions. Appendix B.6 provides more details on the definition and solution to the initial static equilibrium.

The exogenous component of prefecture-level fundamental productivity, $\{\bar{z}_i\}_{i=1}^N$, is calibrated to match prefecture-level GDP share in 2000. We normalize the fundamental productivity in the first location (Beijing) to unity so that $\bar{z}_1 = 1$. The counterpart in the data comes from *China City Statistics Yearbook* in the corresponding year. Similarly, we back out the initial trade costs between ROW and Chinese port cities in both manufacturing sectors, $\tau^j_{ROW,2000}$, by matching the observed trade-to-output ratio in each sector in China.

The parameters that capture the relative importance of unskilled workers and capital in production in each industry, μ^j and λ^j, are calibrated to match the sectoral income shares for unskilled workers and capital in the data, respectively. To allow for technology differences between the ROW and China, we estimate these parameters separately for China and the ROW. In the case of China, the sector-level income share of skilled workers comes from *2005 One Percent Population Survey*, and the share of capital in the value-added comes from China's Input-Output table in 2002. In the case of ROW, the skilled workers' income shares in each sector are computed from the IPUMS One Percent Sample. The capital income shares are represented by those of the U.S. and from the U.S. 2002 Standard Make and Use Table. Table 2 shows the estimated $\{\mu_j, \lambda_j\}_{j=1}^J$ for China and the ROW.

Table 2: Production Function Parameters

Weights	Unskilled M.	Skilled M.	Unskilled S.	Skilled S.
	China			
μ^j	0.18	0.08	0.15	0.02
λ^j	0.83	0.75	0.73	0.59
	ROW			
μ^j	0.06	0.03	0.05	0.01
λ^j	0.40	0.38	0.36	0.43

Notes: This table reports the results of production weights in four sectors for China and the ROW. The weights are calibrated in the initial static equilibrium by targeting sector-level factor income shares. $\mu \in [0,1]$ is the weight on unskilled workers and $\lambda \in [0,1]$ is the weight on capital.

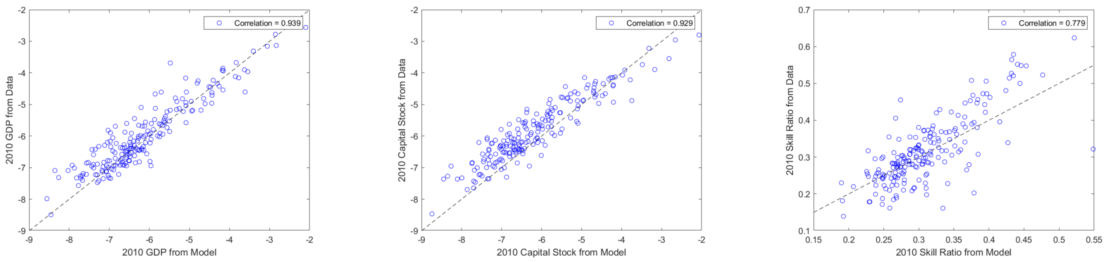
Reassuringly, our estimation shows that the unskilled sectors put more weight on unskilled workers than skilled workers. For example, μ^j, the unskilled worker intensity, equals 0.18 for unskilled manufacturing sectors in China, and 0.08 for the skilled manufacturing sector. Similar patterns exist for the Chinese service sectors (0.15 v.s. 0.02) and for ROW shown in Panel (b) of the same table. Capital intensity and skill intensity also correlate in our estimation. Capital takes up higher weights in unskilled manufacturing sectors than in skilled ones in both China and the ROW. This correlation subsequently determines the pattern of comparative advantage in the quantitative analysis presented later. The two facts that 1) the ROW is relatively more abundant in capital in the data, and 2) the unskilled sector is capital-intensive in the production functions, imply that China specializes in the skilled sector when trading with the ROW, similar to what an Heckscher-Ohlin model predicts.

Amenities and Skill Upgrading Costs The last group of parameters is calibrated by inverting the model along the transition path. These parameters are the skill upgrading cost ω and location-specific amenities $\{\bar{b}_i\}$. We need to solve the entire transition path from the initial static equilibrium towards a steady-state defined by the geography in the year 2015 in order to calibrate these parameters. This is because the data moments we rely on, as explained later, are related to the individuals' forward looking education and migration decisions in the model.

Specifically, we choose ω to match the aggregate skill ratio of 0.36 in the year 2010 ($t = 3$), as indicated by the population Census in China that year. Our calibrated skill upgrading cost is 58 percent of the average lifetime utility among unskilled workers in the initial period. The high upgrading cost comes from two patterns in the data: on the one hand, the skill premium is high in the data at 1.44 in the year 2005. On the other hand, the supply of skills had been low during the same period. Intuitively, the skill upgrading costs encompass not just the financial costs of acquiring a high school or college education but also the fierce selection induced by the strict quota system in Chinese secondary and tertiary education, manifested through the High School or College Entry Exams.

The location fundamental amenity, $\{\bar{b}_i\}$, is calibrated to match the population share of each prefecture in the year 2010. Unlike the location fundamental productivity that only requires solving the initial static equilibrium, simulating the population distribution requires solving the entire transition path in levels. Intuitively, the population distributions in any $t > 1$ are functions of future option values of each location and, therefore, require information on the entire transition path.

The quantification strategy described above aligns reasonably well with the untargeted data moments. Figure 4 compares the model-predicted spatial distribution of total output, capital stock, and skill ratio with their data counterparts, none of which is our calibration target. The model matches the data well, showing correlations ranging from 0.65 to 0.85.



(a) GDP (b) Capital Stock (c) Skill Ratio

Figure 4: Model Fit

Notes: These figures compare the baseline model simulation with the data. Each dot represents a prefecture in China, and the black dotted line is the 45-degree line. All the variables in the model and the data refer to the cross-section in the year 2010. Variables in panel (a)-(b) are in the logarithmic functions.

4 Quantitative Results

In this section, we discuss the quantitative implications of physical and human capital accumulation across space. To highlight the impacts of factor accumulation, we consider four model setups: 1) the "benchmark" model inclusive of both physical and human capital accumulation; 2) the "upskilling" model with skill acquisition but no capital accumulation; 3) the "investment" model with endogenous capital accumulation but no upskilling; and (4) the "basic" model without capital accumulation and skill acquisition. Comparing the results across these four models could reveal the contribution of each channel as well as any potential interactions.

We implement the alternative model specification as follows. In the "upskilling" simulation, unskilled workers are allowed to upskill as in the baseline model, but capital stock in each location is fixed at the initial level. In particular, we set the landlords' investment to cover the depreciated capital in each period, thereby fixing the level of capital stock. In the "investment" model, capital accumulates endogenously, but the total skill ratio in the country is fixed by assuming the: (1) skill upgrading cost κ_s is infinite; (2) exiting workers will be replaced by a new worker with the same type at the same location. These additional assumptions imply that the labor supply equations

(11)-(13) become:

$$L^l_{it+1} = \sum_{n=1}^{N} D^l_{in,t}\left(L^l_{nt} - D^{ls}_{nt}L^l_{nt}\right)$$
$$L^s_{it+1} = \sum_{n=1}^{N} D^s_{in,t}D^{ls}_{nt}L^l_{nt}$$
$$L^h_{it+1} = \sum_{n=1}^{N} D^h_{in}\left(L^h_{nt} + L^s_{nt}\right).$$

In the "basic" model, we apply both previous restrictions so that local capital stocks and China's total skill ratio are exogenously fixed.

We first analyze two simple shocks, a uniform increase in productivity and a uniform reduction in trade and migration frictions, and illustrate how the impacts of shocks depend on our model mechanisms. After understanding how the model works, we use it to quantitatively analyze two observed economic shocks: China's ascension to the WTO and its large-scale investment in domestic infrastructure.

4.1 Hypothetical Shocks

4.1.1 A Uniform Productivity Shock

We first consider a permanent and expected productivity shock that increases all Chinese cities' local productivity, $\bar{z}_i$, by 20 percent from the year 2000. We evaluate the aggregate and spatial impacts of the productivity shock under different model specifications. We focus on how the shock affects real wage, total output, and total consumption both at the aggregate and the local level.[11] Table 4 reports these results in the steady state. We will start with the spatial impacts and then move on to the aggregate impacts.

Spatial Impacts We evaluate how spatial distributions of the productivity shock's impacts depend on factor accumulations. Table 3 reports how the local impacts on pro-

[11]Following Caliendo et al. (2019), the value of an unskilled worker at i period t can be represented as $V^l_{it} = \sum_{o=t}^{\infty}(\beta\xi)^{o-t}\ln\frac{c_{io}}{(D_{ii,o})^{\rho}\left(D^{ll}_{io}\right)^{\psi}}$, where $D_{ii,o}$ is the stay rate and D^{ll}_{io} is the probability of staying unskilled. We define the consumption change Δ such that the counterfactual utility of unskilled worker $V^{l'}_{it} = \sum_{o=t}^{\infty}(\xi\beta)^{o-t}\ln\left[\left((1+\Delta)c^l_{io}\right)/\left((D_{ii,o})^{\rho}\left(D^{ll}_{io}\right)^{\psi}\right)\right] = \sum_{o=t}^{\infty}(\xi\beta)^{o-t}\ln(1+\Delta) + \sum_{o=t}^{\infty}(\xi\beta)^{o-t}\ln\left[\left(c^l_{io}\right)/\left((D_{ii,o})^{\rho}\left(D^{ll}_{io}\right)^{\psi}\right)\right] = \sum_{o=t}^{\infty}(\xi\beta)^{o-t}\ln(1+\Delta) + V^l_{it}$, where V^l_{it} is the baseline utility of unskilled worker. As any newborn will be an unskilled worker and her value incorporates skill upgrading, we interpret the consumption change of unskilled workers as the total welfare change.

duction factors, skill premium, and welfare are distributed across space within China. we use the across-location standard deviation and the range to measure the dispersion of spatial impacts. As rows 2 and 4 in the table show, the uniform productivity shock generates uniform spatial impacts given fixed local capital stocks. Adding capital accumulation induces strong heterogeneous spatial impacts, and incorporating upskilling in addition further promotes spatial heterogeneity. The reason behind these enlarged spatial heterogeneities is that initially more productive cities now are able to accumulate capital faster than undeveloped cities. Due to the faster capital accumulation, those productive cities also attract more migrants as marginal returns on labor increase. Eventually, they reap a larger gain from the productivity shock than less-productive cities. Allowing both factor accumulation strengthens their initial advantages even more due to the interaction effect, resulting in even greater spatial heterogeneity as suggested by larger standard deviations and range in the first row in table 3.

Table 3: Spatial Dispersion of Impacts of Uniform Productivity Shock

	Standard Deviation of Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	15.62	0.41	5.68	3.84	6.77	1.71	16.96	2.27
With Upskilling	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
With Investment	13.65	1.58	4.10	3.34	5.35	1.09	14.46	1.53
Basic Model	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.01

	Max %Changes - Min %Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	86.22	3.03	32.04	19.83	35.46	10.57	90.61	25.40
With Upskilling	0.00	0.01	0.07	0.02	0.03	0.01	0.05	0.05
With Investment	73.65	8.46	22.78	17.36	27.64	6.57	76.45	11.62
Basic Model	0.00	0.14	0.16	0.06	0.06	0.10	0.12	0.12

Notes: This table reports the spatial dispersions of % changes in local economic outcomes in China induced by a uniform 20% increase in productivity under different model specifications. The upper panel shows the standard deviation of perfecture-level impacts and the lower panel shows the range of perfecture-level impacts. From left to right economic outcomes are total capital stock, total skill ratio, real unskilled wage, real skilled wage, average skill premium, total GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is $100\% \times \Delta$ as defined in footnote 11.

Aggregate Impacts We then consider how the productivity boom affects the China economy as a whole, and how the impacts depend on factor adjustments. Firstly, capital accumulation augments the positive welfare impacts of productivity improvements and further worsens between-skill inequality. The welfare gains measured by real wage changes almost double (from 20.00 to 35.97 percent) for unskilled workers and triple for skilled workers (from 20.00 to 65.24 percent), and the average skill premium increases by 20.27 percent. The "welfare amplification" effect is due to the accelerated rate of capital accumulation, which arises from the fact that the productivity boom increases the real return on investment. Meanwhile, more capital stocks raise the return on skill due to capital-skill complementarity, driving up the average skill premium.

Secondly, conditional on capital accumulation, allowing upskilling magnifies the shock-driven capital growth and welfare impacts further, but attenuates the skill premium's rise. The attenuation of skill premium change is a result of two competing forces. On the one hand, the productivity shock increases the total skill ratio (by 2.29 percent) once workers are allowed to upskill, mainly due to the faster capital accumulation. This increased supply of skills weakens the shock-induced skill premium increase. On the other hand, more skilled workers also encourage more capital stock. As shown from row 3 and row 1 in table 4, the increase in total capital stock changes from 81.21 percent to 92.59 percent. This additional boost in capital stock raises the return on skill. In the end, the former force dominates the latter, diminishing the positive impact on skill premium from 20.27 to 16.92 percent.

Thirdly, in the absence of capital accumulation, whether incorporating skill acquisition does not change the aggregate impacts of the productivity shock. Most surprisingly, the productivity shock drives almost no upskilling even workers are allowed to do so. This suggests that the response of skill acquisition is closely related to capital accumulation adjustment.

The second and third results together imply that our model exhibits a strong interaction effect between capital accumulation and skill acquisition. When both factor accumulations are incorporated, they reinforce each other's response to the productivity shock, augmenting the welfare impacts. As shown in the first two columns in table 4, more capital accumulation encourages more skill upgrading, and vice versa. This interaction between factor accumulation has important implications for quantifying impacts on other

Table 4: Aggregate Impacts of Uniform Productivity Shock

	%Changes in Steady State						
	Capital	S.R.	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	92.59	2.29	47.16	72.98	16.92	91.05	54.70
With Upskilling	-	0.00	20.00	20.00	0.00	20.00	20.00
With Investment	81.21	-	34.97	65.24	20.27	71.88	35.36
Basic Model	-	-	20.00	20.00	0.00	20.00	20.00

Notes: This table reports the steady state % changes of aggregate economic outcomes in China induced by a uniform 20% increase in productivity under different model specifications. From left to right economic outcomes are total capital stock, total skill ratio, real unskilled wage, real skilled wage, average skill premium, total GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is $100\% \times \Delta$ as defined in footnote 11.

economic outcomes. For example, consider the impacts on unskilled workers' real wages. From the basic model to the full model simulation, capital accumulation and upskilling together contribute to $47.16 - 20 = 27.16$ percent of additional impact. Allowing capital accumulation alone only explains 57 percent [12] of such additional impact, and upskilling explain 0 percent. Therefore, the remaining 42 percent of the additional impacts is due to the interactive response of factor accumulations. Our finding also has an important implication for quantifying the impact on skill premium: allowing capital accumulation alone overestimates the impact and allowing upskilling alone underestimates it. Therefore, both factor accumulation channels are crucial to correctly quantify the impacts on skill premium. Ignoring either mechanism will underestimate the welfare impact and bias the impact evaluation on skill premium.

4.1.2 A Uniform Infrastructure Improvement

In this section, we consider a permanent and uniform improvement in infrastructure starting from 2015 that will reduce domestic bilateral trade costs by 20 percent. Specifically, the counterfactual domestic trade costs $\tau'_{ni,t}$ is given by $\tau'_{ni,t} = 1 + (100\% - 20\%)(\tau_{ni,t} - 1)$ for the year 2015 and afterward, where $\tau_{ni,t}$ is the baseline trade cost.

[12] $(34.97 - 20)/27.16 \times 100\% = 58\%$

Aggregate Impacts Table 5 summarizes the steady-state aggregate impacts on production factors, skill premium, and various welfare measures.

In the basic model simulation, the infrastructure improvement increases the average real wage of both skilled and unskilled workers by around 1.2 percent. The total welfare, measured by consumption change, also increases by 1.21 percent. The change in connectivity is also skill-neutral in the basic model, as it leaves a negligible impact on skill premium.

However, as shown in the third row in table 5, incorporating capital accumulation significantly enlarges the welfare impacts and renders the infrastructure shock skill-biased. More specifically, the positive impact of unskilled real wage changes from 1.20 percent to 2.01 percent, a 68 percent additional impact contributed by capital accumulation. The impact on skilled real wage amplifies even larger, from 1.21 percent to 3.28 percent. These larger welfare impacts are again due to faster capital accumulation driven by the infrastructure shock. Furthermore, the shock increases skill premium by 1.17 percent, suggesting an unequal gain from infrastructure improvement. Intuitions behind this distributional impact are as follows: as the infrastructure improvement reduces the costs of sourcing varieties from other locations (p_{it}), and hence the investment cost, it promotes capital accumulation. As shown in the table, the infrastructure improvement increases total capital stock by 3.50 percent once capital investment is possible. More capital stocks increase skilled wages relatively more due to capital-skill complementarity. Therefore, a skill-neutral infrastructure shock in the basic model becomes skill-biased once we incorporate endogenous capital adjustment.

We again find strong evidence of interactive responses of factor accumulation to a given shock. As shown in table 5, allowing upskilling alone barely changes the aggregate impacts (from row 4 to row 2), but it exerts strong influences on welfare gains and skill premium changes conditional on endogenous capital accumulation (from row 3 to row 1). Specifically, adding upskilling conditional on investment further enlarges the welfare impacts but dampens skill premium increases. The attenuation in skill premium's increase is due to upskilling balancing the skill supply in response to the shock, as many unskilled workers are now able to upgrade their skill type to enjoy the skill premium rise. However, in the absence of capital accumulation, skill upgrading barely affects the skill ratio and skill premium. In the case of unskilled workers' real wage changes, we find that the

Table 5: Aggregate Impacts of Uniform Infrastructure Improvement

	%Changes in Steady State						
	Capital	S.R.	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	3.08	0.11	2.54	3.51	0.97	2.98	1.39
With Upskilling	-	-0.00	1.23	1.20	-0.04	1.11	1.25
With Investment	3.50	-	2.01	3.28	1.17	3.28	1.31
Basic Model	-	-	1.20	1.21	-0.00	1.12	1.25

Notes: This table reports the steady state % changes of aggregate economic outcomes in China induced by a uniform infrastructure shock under different model specifications. From left to right economic outcomes are total capital stock, total skill ratio, real unskilled wage, real skilled wage, average skill premium, total GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is $100\% \times \Delta$ as defined in footnote 11.

interaction effect between factor accumulation contributes to 40 percent[13] of the total welfare gain amplification from the basic to full model simulation.

Spatial Impacts In table 6, we show that the spatial heterogeneity of the impacts also critically depends on the factor accumulation. Compared to the full model simulation, the basic model simulation exhibits much smaller spatial variations of benefits from infrastructure improvement. Adding capital accumulation significantly differentiates the top winner and the last winner. For example, the top winner only gains 1.96 percent more consumption than the last winner in basic model simulation, but the gap widens to 4.63 percent with capital accumulation. Further incorporating upskilling enlarges spatial variation slightly, but adding upskilling alone almost does not affect it, suggesting an interaction effect between capital accumulation and skill acquisition.

In summary, in this section, we use two hypothetical economic shocks to illustrate how physical and human capital accumulation determine the aggregate and spatial impacts of them. The results show that the impacts critically depend on both capital accumulation and skill acquisition. Omitting either mechanism will bias the quantification results. More importantly, there is an interaction effect at present only when both factor accumulation channels are modeled. With these results in mind, in the next section, we will look at

[13]Capital accumulation explains $(2.01 - 1.20)/(2.54 - 1.20) \times 100\% = 60\%$ of the total amplification, skill acquisition contributes $(1.23 - 1.20)/(2.54 - 1.20) \times 100\% = 2\%$, and the interaction between capital and upskilling contributes the remaining 38%

Table 6: Spatial Dispersion of Impacts of Uniform Infrastructure Improvement

	Standard Deviation of Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	3.03	0.18	1.78	0.99	1.28	0.36	0.18	0.46
With Upskilling	0.00	0.02	0.53	0.40	0.33	0.10	0.02	0.20
With Investment	2.82	0.33	1.51	0.92	1.13	0.27	0.18	0.42
Basic Model	0.00	0.13	0.61	0.39	0.35	0.05	0.02	0.20

	Max %Changes - Min %Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	25.15	1.78	15.33	8.45	10.05	1.99	1.62	6.23
With Upskilling	0.00	0.13	2.85	2.01	1.57	0.55	0.10	1.97
With Investment	22.07	3.05	13.74	7.10	8.28	3.07	1.52	4.63
Basic Model	0.00	0.73	3.31	2.00	1.72	0.48	0.10	1.96

Notes: This table reports the spatial dispersion of % changes in local economic outcomes in China induced by a uniform infrastructure shock under different model specifications. The upper panel shows the standard deviation of perfecture-level impacts and the lower panel shows the range of perfecture-level impacts. From left to right economic outcomes are total capital stock, total skill ratio, real unskilled wage, real skilled wage, average skill premium, total GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is $100\% \times \Delta$ as defined in footnote 11.

some realistic events that happened in early 2000 China and quantify their impacts based on our models.

4.2 Impacts of Trade Liberalization

In this part, we discuss the aggregate and the spatial impacts of trade in the context of China's WTO accession. Specifically, we compare the baseline economy with observed trade liberalization after the WTO accession to a counterfactual economy where the trade costs between China and the ROW were kept at the pre-WTO levels in the year 2000. We run this counterfactual experiment separately for each model setup.

Aggregate Impacts Since China is less capital-abundant relative to the ROW in the early 2000s and skilled manufacturing uses less capital than unskilled one, China has a comparative advantage in the skilled manufacturing sector. Thus we expect that China's trade liberalization benefits skilled workers more than unskilled workers because of the Stolper-Samuelson theorem. In the basic model, as shown in row 4 in table 7, trade liberalization brings positive but unequal welfare gains across skill types. The unskilled real wage increases by 1.29 percent and the skilled one increases by 1.69, increasing the skill premium by 0.41 percent. This result is consistent with the prediction from the Stolper-Samuelson theorem.

Table 7: Aggregate Impacts of Trade Liberalization

	%Changes in Steady State						
	Capital	S.R.	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	1.33	0.11	1.61	2.72	1.17	1.33	1.90
With Upskilling	-	0.15	0.99	1.94	0.95	1.83	1.25
With Investment	2.41	-	2.09	3.17	1.24	2.27	2.04
Basic Model	-	-	1.29	1.69	0.41	1.73	1.28

Notes: This table reports the steady state % changes of aggregate economic outcomes in China induced by trade liberalization under different model specifications. From left to right economic outcomes are total capital stock, total skill ratio, real unskilled wage, real skilled wage, average skill premium, total GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is 100% $\times\, \Delta$ as defined in footnote 11.

Including skill acquisition allows China to strengthen its comparative advantage in the skilled sector when it trades with the ROW. As shown in the second row in table 7, trade liberalization encourages skill upgrading in China, increasing skill ratio by 0.15 percent and skill premium by 0.95 percent. These results are consistent with the Stolper-Samuelson theorem. Compared to the basic model simulation, allowing skill adjustment enlarges the between-skill unequal gains from trade but barely affects the total welfare gains.

Next, modelling capital accumulation but no upskilling yields much higher welfare gains from trade liberalization. The unskilled real wage increases by 2.09 percent, a 1.5 times amplification from 1.29 in basic model simulation. The skilled real wage also increases by 3.17 percent from trade liberalization. More interestingly, trade liberalization increases skill premium by 1.24 percent under capital accumulation, a much stronger positive impact compared to the basic model simulation. In fact, adding capital investment has two competing implications on trade-induced impacts on skill premium. On the one hand, with endogenous capital, trade liberalization reduces the cost of sourcing investment goods and thus encourages capital investment. Due to capital-skill complementarity, the accelerated rate of capital accumulation substantially improves skilled workers' welfare gain. On the other hand, the trade-induced capital gain to some extent weakens China's comparative advantage in the skilled sector, as the skilled sector is less capital intensive than the unskilled one, thus attenuating the positive impact on skill premium. In the end, the former is much stronger than the latter, resulting in a magnified positive impact on skill premium.

Finally, including both capital investment and upskilling generates a negative interaction effect. In the case of trade between China and the ROW, capital investment weakens China's comparative advantage in the skilled sector, but skill upgrading strengthens it. When both mechanisms are modeled, capital and skill adjustment weaken each other's response to the trade liberalization. As shown in the first two columns in table 7, capital accumulation depresses the trade-induced skill upgrading, and the skill upgrading also diminishes the positive impact on capital stock. As a result, trade liberalization induces moderate increases in welfare gains and skill premium.

Spatial Impacts Although trade liberalization creates aggregate welfare improvement in China, the local gains are highly unequal across regions, and the degree of inequality especially depends on our model mechanism. Table 8 reports cross-city inequality measures of economic gains from trade liberalization. The main message here is that endogenous factor accumulation significantly widens the benefit gaps across cities. Coastal cities in China, due to their geographic closeness to the ROW, reap larger gains from trade liberalization than inland cities. In appendix C.2, we show the plots of the trade impacts and their clear spatial patterns. Incorporating physical and human capital accumulation further reinforces their geographic advantage, as trade liberalization allows them to source cheap varieties from the ROW and accumulate more capital stocks.

In particular, as shown in the table 8, the top winner from trade liberalization accumulates 23 more percent of capital than the last winner under the full model simulation. Consequently, the trade shock induces a larger migrant (also skilled migrant) inflow towards coastal cities and welfare gains in the full model simulation than in the basic model simulation. For example, compare Dalian, a coastal city and one of the top winners, and

Table 8: Spatial Dispersion of Trade Liberalization's Impacts

	Standard Deviation of Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	4.61	0.17	2.58	1.17	1.86	0.68	0.22	0.44
With Upskilling	0.00	0.04	0.66	0.38	0.43	0.07	0.04	0.17
With Investment	4.55	0.59	2.19	1.30	1.69	0.40	0.27	0.55
Basic Model	0.00	0.04	0.65	0.40	0.42	0.04	0.05	0.18

	Max %Changes - Min %Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	23.01	1.53	13.80	5.54	8.63	3.81	1.17	2.21
With Upskilling	0.00	0.26	3.31	2.02	2.10	0.41	0.27	0.90
With Investment	24.64	3.31	12.83	6.92	8.63	2.39	1.39	3.15
Basic Model	0.00	0.39	3.33	2.05	2.04	0.38	0.35	0.90

Notes: This table reports the % changes of aggregate economic outcomes in China induced by liberalization under different model specifications. From left to right economic outcomes are capital stock, skill ratio, real unskilled wage, real skilled wage, skill premium, GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is $100\% \times \Delta$ as defined in footnote 11.

Chengdu, an inland large city and one of the last winners of trade liberalization. In the basic model simulation, Dalian attracts 1.5 percent more migrants and consume 1.7 percent more due to the trade liberalization, while Chengdu loses 1.1 percent of its workforce and consumes only 1.0 percent more. These trade-induced unequal consumption gains and inland-to-coast migration pattern already suggest that coastal cities are the main beneficiaries of trade liberalization. But factor accumulations significantly widen the unequal gains from trade. In the full model simulation, the trade-induced workforce change is a striking 8.8 percent increase in Dalian, and a 5.0 percent drop in Chengdu; local consumption gains in Dalian become 3.1 percent and that in Chengdu is still at 1.1 percent. These widen unequal gains between the coast and the inland are not only in population and consumption, but also in other outcomes such as capital and skilled workers.

4.3 Impacts of Infrastructure Investment

During the early 2000s, China experienced large-scale domestic infrastructure improvement, resulting in lower domestic trade and migration costs. How does such infrastructure improvement affect welfare and the skill premium in China? Do the impacts depend on production factor adjusments? In this section, we try to answer those questions. We consider a counterfactual economy in which China's domestic trade costs and migration costs are fixed at the initial level in 2000.

Aggregate Impacts Table 9 summarizes the aggregate impacts of such infrastructure improvement. The results are generally consistent with previous analyses of uniform infrastructure improvement. Using the basic model, we find that the infrastructure improvement increases unskilled real wages by 1.67 percent and skilled wages by 1.73 percent. Allowing endogenous capital accumulation alone considerably enlarges welfare impacts: the welfare gains measured by unskilled and skilled wage changes increase to 3.04 and 4.36 percent respectively. Allowing upskilling in addition to capital accumulation further augments welfare impacts: the infrastructure improvement increases unskilled and skilled wages now by 3.65 and 4.82 percent respectively. However, adding upskilling alone only generates limited welfare changes from the basic model simulation. Such different roles of upskilling with or without capital accumulation are evidence of the interaction effect between capital and skill accumulation.

Table 9: Aggregate Impacts of Infrastructure Improvement

	%Changes in Steady State						
	Capital	S.R.	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	2.47	0.14	3.65	4.82	1.16	2.38	4.65
With Upskilling	-	0.14	1.92	1.55	-0.38	1.47	3.20
With Investment	3.46	-	3.04	4.36	1.31	3.28	3.57
Basic Model	-	-	1.67	1.73	0.23	1.56	2.50

Notes: This table reports the steady state % changes of aggregate economic outcomes induced by China's infrastructure improvement in the early 2000s. From left to right economic outcomes are total capital stock, total skill ratio, real unskilled wage, real skilled wage, average skill premium, total GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is $100\% \times \Delta$ as defined in footnote 11.

Table 9 also shows that the infrastructure improvement impacts skill premium distinctly given different model assumptions. The basic model simulation suggests that the infrastructure expansion is slightly skill-biased or skill-neutral, as it increases skill premium by only 0.23 percent. The expansion becomes strongly skill-biased given endogenous capital accumulation, with the infrastructure improvement raising skill premium by 1.31 percent. The reason is that the infrastructure improvement induces faster capital accumulation, increasing the return on skill. It then becomes unskill-biased under the simulation with skill acquisition but no capital investment. As infrastructure improvement increases the total skill ratio, the resulting larger supply of skills reduces the skill premium. A more interesting result here is that infrastructure improvement encourages skill upgrading. The main reason is that the infrastructure expansion reduces geographic travel costs for both skill-type workers in the same absolute magnitude, but lowers skilled workers' migration friction relatively more since skilled workers face lower migration policy barriers than unskilled workers. In other words, the infrastructure improvement exaggerates unskilled workers' disadvantage in the discriminatory migration policy. Therefore, unskilled workers are more willing to transfer to skilled workers given the lower travel costs, resulting in a positive change in the total skill ratio. Finally, the infrastructure expansion becomes skill-biased again in the full model simulation as the positive effect from capital growth dominates the negative effect from a larger skill supply. But the skill supply adjustment still attenuates the positive impact on skill premium from 1.31 to 1.16 percent.

Spatial Impacts Lastly, we examine the spatial impact of China's infrastructure expansion. Plots in appendix C.2 show that the western or undeveloped cities gain the most from the expansion between 2000-2015, as those cities previously had low market access but experienced relatively larger improvement in connectiveness than other cities. Next, we compare across model setups the spatial dispersions of impacts of infrastructure improvement in table 10. Consistent with our main conclusion, capital accumulation generates highly heterogeneous spatial impacts on welfare but skill acquisition alone contributes little. The interaction between capital and skill accumulation further enlarges the spatial dispersion.

Table 10: Spatial Dispersion of Infrastructure Improvement's Impacts

	Standard Deviation of Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	22.38	1.21	13.59	4.02	5.05	2.06	1.27	2.30
With Upskilling	0.00	1.06	3.87	1.50	1.56	2.29	0.29	1.34
With Investment	18.67	2.48	10.58	3.76	3.92	1.78	1.21	2.11
Basic Model	0.00	1.23	4.04	1.18	1.51	1.76	0.27	0.88

	Max %Changes - Min %Changes in Steady State							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	155.81	8.29	84.45	26.42	34.97	13.01	7.18	17.15
With Upskilling	0.00	7.80	30.23	9.89	12.57	18.40	2.46	11.28
With Investment	126.12	14.38	78.19	24.44	26.93	11.00	6.96	15.73
Basic Model	0.00	9.99	29.59	8.64	12.48	17.57	2.27	7.52

Notes: This table reports the spatial dispersion of % changes in local economic outcomes induced by China's infrastructure improvement in the early 2000s. The upper panel shows the standard deviation of prefecture-level impacts and the lower panel shows the range of prefecture-level impacts. From left to right economic outcomes are capital stock, skill ratio, real unskilled wage, real skilled wage, skill premium, GDP, and consumption equivalence. The "Full Model" is the baseline model; 'With Upskilling' refers to the model with skill acquisition but constant local capital stock; 'With Investment' refers to the model with endogenous capital accumulation but skill upgrading cost is infinite: $\kappa_s = \infty$; 'Basic Model' refers to the one with constant capital stock and $\kappa_s = \infty$. The real unskilled (skilled) wage is a weighted average with the weight of each city being the share of local unskilled (skilled) labor. GDP is the China's total output $\sum_{i=1}^{N-1} X_i^*$. C.E. is 100% $\times\ \Delta$ as defined in footnote 11.

4.4 Place-targeting Policies

Given the previous results, modelling endogenous capital and skill formation also has policy implications. Consider a place-targeting policy that the government provides exogenous wage subsidies to poor cities in order to improve the workers' welfare there. The required subsidy level to achieve the same target goal could differ, depending on whether factor accumulation is modelled. In this section, we consider a simple place-targeting policy, in which an absent government exogenously subsidizes worker wages in a targeted poor location to improve the location's overall welfare by 10 percent. We assume the government's wage subsidy simply as a multiplier on local equilibrium wages for all worker types. The policy can roughly represent the "targeted poverty alleviation" or other anti-poverty campaigns implemented in China.

The required subsidy level indeed crucially depends on factor adjustments. As shown in figure 5, endogenous capital reduces the required amount of wage subsidy compared to simulations without capital investment, as capital accumulation induces a similar amplification effect on welfare as before. On the other hand, allowing skill upgrading significantly increases the required subsidy level. Since unskilled workers can become skilled ones, and skilled workers tend to migrate from the poor cities to large cities where capital stocks are larger and returns on skill are higher, poor cities lose more population in this case than in the simulation without upskilling. Therefore, the government needs additional wage subsidies to prevent workforce loss and improve welfare. Conversely, we can expect that the required subsidy will be less for big cities when unskilled workers can acquire skills.

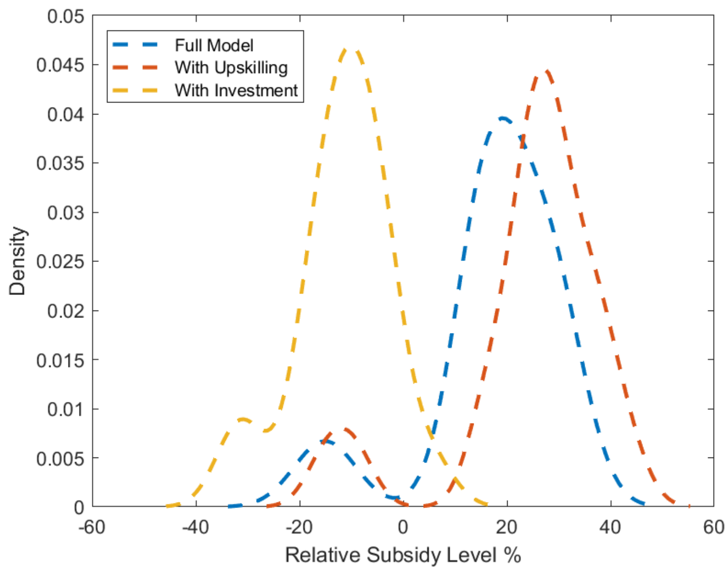


Figure 5: Distribution of Required Wage Subsidy

Notes: This figure shows the kernel density of required relative wage subsidy to improve local welfare by 10 percent for the 20 poor cities in steady state. The required relative wage subsidy in each model setup is calculated as the required wage subsidy in that model setup minus that in the basic model simulation. The poor city is defined as those in the bottom 5 percentile in the 2000 GDP per capita distribution. Local welfare is measured by the average steady-state utility across worker types.

5 Conclusion

In this study, we develop a dynamic spatial framework to understand how economic shocks or events affect spatial factor accumulation and how the aggregate and spatial impacts depend on factor accumulation. The model features capital-skill complementarity, capital accumulation, and endogenous skill acquisition. Different skill types of workers are differentiated by their spatial mobilities and their roles in the production function. We then apply our framework to China's economy.

We find that physical and human capital accumulation interact with each other in response to economic shocks, and the impacts of shocks crucially depend on the presence of factor accumulation. Ignoring both physical and human capital accumulation substantially underestimates the welfare impacts of shocks or events such as trade liberalization, infrastructure improvement, and a productivity jump. The interaction adjustments between capital and skill formation can be context-specific: it is positive in the productivity shock but negative in trade liberalization. Our quantitative analysis also shows that the skill premium responds differently, depending on whether capital or human capital can adjust.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

_ **and** _ , "The welfare effects of transportation infrastructure improvements," *The Review of Economic Studies*, 2022, *89* (6), 2911–2957.

Armington, Paul S., "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 1969, *16* (1), 159–178.

Artuc, Erhan, Irene Brambilla, and Guido Porto, "Patterns of labour market adjustment to trade shocks with imperfect capital mobility," *The Economic Journal*, 2022, *132* (646), 2048–2074.

Autor, David, David Dorn, and Gordon Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *The American Economic Review*, 2013, *103* (6), 2121–2168.

_, _, and _, "On the Persistence of the China Shock," *Brookings Papers on Economic Activity*, 2021, pp. 381–447.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the road: Access to transportation infrastructure and economic growth in China," *Journal of Development Economics*, 2020, *145*, 102442.

Blanchard, Emily and Gerald Willmann, "Trade, education, and the shrinking middle class," *Journal of International Economics*, 2016, *99*, 263–278.

Borsook, I., "Earnings, ability and international trade," *Journal of International Economics*, 1987, *22* (3), 281–295.

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

Cai, Sheng, Lorenzo Caliendo, Fernando Parro, and Wei Xiang, "Mechanics of spatial growth," Technical Report, National Bureau of Economic Research 2022.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 2019, *87* (3), 741–835.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade theory with numbers: Quantifying the consequences of globalization," in "Handbook of International Economics," Vol. 4, Elsevier, 2014, pp. 197–261.

Donaldson, Dave and Richard Hornbeck, "Railroads and American economic growth: A "market access" approach," *The Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Duffy, John, Chris Papageorgiou, and Fidel Perez-Sebastian, "Capital-skill complementarity? Evidence from a panel of countries," *Review of Economics and Statistics*, 2004, *86* (1), 327–344.

Faber, Benjamin, "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Falvey, Rod, David Greenaway, and Joana Silva, "Trade liberalisation and human capital adjustment," *Journal of International Economics*, 2010, *81* (2), 230–239.

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

Feenstra, Robert C and Gordon H Hanson, "Globalization, Outsourcing, and Wage Inequality," *The American Economic Review*, 1996, *86* (2), 240–245.

Findlay, Ronald and Henryk Kierzkowski, "International Trade and Human Capital: A Simple General Equilibrium Model," *Journal of Political Economy*, 1983, *91* (6), 957–978.

Goldberg, Pinelopi Koujianou and Nina Pavcnik, "Distributional effects of globalization in developing countries," *Journal of Economic Literature*, 2007, *45* (1), 39–82.

Head, Keith and John Ries, "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *American Economic Review*, September 2001, *91* (4), 858–876.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding, "Inequality and unemployment in a global economy," *Econometrica*, 2010, *78* (4), 1239–1283.

Hu, Naiyuan and Lin Ma, "Educational Migration in China."

Kleinman, Benny, Ernest Liu, and Stephen J Redding, "Dynamic spatial general equilibrium," *Econometrica*, 2023, *91* (2), 385–424.

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.

Ma, Lin and Yang Tang, "Geography, trade, and internal migration in China," *Journal of Urban Economics*, 2020, *115*, 103181.

— **and** — , "The Distributional Impacts of Transportation Networks in China," *Journal of International Economics*, 2024.

Parro, Fernando, "Capital-skill complementarity and the skill premium in a quantitative model of trade," *American Economic Journal: Macroeconomics*, 2013, *5* (2), 72–117.

Ravikumar, B, Ana Maria Santacreu, and Michael Sposi, "Capital accumulation and dynamic gains from trade," *Journal of International Economics*, 2019, *119*, 93–110.

Redding, Stephen J and Matthew A Turner, "Transportation costs and the spatial organization of economic activity," *Handbook of regional and urban economics*, 2015, *5*, 1339–1398.

Tombe, Trevor and Xiaodong Zhu, "Trade, migration, and productivity: A quantitative analysis of China," *American Economic Review*, 2019, *109* (5), 1843–72.

Young, Alwyn, "Gold into base metals: Productivity growth in the People's Republic of China during the reform period," *Journal of Political Economy*, 2003, *111* (6), 1220–1261.

Zhang, Jun, Guiying Wu, and Jipeng Zhang, "The estimation of China's provincial capital stock: 1952–2000," *Economic Research Journal (Chinese)*, 2004, *10* (1), 35–44.

A Details of the Model

A.1 Price and Trade

Denote p_{it} as the price index at location i. By the nested preference structure and given the price of sector j's goods supplied by exporter n to i, $p^j_{in,t}$, the price index at i is

$$p_{it} = \prod_{j=1}^{J} \left(p^j_{it}\right)^{\gamma^j}, \tag{28}$$

where the sector-level price p^j_{it} is given by

$$p^j_{it} = \left[\sum_{n=1}^{N} \left(p^j_{in,t}\right)^{-\theta}\right]^{-\frac{1}{\theta}}. \tag{29}$$

The share of importer i's expenditure within industry j on goods supplied by exporter n is

$$\pi^j_{in,t} = \frac{(p^j_{in,t})^{-\theta}}{\sum_{m=1}^{N}(p^j_{im,t})^{-\theta}}. \tag{30}$$

A.2 Worker's Problem

A.2.1 Skilled Worker's Problem

The recursive problem of a skilled worker is given by

$$v^h_{it} = \ln\left(b_{it}\frac{w^h_{it}}{p_{it}}\right) + \max_{\{n\in N\}}\left\{\xi\beta\mathbb{E}\left[v^h_{nt+1}\right] - \kappa^h_{ni,t} + \rho\varepsilon_{nt}\right\},$$

where ε_{nt} is i.i.d. across individual and time and follows a Gumbel distribution with distribution function $F(\varepsilon) = \exp(-\exp(-\varepsilon - \bar{\gamma}))$, where $\bar{\gamma}$ is the Euler's constant. Let's denote $V^h_{it} \equiv \mathbb{E}\left[v^h_{nt}\right]$ and $\Phi^h_{ni} \equiv \xi\beta\mathbb{E}\left[v^h_{nt}\right] - \kappa^h_{ni,t}$. We aim to solve for the expected continuation value $\mathbb{E}\left[\max_n\{\Phi^h_{ni} + \rho\varepsilon_{nt}\}\right]$.

Note that the distribution function of $\max_n\{\Phi^h_{ni} + \rho\varepsilon_{nt}\}$ is

$$\begin{aligned}
\Pr\left(\max_n\{\Phi^h_{ni} + \rho\varepsilon_{nt}\} < x\right) &= \prod_n \Pr\left(\Phi^h_{ni} + +\rho\varepsilon_{nt} < x\right) \\
&= \prod_n \exp\left\{-\exp\left(-\frac{x - \Phi^h_{ni} + \rho\bar{\gamma}}{\rho}\right)\right\} \\
&= \exp\left\{-\sum_n \exp\left(-\frac{x - \Phi^h_{ni} + \rho\bar{\gamma}}{\rho}\right)\right\},
\end{aligned}$$

where

$$\begin{aligned}
\sum_n \exp\left(-\frac{x - \Phi^h_{ni} + \rho\bar{\gamma}}{\rho}\right) &= \exp\left[\ln\left(e^{\left(-\frac{x+\rho\bar{\gamma}}{\rho}\right)}\sum_n e^{\left(\frac{\Phi^h_{ni}}{\rho}\right)}\right)\right] \\
&= \exp\left[-\frac{x+\rho\bar{\gamma}}{\rho} + \ln\sum_n \exp\left(\frac{\Phi^h_{ni}}{\rho}\right)\right] \\
&= \exp\left[-\frac{x + \rho\bar{\gamma} - \rho\ln\sum_n \exp\left(\frac{\Phi^h_{ni}}{\rho}\right)}{\rho}\right].
\end{aligned}$$

Therefore, $\max_n\{\Phi^h_{ni} + \rho\varepsilon_{nt}\}$ also follows a Gumbel distribution with the location parameter $-\rho\bar{\gamma} + \rho\ln\sum_n \exp\left(\frac{\Phi^h_{ni}}{\rho}\right)$ and scale parameter ρ. Using the property that a Gumbel distribution with location parameter μ and scale parameter λ has the expectation as $\mu + \bar{\gamma}\lambda$, we have

$$\begin{aligned}
\mathbb{E}\left[\max_n\{\Phi^h_{ni} + \rho\varepsilon_{nt}\}\right] =& \rho\ln\sum_n \exp\left(\frac{\Phi^h_{ni}}{\rho}\right) \\
=& \rho\ln\sum_n \exp\left[\left(\xi\beta\mathbb{E}\left[v^h_{nt}\right] - \kappa^h_{ni,t}\right)/\rho\right]
\end{aligned}$$

and therefore

$$V^h_{it} = \ln\left(\frac{b_{it}w^h_{it}}{p_{it}}\right) + \rho\ln\left\{\sum_{n=1}^{N} \exp\left[(\xi\beta V^h_{n,t+1} - \kappa^h_{ni,t})/\rho\right]\right\}.$$

A.3 Unskilled Worker's Problem

The derivation of the solution to the unskilled worker's problem is similar. First, following the exact procedure as before, the expected continuation value of being skill type d worker is

$$\widetilde{V}_{it}^d \equiv \mathbb{E}\left[\widetilde{v}_{it}^d\right] = \rho \ln \sum_{n=1}^{N} \exp\left[(\xi\beta V_{n,t+1}^d - \kappa_{ni,t}^l)/\rho\right], \ d \in \{l, s\}.$$

Assume the idiosyncratic education shock ζ_{it}^d is i.i.d. across individual and time and also follows the Gumbel distribution $F(\varepsilon) = \exp(-\exp(-\varepsilon - \bar{\gamma}))$, similarly using the result of Gumbel distribution we have

$$V_{it}^l = \ln\left(\frac{b_{it} w_{it}^l}{p_{it}}\right) + \psi \ln\left[\exp\left(\frac{\widetilde{V}_{it}^l}{\psi}\right) + \exp\left(\frac{\widetilde{V}_{it}^s - \omega}{\psi}\right)\right].$$

A.4 Firm's Problem

In this part, we drop the time notation for brevity. The problem of a producer in sector j at location i is given by:

$$\min_{l,s,k} \ w_i^l l_i^j + w_i^h h_i^j + r_i k_i^j,$$

subject to:

$$z_i\left[(\mu^j)^{\frac{1}{\sigma}}(L_i^{ij})^{\frac{\sigma-1}{\sigma}} + (1-\mu^j)^{\frac{1}{\sigma}}(L_i^{ej})^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \geq q_i^j \tag{31}$$

$$L_i^{ej} = \left[(\lambda^j)^{\frac{1}{\eta}}(k_i^j)^{\frac{\eta-1}{\eta}} + (1-\lambda^j)^{\frac{1}{\eta}}(L_i^{hj})^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}. \tag{32}$$

First order conditions for s_i^j and k_i^j yield:

$$k_i^j = \frac{\lambda_j}{1-\lambda_j}\left(\frac{w_i^s}{r_i}\right)^{\eta} L_i^{hj}. \tag{33}$$

Using this expression to replace k_i^j in equation (32) and define the price w_i^{ej} for composite input L_i^{ej} such that $w_i^{ej} L_i^{ej} = r_i k_i^j + w_i^h L_i^{hj}$, we obtain:

$$w_i^{ej} = \left[\lambda^j (r_{it})^{1-\eta} + (1-\lambda^j)(w_{it}^h)^{1-\eta}\right]^{\frac{1}{1-\eta}}. \tag{34}$$

Similarly, first order conditions for L_i^{lj} and L_i^{hj} give:

$$L_i^{lj} = \frac{\mu_j}{1-\mu_j}\left(\frac{w_i^{ej}}{w_i^l}\right)^{\sigma} L_i^{hj}. \tag{35}$$

Using equation (35) and define the unit cost c_i^j for the variety q_i^j such that $c_i^j q_i^j = w_i^l L_i^{lj} + w_i^{ej} L_i^{ej}$, we obtain:

$$c_i^j = \frac{1}{z_i}\left[\mu^j (w_i^l)^{1-\sigma} + (1-\mu^j)(w_i^e)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{36}$$

A.5 Numerical Algorithm for Solving Steady State

We first write down the corresponding equilibrium conditions in the steady state. The value function (7)(8)(6) become:

$$V_i^{l*} = \ln b_i + \ln \frac{w_i^{l*}}{p_i^*} + \psi \ln\left[\exp\left(\frac{\tilde{V}_i^{l*}}{\psi}\right) + \exp\left(\frac{\tilde{V}_i^{s*}-\omega}{\psi}\right)\right] \tag{37}$$

with $\tilde{V}_i^{d*} = \rho \ln \sum_{g=1}^{N} \exp\left[(\xi\beta V_g^{d*} - \kappa_{gi}^{d*})/\rho\right],\ d = l, s.$

$$V_i^{s*} = \ln b_i + \ln \frac{\iota w_{i*}^l}{p_i^*} + \rho \ln \sum_{g=1}^{N} \exp\left[(\xi\beta V_g^{h*} - \kappa_{gi}^h)/\rho\right], \tag{38}$$

$$V_i^{h*} = \ln b_i + \ln \frac{w_i^{h*}}{p_i^*} + \rho \ln \sum_{g=1}^{N} \exp\left[(\xi\beta V_g^{h*} - \kappa_{gi}^h)/\rho\right], \tag{39}$$

the skill upgrading matrix (9) and migration matrix (10) become

$$D_i^{ls*} = \frac{\exp\left[(V_i^{s*}-\omega)/\psi\right]}{\exp\left[(V_i^{l*})/\psi\right] + \exp\left[(V_i^{s*}-\omega)/\psi\right]}, \tag{40}$$

$$D_{in}^{d*} = \frac{\exp\left[(\xi\beta V_i^{d*} - \kappa_{in}^{d*})/\rho\right]}{\sum_{g=1}^{N} \exp\left[(\xi\beta V_g^{d*} - \kappa_{gn}^{d*})/\rho\right]},\ d = l, s, h \tag{41}$$

and the labor supply (11) - (13) become

$$L_i^{l*} = \xi \sum_{n=1}^{N} D_{in}^{l*} \left(L_n^{l*} - D_n^{ls*} L_n^{l*} \right) + \left(L_i^{l*} + L_i^{s*} + L_i^{h*} \right) (1 - \xi), \tag{42}$$

$$L_i^{s*} = \xi \sum_{n=1}^{N} D_{in}^{s*} D_n^{ls*} L_n^{l*}, \tag{43}$$

$$L_i^{h*} = \xi \left(\sum_{n=1}^{N} D_{in}^{h*} \left(L_n^{h*} + L_n^{s*} \right) \right). \tag{44}$$

The market clearing conditions (20)(21)(22)(23) become:

$$X_i^{j*} = \sum_{n=1}^{N} S_{ni}^{j*} \left[\gamma_j \sum_{m=1}^{J} X_n^{m*} \right], \tag{45}$$

$$w_i^{l*} = \frac{\sum_{j=1}^{J} \phi_i^{lj*} X_i^{j*}}{\tilde{L}_i^{l*}} \tag{46}$$

$$w_i^{h*} = \frac{\sum_{j=1}^{J} \phi_i^{hj*} X_i^{j*}}{L_i^{h*}} \tag{47}$$

$$r_i^* = \frac{\sum_{j=1}^{J} \phi_i^{kj*} X_i^{j*}}{k_i^*}, \tag{48}$$

with steady-state factor income shares given by:

$$\phi_i^{lj*} = \left[1 + \frac{1-\mu^j}{\mu^j} \left(\frac{w_i^{l*}}{w_i^{e*}} \right)^{\sigma-1} \right]^{-1} \tag{49}$$

$$\phi_i^{hj*} = \left[1 + \frac{\mu^j}{1-\mu^j} \left(\frac{w_i^{e*}}{w_i^{l*}} \right)^{\sigma-1} \right]^{-1} \left[1 + \frac{\lambda^j}{1-\lambda^j} \left(\frac{w_i^{h*}}{r_i^*} \right)^{\eta-1} \right]^{-1} \tag{50}$$

$$\phi_i^{kj*} = \left[1 + \frac{\mu^j}{1-\mu^j} \left(\frac{w_i^{e*}}{w_i^{l*}} \right)^{\sigma-1} \right]^{-1} \left[1 + \frac{1-\lambda^j}{\lambda^j} \left(\frac{r_i^*}{w_i^{h*}} \right)^{\eta-1} \right]^{-1}. \tag{51}$$

The trade share adopts the following expression:

$$\pi_{ni}^{j*} = \frac{(p_{ni}^{j*})^{-\theta}}{\sum_{g=1}^{N} (p_{nm}^{j*})^{-\theta}}, \tag{52}$$

where

$$p_{ni}^{j*} = \frac{\tau_{ni}}{z_i^*}\left[\mu^j(w_i^{l*})^{1-\sigma} + (1-\mu^j)\left[\lambda^j(r_i^*)^{1-\eta} + (1-\lambda^j)(w_i^{h*})^{1-\eta}\right]^{\frac{1-\sigma}{1-\eta}}\right]^{\frac{1}{1-\sigma}}.$$

The capital accumulation condition becomes:

$$k_i^* = \xi\beta(1-\delta+\frac{r_i^*}{p_i^*})k_i^*, \text{ with } p_i^* = \prod_{j=1}^{J}\left[\sum_{n=1}^{N}(p_{in}^{j*})^{-\theta}\right]^{-\frac{\gamma_j}{\theta}}. \tag{53}$$

Given these conditions, the algorithm is as follows.

(1) Start with an initial guess of value functions $\{V_i^{l(0)}, V_i^{s(0)}, V_i^{h(0)}\}_{i=1}^N$ and factor allocations $\{L_i^{l(0)}, L_i^{s(0)}, L_i^{h(0)}, k_i^{(0)}\}_{i=1}^N$.

(2) Given $\{V_i^{l(0)}, V_i^{s(0)}, V_i^{h(0)}\}_{i=1}^N$, compute skill upgrading probability $\{D_i^{ls}\}_{i=1}^N$ and migration shares $\{D_{ig}^{ed}\}_{i=1}^N$ by (40) and (41), and then solve new labor allocations by (42), (43) and (44) to obtain $\{L_i^{l(1)}, L_i^{s(1)}, L_i^{h(1)}\}_{i=1}^N$. The total local unskilled labor supply is obtained as $\tilde{L}_i^{l(1)} = L_i^{l(1)}(1-D_i^{ls}) + \iota L_i^{s(1)}$

(3) Given $\{\tilde{L}_i^{l(1)}, L_i^{h(1)}, k_i^{(0)}\}_{i=1}^N$, solve factor prices $\{w_i^l, w_i^h, r_i\}_{i=1}^N$ from markets clearing conditions as follows:

- (a) set an initial guess of factor prices $\{w_i^l, w_i^h, r_i\}_{i=1}^N$,
- (b) compute factor incomes shares $\{\phi_i^{lj}, \phi_i^{hj}, \phi_i^{kj}\}_{i=1}^N$ from (49), (50), (51),
- (c) compute prices $\{p_{ni}\}_{n=1,i=1}^{N,N}$ and trade shares $\{\pi_{ni}\}_{i=1,n=1}^{N,N}$ from (52),
- (d) solve total output X_i^{j*} by (45)
- (e) obtain new factor prices $\{w_i^l, w_i^h, r_i\}_{i=1}^N$ by (46), (47), (48),
- (f) iterate until factor prices converge.

(4) Use $\{w_i^l, w_i^h, r_i\}_{i=1}^N$ to compute price index $\{p_i\}_{i=1}^N$ and solve new captial $\{k_i^{(1)}\}_{i=1}^N$ by (53).

(5) Given $\{V_i^{l(0)}, V_i^{s(0)}, V_i^{h(0)}, w_i^{l(1)}, w_i^{h(1)}, p_i\}_{i=1}^N$, solve new value functions $\{V_i^{l(1)}, V_i^{s(1)}, V_i^{h(1)}\}_{i=1}^N$ by (37) - (39).

(6) Update $\{V_i^{l(0)}, V_i^{s(0)}, V_i^{h(0)}, L_i^{l(0)}, L_i^{s(0)}, L_i^{h(0)}, k_i^{(0)}\}_{i=1}^N$ from $\{v_i^{l(1)}, v_i^{s(1)}, L_i^{l(1)}, L_i^{s(1)}, L_i^{h(1)}, k_i^{(1)}\}_{i=1}^N$.

(7) Repeat steps (2)-(6) until value functions $\{V_i^l, V_i^s, V_i^h\}_{i=1}^N$ converge.

A.6 Numerical Algorithm for Solving Path Equilibrium

Given the initial allocations of labor and capital, $\{L^l_{i0}, L^s_{i0}, L^h_{i0}, k_{i0}\}$, we solve a transition path of length T towards a steady state using a shooting algorithm as follows.

(1) Start with an initial guess of value functions and capital stocks $\{V^{l(0)}_{it}, V^{s(0)}_{it}, V^{h(0)}_{it}\}^{N,T}_{i=1,t=1}$, where $V^{l(0)}_{iT}$, $V^{s(0)}_{iT}$ and $V^{h(0)}_{iT}$ are approximated by steady-state level of value functions.

(2) Given $\{V^{l(0)}_{it}, V^{s(0)}_{it}, V^{h(0)}_{it}\}^{N,T}_{i=1,t=1}$, solve upskilling probability $\{D^{ls}_{it}\}^{N,T}_{i=1,t=1}$ and migration shares $\{D_{in,t}\}^{N,T}_{i=1,t=1}$ from (9) and (10).

(3) Use $\{D^{ls}_{it}, D_{in,t}\}^{N,T}_{i=1,t=1}$ and $\{L^l_{i0}, L^s_{i0}, L^h_{i0}\}$ to solve $\{L^l_{it}, L^s_{it}, L^h_{it}\}^{N,T}_{i=1,t=1}$ by (11) - (13). Then compute total local unskilled labor supply as $\tilde{L}^l_{it} = L^l_{it}(1 - D^{ls}_{it}) + \iota L^s_{it}$.

(4) For each time period t, use current state variables $\{\tilde{L}^l_{it}, L^h_{it}, k^{(0)}_{it}\}^N_{i=1}$ to solve factor prices $\{w^l_{it}, w^h_{it}, r_{it}\}^N_{i=1}$:

- (a) set an initial guess of factor prices $\{w^l_{it}, w^h_{it}, r_{it}\}^N_{i=1}$,
- (b) compute factor incomes shares $\{\phi^{lj}_{it}, \phi^{hj}_{it}, \phi^{kj}_{it}\}^N_{i=1}$ from (16), (17), (18),
- (c) compute trade shares $\{\pi^j_{ni,t}\}^N_{i=1}$ from (19) and (30),
- (d) solve total output by (23),
- (e) solve new factor prices by (20), (21), (22),
- (f) iterate until factor prices converge.

(5) Use solved factor prices $\{w^l_{it}, w^h_{it}, r_{it}\}^{N,T}_{i=1,t=1}$ to compute $\{p_{ni,t}\}^{N,T}_{i=1,t=1}$ by (19). Then obtain price index $\{p_{nt}\}^{N,T}_{n=1,t=1}$ by (29) and solve new captial allocations sequence $\{k^{(1)}_i\}^{N,T}_{i=1,t=1}$ from k_{i0} and (14).

(6) Set $\{V^{l(1)}_{iT}, V^{s(1)}_{iT}, V^{h(1)}_{iT}\} = \{V^{l(0)}_{iT}, V^{s(0)}_{iT}, V^{h(0)}_{iT}\}$. Given $\{w^l_{it}, w^h_{it}, p_i\}^{N,T}_{i=1,t=1}$ and $\{V^{l(1)}_{iT}, V^{s(1)}_{iT}, V^{h(1)}_{iT}\}$, solve new value functions $\{V^{l(1)}_{it}, V^{s(1)}_{it}, V^{h(1)}_{it}\}^{N,T-1}_{i=1,t=1}$ backward by (7), (8) and (6).

(7) Update $\{V^{l(0)}_{it}, V^{s(0)}_{it}, V^{h(0)}_{it}\}^{N,T}_{i=1,t=1}$ from $\{V^{l(1)}_{it}, V^{s(1)}_{it}, V^{h(1)}_{it}\}^{N,T}_{i=1,t=1}$.

(8) Repeat steps (2)-(7) until value functions $\{V^l_{it}, V^s_{it}, V^h_{it}\}^{N,T}_{i=1,t=1}$ converge.

B Details of Data and Quantification

B.1 Data Sources for China

1. The **2000 Census** and **2010 Census** in China. These datasets provide prefecture-level population and skill ratios in the years 2000 and 2010. We aggregate the 2010 skill ratios at the country level, which is then used to identify the skill upgrading cost.

2. The **China's 2002 Industrial Classification for National Economic Activities** (GB/T 4754-2002) provides a detailed classification of 96 industries at a two-digit level. We exclude industries in agriculture and mining and the waste processing industry, resulting in a total number of 82 industries.

3. The **One Percent Population Survey** in 2005. We use this dataset to obtain prefecture-level bilateral migrant flows in 2005, the industry-level ratio of total skilled workers' income to total workers' income for 96 industries (industrial skill intensities), and prefecture-level skill premiums in 2005.

4. The **City Statistical Yearbooks** of China, from which we obtain prefecture-level GDP in 2000 and 2010, gross fixed capital formation from 1994 to 2000, and yearly investment price index for 1994-2000. We use these data on investment to construct prefecture-level capital stocks and then prefectural capital shares in 2000.

5. The **2002 China Input-Output Table**. The IO Table provides final consumption and capital income shares in the value-added for 42 industries at the two-digit level. We excluded agricultural and mining industries and manually mapped the remaining 37 industries with the 82 industries in the GB/T 4754-2002 classification so that the industry classification is consistent.

B.2 Data Sources for the Rest of the World (ROW)

1. The **OECD Statistics**. This database provides the initial population aged from 25 to 64 in the year 2000 for 33 countries, including China. We combine the total population from this source and the prefecture population share in the 2000 census to compute the prefectural population in the initial state. We aggregate the

remaining countries' populations as the population of the ROW. We also observe country-level shares of unskilled workers from the same database out of the total workers. We obtain the initial skill ratio of the ROW as the ratio of the ROW's total skilled workers to its total workers.

2. The **Penn World Table**. We use PWT version 10.0 to obtain initial capital stocks in the year 2000 for countries in the ROW and China. Each country's capital stock is in units of 2000 USD, where we use the exchange rate for the year 2000 from the National Account data in the same database. The initial capital stock of the ROW is computed as an aggregate of the capital stocks of all 32 countries in the list. We infer the prefectural-level capital stock by combining the total capital stock from PWT and the prefectural capital shares calculated from the City Statistical Yearbooks of China.

3. The **World Input-Output Database (WIOD)**. We use the WIOD 2016 Release to obtain China's sectoral trade-to-GDP ratios from 2000 to 2006. The World Input-Output Tables provide intercountry trade flows for 56 industries, including 19 manufacturing industries. China's national IO tables from 2000 to 2006 provide China's sectoral value added.

4. The **IPUMS USA**. We use the one-percent sample of the U.S. 2000 Census from IPUMS USA to infer the skilled workers' income share in each sector. We match China's 42 industries with the NAICS 2007 code to ensure consistent sector classification. We define skilled workers as workers with an education level of 12 grade or above, i.e., high school graduates or college graduates.

5. The **2002 Standard Make and Use Tables** of the U.S. provide capital and labor income share in the total value added at the 6-digit industry level. Again, we match China's 42 industries with the U.S. industry code in the 2002 Make and Use Tables. The labor income share is then divided between skilled and unskilled workers using the results from IPUMS USA.

Table 11: List of Countries in the ROW

Australia	Belgium	Canada	Costa Rica	Czech Republic
Denmark	Estonia	Finland	France	Germany
Greece	Hungary	Ireland	Italy	Japan
Korea	Latvia	Lithuania	Luxembourg	Mexico
Netherlands	New Zealand	Poland	Portugal	Slovak Republic
Slovenia	Spain	Sweden	Switzerland	Türkiye
United Kingdom	United States			

Notes: This table lists 32 OECD countries that are selected as the ROW because of their data availability.

Table 12: List of Port Cities

Tianjin	Tangshan	Qinhuangdao	Dalian	Dandong	Jinzhou	Shanghai
Suzhou	Nantong	Ningbo	Wenzhou	Jiaxing	Fuzhou	Xiamen
Quanzhou	Qingdao	Yantai	Weihai	Guangzhou	Shenzhen	Zhuhai
Shantou	Foshan	Jiangmen	Zhanjiang	Huizhou	Haikou	

Notes: This table lists the 27 prefectures that 1) import and export from the international markets in the Chinese Customs database and 2) are on the coast

B.3 Additional Tables

B.4 Sector Classification

We classify the 82 industries from GB/T 4754-2002 Chinese Classification (GB hereafter) into four sectors based on skill intensities: skilled manufacturing sector, unskilled manufacturing sector, skilled service sector, and unskilled service sector. Specifically, we compute the skill intensity of each industry by taking the ratio of skilled workers' income to total labor income for each industry. Then, we rank manufacturing and service industries separately by skill intensity. We treat industries above the median skill intensity as the *skilled-intensive industries* and group them to define the skilled sector. Those below the median skill intensity are aggregated as unskilled-intensive sectors. The skill intensities of each industry are estimated using the *One Percent Population Survey*. Table 13 shows the corresponding result.

To obtain sector-level capital and labor income share, we utilize the 2002 China Input-Output Table and match its 37 industries with 82 GB industries. Usually, the 2002 IO table industries each contain multiple GB industries. Since we define skilled sectors based on the GB system, if all GB industries within one IO table industry are classified as skilled industries, then the IO table industry is also considered skilled. For one IO table industry containing both skilled and unskilled GB industries, we consider the whole industry as a skilled one if there are more skilled GB industries within it than unskilled GB industries. Then, we aggregate IO Table industries into four sectors by skill intensity and compute the corresponding sectoral capital income shares as the total sectoral capital income ratio to sectoral value added.

Next, we use the WIOD World Input-Output table to obtain China's sector-level trade-to-GDP ratios. We only consider trade in the manufacturing sector and trade flows between China and 32 countries included in the ROW. To obtain sectoral imports and exports of China, we manually map the 19 manufacturing industries in the WIOD with 16 Chinese manufacturing industries in GB 42 industry classifications and define the skilled and unskilled sectors. From the WIOD, we also use China's national IO tables from 2000 to 2006 to obtain China's sectoral value added. Given imports and exports data and the value added, we compute China's sectoral import/export-to-GDP ratio between 2000 and 2014 by taking the ratio of import/export to value added.

Lastly, we use the U.S. sectoral income shares to represent those of the ROW. We manually match China's 42 IO table industries with the U.S. industry code from the 2002 Standard Make and Use Tables. Table 13 shows matching results. Then, we aggregate those industries into four sectors as before and compute skilled workers' income share as the ratio of total skilled workers' income to total workers' income for each sector.

B.5 Estimate Migration Costs

We provide the details on estimating the migration costs here from the *One Percent Population Survey* in 2005. The survey records the individual's current location and asks for the location one year and five years ago. In addition to the location history, we also observe the place of hukou registration.

We first note that the stock of migrants from location i to location g consists of current and past movers from i that choose to stay in g. To be specific, the ratio of migrant stock in location g with origin i to the origin city's stock of workers at time t can be expressed as

$$\bar{D}^d_{gi,t} = \frac{L^d_{it} D^d_{gi,t} + \sum_{\tau=1}^{\infty} L^d_{it-\tau} D^d_{gi,t-\tau} (D^d_{gg,t-\tau})^{\tau}}{L^d_{it}}, \ d = l, s. \tag{54}$$

Table 13: Sector Classification

Panel A: Manufacturing Sector

Unskilled					Skilled				
IO2002	Description	GB2002	USE2002	Skill	IO2002	Description	GB2002	USE2002	Skill
8	Clothing	C18-19	3140-3160	0.20	14	Primary metals	C32-33	331A-3315	0.48
9	Wood&Furniture	C20-21	3210,3370	0.20	18	Electrical equipment	C39	3351-3359	0.49
21	Other manufactures	C42	3399	0.20	12	Chemicals	C26-30	3251-3260	0.50
7	Textile	C17	3130	0.23	16	Machinery	C35-36	3331-3339	0.52
13	Nonmetallic mineral products	C31	3270	0.27	20	Instruments	C41	3345	0.56
15	Manufactures of metal	C34	3321-332B	0.31	17	Transportation equipment	C37	3361-336B	0.57
6	Food&Tabacco	C13-16	3110-3122	0.36	19	Telecommunication equipment	C40	3341-3344,3346	0.60
10	Paper&Printing	C22-24	3221,3222,3230	0.37	11	Petroleum	C25	3240	0.70

Panel B: Service Sector

Unskilled					Skilled				
IO2002	Description	GB2002	USE2002	Skill	IO2002	Description	GB2002	USE2002	Skill
26	Construction	E47-50	2301-2303	0.29	23	Electric power	D44	2211	0.76
38	Other services	N81-O83	8111-8140	0.35	33	Real estate	K72	5310	0.78
31	Accommodation&Food	I66-67	7210-7220	0.36	41	Culture,Sports & Entertainment	R88-92	5111.5120-5161,71A0-7130	0.78
27	Transportation	F51-58	4810-48A0,4920-4930	0.39	34	Rental&Business	L73-74	5321-5613	0.79
30	Wholesale&Retail	H63,H65	4200,4A00	0.47	37	Technical services	M76-N79	5413,5416	0.85
24	Natural gas	D45	2212	0.62	40	Health&Welfare	Q85-87	6210-6240	0.87
25	Water	D46	2213	0.68	42	Government	S93-97	S001-S009	0.89
28	Postal	F59	4910	0.74	29	Computer service	G60-62	5112,5170-5190,5415	0.90
36	Scientific service	M75	5417	0.92					
32	Finance	J68-71	52A0-5250	0.93					
39	Education	P84	6100	0.93					

Notes: This table shows the composition of the skilled and unskilled manufacturing sectors, and skilled and unskilled service sectors. Column "IO2002" shows the industry code in the 2002 China Input-Output table; "GB 2002" is from the industry classification system of China's 2002 Industrial Classification for National Economic Activities; "USE2002" refers to the code from the U.S. 2002 Standard Make and Use Tables; "Skill" shows the share of skilled labor income out of total labor income and thus indicates industry skill intensity. Both "GB 2002" and "USE2002" are manually matched with the IO2002 industry by description.

In the numerator on the right-hand side, the first term is migration flow from the origin i at period t, where $D^d_{gi,t}$ is the migration probability, and L^d_{it} is the population at the origin i. In addition to the most recent movers, the current stock of migrants from location i also includes those who moved τ periods ago $L^d_{it-\tau}D^d_{gi,t-\tau}$ and choose to stay in location g thereafter with a probability $(D^d_{gg,t-\tau})^\tau$. The second term in the numerator counts these migrants retrospectively from $\tau = 1$ to distance history.

Assume that the migrant stocks are observed at a steady state so that $D^d_{gi,t} = D^d_{gi}$, then this ratio can be simplified as:

$$\bar{D}_{gi}^d = \frac{D_{gi}^d}{1 - D_{gg}^d}, \tag{55}$$

where D_{gi}^d is defined in the model as

$$D_{gi}^d = \frac{\exp\left[(\beta v_g^d - \kappa_{gi}^d)/\rho\right]}{\sum_{n=1}^N \exp\left[(\beta v_n^d - \kappa_{ni}^d)/\rho\right]}. \tag{56}$$

Therefore, double differencing the migrant stock share yields our main structural equation that can be used to estimate the migration costs:

$$\frac{\bar{D}_{gi}^d}{\bar{D}_{ii}^d}\frac{\bar{D}_{ig}^d}{\bar{D}_{gg}^d} = \frac{D_{gi}^d}{D_{ii}^d}\frac{D_{ig}^d}{D_{gg}^d} = \exp\left[-\frac{1}{\rho}(\kappa_{gi}^d + \kappa_{ig}^d)\right], \tag{57}$$

where we use the result

$$\frac{D_{gi}^d}{D_{ii}^d} = \frac{\exp\left[(\beta v_g^d - \kappa_{gi}^d)/\rho\right]}{\exp\left[(\beta v_i^d)/\rho\right]} = \exp\left\{\left[\beta\left(v_g^d - v_i^d\right) - \kappa_{gi}\right]/\rho\right\}. \tag{58}$$

B.6 Quantification in the Initial Equilibrium

In this section, we elaborate how to calibrate the prefecture-level fundamental productivity, $\{\bar{z}_i\}$, production function parameters, $\{\mu_j, \lambda_j\}$, and the trade costs between China port cities and the ROW, $\{\tau_{ROW,2000}^j\}$, by inverting the initial equilibrium. We drop the time subscript for all variables in the following for clarity, knowing that all variables and conditions are at the initial period.

B.6.1 Define the Initial Equilibrium

Given the initial distribution of local population and capital stock, $\{L_i^l, L_i^s, L_i^h, k_i\}$, the initial static equilibrium is defined as a set of prices $\{w_i^l, w_i^h, r_i\}$ that clear all labor and capital markets at the initial period:

$$w_i^l = \frac{\sum_{j=1}^J \phi_i^{lj} X_i^j}{\tilde{L}_i^l} \tag{59}$$

$$w_i^h = \frac{\sum_{j=1}^J \phi_i^{hj} X_i^j}{L_i^h} \tag{60}$$

$$r_i = \frac{\sum_{j=1}^J \phi_i^{kj} X_i^j}{k_i} \tag{61}$$

where the total revenue in location i industry j, X_i^j, and factor income shares $\{\phi_i^{lj}, \phi_i^{lj}, \phi_i^{lj}\}$, are given by

$$X_i^j = \gamma_j \sum_{n=1}^N \pi_{ni} X_n = \gamma_j \sum_{n=1}^N \pi_{ni} \sum_{s=1}^J X_n^s \tag{62}$$

$$\phi_i^{lj} = \left[1 + \frac{1-\mu^j}{\mu^j}\left(\frac{w_i^l}{w_i^e}\right)^{\sigma-1}\right]^{-1}, \tag{63}$$

$$\phi_i^{hj} = \left[1 + \frac{\mu^j}{1-\mu^j}\left(\frac{w_i^e}{w_i^l}\right)^{\sigma-1}\right]^{-1}\left[1 + \frac{\lambda^j}{1-\lambda^j}\left(\frac{w_i^h}{r_i}\right)^{\eta-1}\right]^{-1}, \tag{64}$$

$$\phi_i^{kj} = \left[1 + \frac{\mu^j}{1-\mu^j}\left(\frac{w_i^e}{w_i^l}\right)^{\sigma-1}\right]^{-1}\left[1 + \frac{1-\lambda^j}{\lambda^j}\left(\frac{r_i}{w_i^h}\right)^{\eta-1}\right]^{-1}. \tag{65}$$

B.6.2 Solution Algorithem

Given the initial conditions $\{L_i^l, L_i^s, L_i^h, k_i\}$, we solve the initial equilibrium as

(1) start with an initial guess of wages and interests $\{w_i^{l(0)}, w_i^{s(0)}, r_i^{(0)}\}_{i=1}^N$,

(2) compute initial factor incomes shares $\{\phi_i^{lj}, \phi_i^{hj}, \phi_i^{kj}\}_{i=1}^N$ from (63), (64), (65) and total output $\{X_i^j\}$ from (62),

(3) solve new factor prices by (59), (60), (61),

(4) and iterate until factor prices converge.

After solving the initial equilibrium, we can compute the equilibrium prefecture-level GDP share, $\frac{\sum_j X_i^j}{\sum_n \sum_j X_n^j}$, average factor income share, $\{\bar{\phi}^{lj}, \bar{\phi}^{hj}, \bar{\phi}^{kj}\}$, and China's total trade-to-output ratio. These outcomes are determined by the local fundamental productivity, the weight parameters in the production function, and the trade costs between China and the ROW respectively. We then adjust those parameters in the initial equilibrium such that the model-generated moments match the data counterparts.

C Additional Quantitative Results

C.1 Short-run Impacts of Trade and Infrastructure

We provide the results of the short-run impacts (at year 2010) of China's trade liberalization and infrastructure expansion here. In general, as capital accumulation and skill formation need to take time, the short-run impacts are not as strong as the long-run impacts, and the interaction between capital and skill formation is not clear as well. Nevertheless, the results are still suggestive.

Table 14 and 16 show that capital adjustment amplifies the short-run welfare gains, and skill upgrading widens the between-skill inequality driven by trade liberalization. Table 15 and 17 also show that the results on spatial dispersions of local impacts are consistent with the main results in the paper.

Table 14: Short-run Aggregate Impacts of Trade Liberalization

	%Changes in 2010						
	Capital	S.R.	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	0.92	0.12	1.40	2.27	0.80	2.34	2.14
With Upskilling	-	0.08	1.21	1.88	0.64	1.78	1.33
With Investment	0.94	-	1.42	2.29	0.79	2.31	1.68
Basic Model	-	-	1.27	1.85	0.53	1.76	1.23

Notes: This table reports the short-run % changes of aggregate economic outcomes in China induced by a uniform infrastructure shock under different model specifications.

Table 15: Spatial Dispersion of Trade Liberalization's Impacts

	Standard Deviation of Changes in 2010							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	0.38	0.12	0.20	0.57	0.82	0.26	0.96	0.50
With Upskilling	0.00	0.08	0.15	0.51	0.70	0.20	0.77	0.34
With Investment	0.39	0.13	0.19	0.56	0.80	0.25	0.94	0.45
Basic Model	0.00	0.09	0.15	0.51	0.68	0.19	0.76	0.34

	Max %Changes - Min %Changes in 2010							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	1.78	0.68	1.43	2.65	4.21	1.79	4.74	2.76
With Upskilling	0.00	0.66	1.21	2.61	3.70	1.45	4.06	1.88
With Investment	1.81	0.86	1.30	2.64	4.32	1.93	4.60	2.32
Basic Model	0.00	0.81	1.14	2.60	3.78	1.54	3.95	1.98

Notes: This table reports the spatial dispersion of % changes of local economic outcomes in China induced by liberalization under different model specifications.

Table 16: Short-run Aggregate Impacts of Infrastructure Improvement

	%Changes in 2010						
	Capital	S.R.	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	0.37	0.12	1.74	1.51	-0.18	1.33	5.56
With Upskilling	-	0.07	1.60	1.33	-0.21	1.26	4.18
With Investment	0.40	-	1.64	1.61	0.06	1.39	4.48
Basic Model	-	-	1.53	1.38	-0.05	1.29	3.77

Notes: This table reports the short-run % changes of aggregate economic outcomes in China induced by infrastructure shock under different model specifications.

Table 17: Spatial Dispersion of Infrastructure Improvement's Impacts

	Standard Deviation of Changes in 2010							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	0.52	0.58	1.31	1.35	1.17	0.84	2.08	3.25
With Upskilling	0.00	0.49	1.12	1.23	1.04	0.79	1.71	2.71
With Investment	0.54	0.67	1.37	1.33	1.17	0.87	2.12	3.13
Basic Model	0.00	0.57	1.19	1.22	1.05	0.82	1.76	2.76

	Max %Changes - Min %Changes in 2010							
	Capital	S.R.	Pop	w^l/p	w^h/p	S.P.	GDP	C.E.
Full Model	4.55	4.50	10.22	9.74	9.31	6.70	13.97	18.83
With Upskilling	0.00	4.07	9.42	8.81	7.44	6.62	11.37	16.05
With Investment	4.64	5.24	10.69	9.62	9.56	6.81	14.47	18.16
Basic Model	0.00	4.85	9.65	8.85	7.59	6.82	11.81	16.24

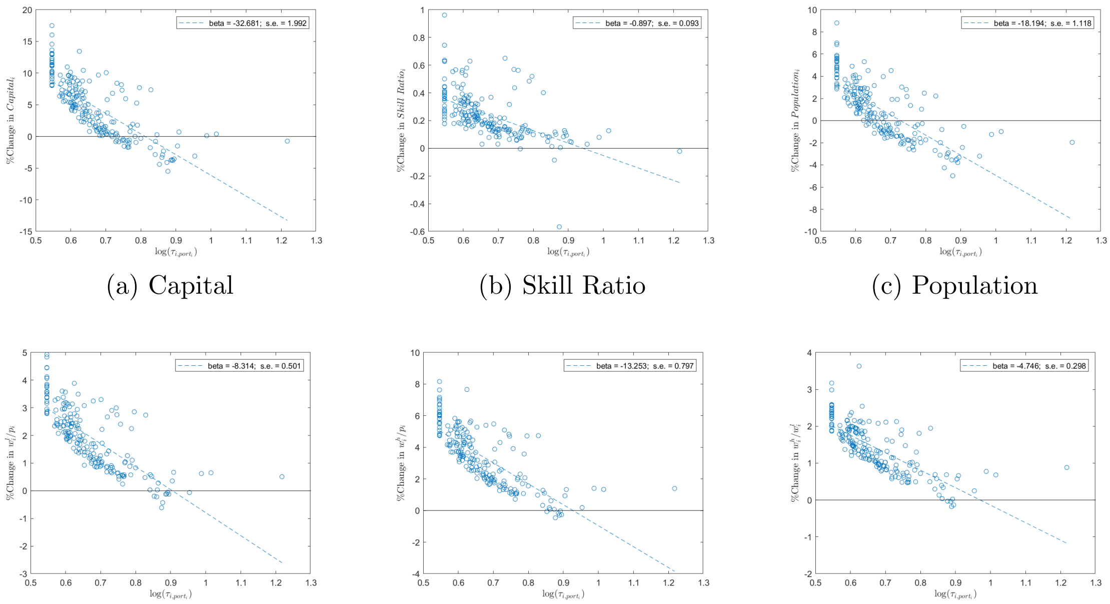
Notes: This table reports the spatial dispersion of % changes of local economic outcomes in China induced by infrastructure improvement under different model specifications.

C.2 Spatial Impact in Steady State

In this part, we provide more details on the spatial impacts of trade liberalization and infrastructure improvement in the full model simulation. We are interested in how the events affect each city in China differently and try to capture some general spatial patterns.

C.2.1 Spatial Impact of Trade Liberalization

We use each city's distance to the ROW, $\log(\tau_{i,port})$, to differentiate cities in China. In general, coastal cities in China are both geographically and economically closer to the ROW and have a shorter distance. Therefore, they have a larger exposure to trade liberalization. Figure 6 shows that coastal cities are the main beneficiary from trade liberalization: they accumulate more capital, attract (skilled) workers from inland cities, and have larger wage increases, as suggested by the negative distance elasticities. Note that the skill premium also increases the most in the coastal cities due to the larger trade-induced capital gains. Figure 7 further confirms this result and shows more spatial details. Here we use the impacts on unskilled wages to illustrate the trade's spatial impacts, with the notion that other outcomes, such as capital changes, population changes, and output changes, follow a similar pattern.



(a) Capital (b) Skill Ratio (c) Population

(d) Unskilled Real Wage (e) Skilled Real Wage (f) Skill Premium

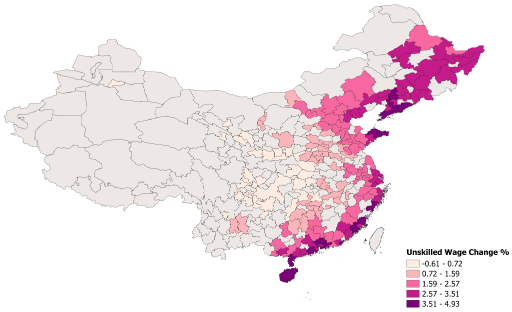
Figure 6: Spatial Impacts of Trade Liberalization in Steady States

Notes: these figures show the spatial impacts of China's accession to WTO on production factors, real wages, and skill premiums across Chinese prefectures in steady states. The results come from the full model simulation. The horizontal axis measures each city's distance to the ROW. Each dot represents a prefecture-level city. The straight lines are linear fits.

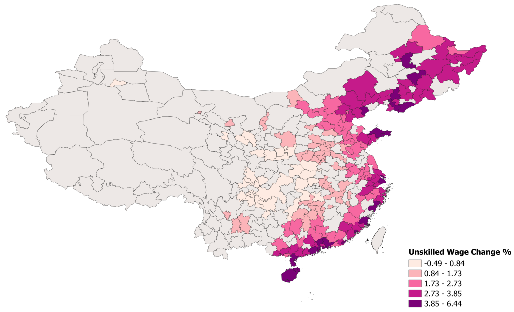
C.2.2 Spatial Impacts of Infrastructure Improvement

In order to measure the spatial impacts of China's infrastructure improvement between 2000 and 2015, we construct a typical "market access" measure for each city at the year 2000 and 2015 and use it to measure each city's exposure to the infrastructure expansion. Cities with larger improvements in domestic connectiveness during this time of period have larger increases in their market access.

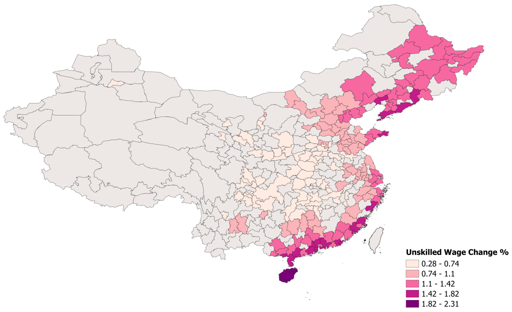
As shown in figure 8 and 9, the main beneficiaries of the infrastructure expansion are central and western cities, whose domestic market access increases the most. In the end, the expansion induces a larger capital gain, population gain, and real wage increase in those cities. Since central and western cities are generally poorer than coastal cities, the



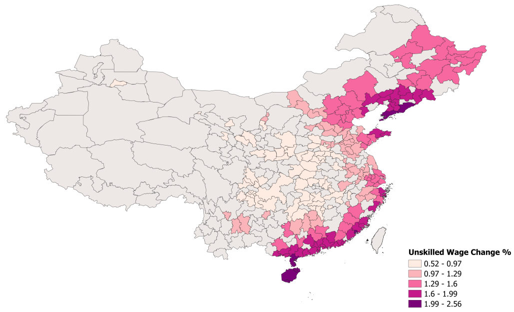
(a) Full Model



(b) With Capital Invest.



(c) With Upskilling

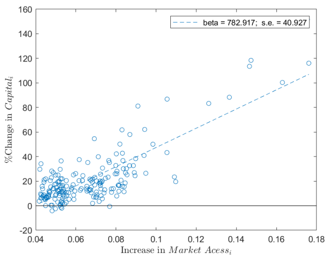


(d) Basic Model

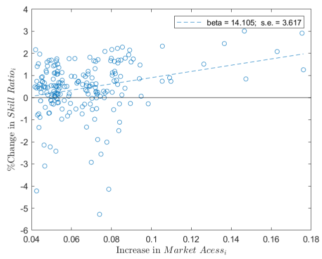
Figure 7: Spatial Impacts of Trade Liberalization on Unskilled Wage

Notes: these figures show the spatial impacts of China's accession to WTO on unskilled wages across Chinese prefectures in steady states. Panel (a) is from the full model simulation; panel (b) from the simulation with only capital investment; panel (c) from the simulation with only skill upgrading; panel (d) from the basic model simulation, without both capital accumulation and upskilling.

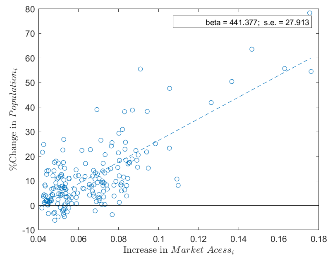
infrastructure expansion also helps to reduce regional inequality in China.



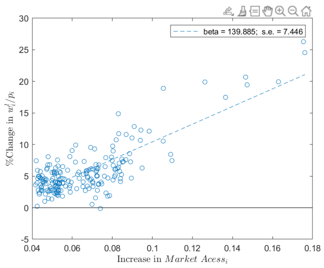
(a) Capital



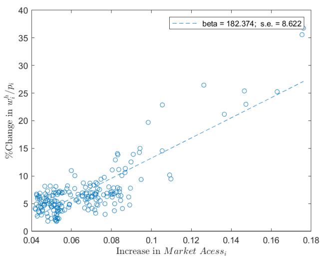
(b) Skill Ratio



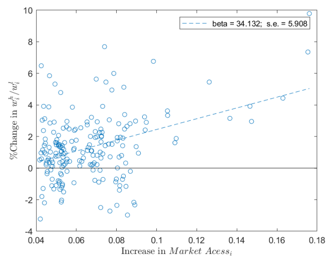
(c) Population



(d) Unskilled Real Wage



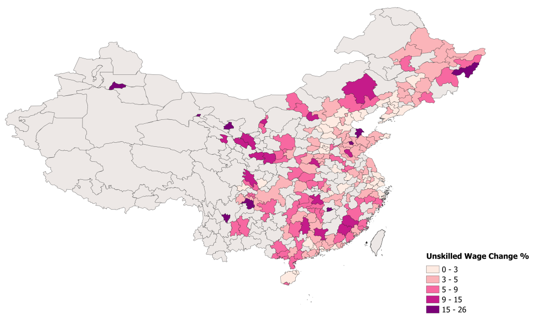
(e) Skilled Real Wage



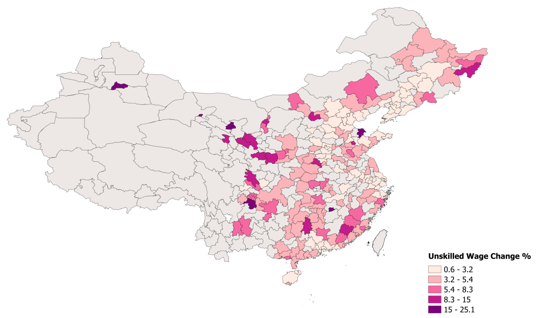
(f) Skill Premium

Figure 8: Spatial Impacts of Infrastructure Improvement in Steady States

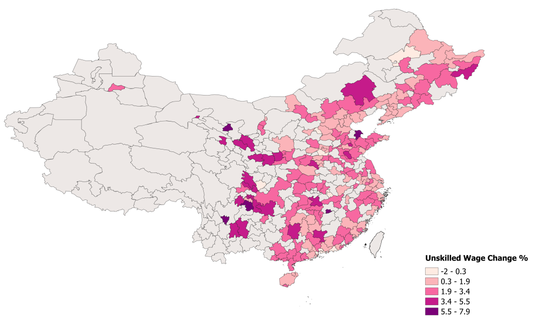
Notes: these figures show the spatial impacts of infrastructure improvement in China during 2000-2015 on production factors, real wages, and skill premiums across Chinese prefectures in steady states. The results come from the full model simulation. The horizontal axis measures each city's increase in market access due to infrastructure improvement. Each dot represents a prefecture-level city. The straight lines are linear fits.



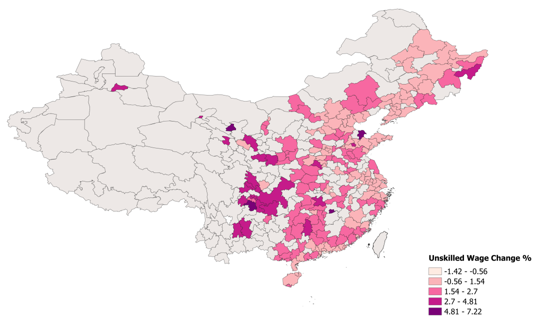
(a) Full Model



(b) With Capital Invest.



(c) With Upskilling



(d) Basic Model

Figure 9: Spatial Impacts of Infrastructure Improvement on Unskilled Wage

Notes: these figures show the spatial impacts of China's infrastructure expansion between 2000-2015 on unskilled wages across Chinese prefectures in steady states. Panel (a) is from the full model simulation; panel (b) from the simulation with only capital investment; panel (c) from the simulation with only skill upgrading; panel (d) from the basic model simulation, without both capital accumulation and upskilling.