2-2024

The long and short-run spatial impacts of trade

Lin MA
Singapore Management University, linma@smu.edu.sg

Yunlong SONG
Singapore Management University, ylsong.2020@phdecons.smu.edu.sg

Yang TANG
Nanyang Technological University

SMU ECONOMICS &
STATISTICS

The Long and Short-Run Spatial Impacts of Trade

Lin Ma, Yunlong Song, Yang Tang

February 2024

Paper No. 03-2024

The Long and Short-Run Spatial Impacts of Trade

Lin Ma Yunlong Song Yang Tang *

February 21, 2024

Abstract

We explore how the spatial impacts of trade evolve over time using a dynamic spatial model that incorporates capital accumulation and skill acquisition. We show that in the short run, the spatial impacts of trade mainly depend on the initial conditions, especially the endowments of physical and human capital across locations. However, in the long run, trade shocks shape the distribution of production factors across space through factor accumulation and migration, resulting in significantly different spatial impacts. In the context of China's WTO accession, we find that international trade is seven times more effective in driving the population towards coastal areas in the long run than in the short run. The skill composition of trade-induced migration exhibits a reversal over time. Furthermore, we demonstrate that policies designed to alleviate the localized impacts of globalization would be misdirected and underfunded if policymakers overlook the intertemporal variations in the spatial impacts of trade.

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

JEL Classification: F12;O11;R12

*Respectively: Singapore Management University (linma@smu.edu.sg); Singapore Management University (ylsong.2020@phdecons.smu.edu.sg); Nanyang Technological University (tangyang@ntu.edu.sg). 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

International trade exerts profound distributional impacts across regions within a country. While a country gains from trade on average, workers in a particular region could benefit or lose, depending on the location's sectoral composition, factor endowments, and position in internal geography within the country.[1] For example, as shown in Autor et al. (2013), U.S. labor markets specializing in industries that bore the brunt of Chinese import competition performed poorly relative to other locations that escaped the "China Shock". Across the Pacific in China, export booms benefited the coastal locations much more than the inland ones, drawing resources away from the latter and contributing to rising spatial inequality (Fan, 2019; Ma and Tang, 2020). The uneven spatial impacts fueled heated political debates and were one of the major reasons behind the recent populist movement against globalization around the world (Feigenbaum and Hall, 2015; Colantone and Stanig, 2018; Autor et al., 2020). Understanding the spatial impacts of trade, therefore, is crucial for designing policies that could effectively alleviate the localized impacts of globalization.

We contribute to this literature by highlighting a previously overlooked dimension: time. We show that the answers to key questions on the spatial impacts of trade, and consequently, the design of place-based policies, vary depending on the time horizon. For instance, one pivotal question about the spatial impacts of trade concerns how skilled and unskilled workers migrate differently in response to trade shocks. Consider the case of China upon joining the World Trade Organization (WTO) in 2001. At that time, China enjoyed comparative advantages in unskilled-intensive industries, and the export boom led to the migration of unskilled workers towards coastal locations in the short run. In the long run, however, the export boom also allows for faster accumulation of physical capital — better infrastructure, buildings, and machinery — at the coastal locations. Higher capital stocks complement the productivity of skilled workers more than unskilled ones, making coastal locations more attractive to skilled workers over time. In this specific example, the skill composition of migrants to coastal locations reverses as time goes by. The variation in spatial impacts of trade stems from two opposing forces. In the short run, spatial impacts are predominantly

[1] In the context of the U.S., for example, see Autor et al. (2013) for empirical evidence and Caliendo et al. (2019) for quantitative exercise. Also, see Fan (2019) in the context of China.

determined by initial conditions. However, in the long run, factor endowments respond to trade shocks through accumulation and redistribution. The endogenous response of production factors changes the channels through which trade shocks manifest themselves, thereby leading to significantly different spatial impacts of trade over the long term. How do other variables, such as population distribution, skill premium, and skill acquisition, respond to trade shocks in the short and long run? What are the implications for place-based policies designed to mitigate the negative impacts of globalization? We aim to provide some initial answers to these questions.

To systematically study the spatial impacts of trade over a long time horizon, we develop a dynamic spatial model with endogenous physical and human capital accumulation. Our framework highlights the spatial variation in the accumulation and distribution of physical and human capital in response to trade shocks. 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 face type-specific migration frictions, and production exhibits capital-skill complementarity. The supply 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 of skilled workers and the skill upgrading decisions made by local unskilled workers. In the short run, the migration frictions and the gradual nature of factor accumulation imply that the initial factor endowments mainly determine the spatial impacts of trade. In the long run, however, the distribution of production factors shifts across space due to capital accumulation, skill acquisition, and migration in response to economic shocks. As a result, the model can deliver spatial impacts of trade that change over time.

We quantify our model in the context of China, a country that experienced drastic trade liberalization and massive internal migration in the past two decades. We model four sectors that differ in trade costs and factor intensity and map the geographical units 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 along the transition path. Consistent with previous literature, migration costs are significant, equivalent to 28.7 to 42.1 percent of one's lifetime utility. The skill upgrading costs are also considerable, equivalent to 48.2 percent of an unskilled worker's lifetime utility. The sizable frictions across space and skill types predict a slow response to a given shock and a potentially large difference between the long and the short-run spatial responses to trade shocks.

We focus on three key questions that characterize the spatial impacts of trade: 1) how do trade shocks affect the population distribution across space? 2) how do skilled and unskilled workers migrate in response to trade shocks? and 3) how does the skill premium in different locations respond to trade shocks? To answer these questions, we compare the baseline simulation that captures factual trade liberalization to a counterfactual economy without trade liberalization. We then measure the impacts of trade on population movement, skill composition, and skill premium and study how the impacts change over time.

The answers to all three questions depend on the time horizon. Regarding population movement, in the short run, the distance elasticity of the population is -0.035: for every 1 percent reduction in the distance to the ports, international trade leads to 0.035 percent more population. In the long run, the distance elasticity increases seven times in absolute values to -0.248. This suggests that international trade is seven times more effective in driving the population towards coastal locations in the long run compared to the short run. The disparity in long- and short-run distance elasticity arises from accelerated capital formation in coastal locations, fueled by increased returns on capital following the export boom. We find that shutting down capital accumulation would reduce the gap between the long and the short-run elasticity by around 85 percent. On the contrary, shutting down skill upgrading would further widen the gap in distance elasticity between the long- and short-run by 26.8 percent. This is because, without local skill upgrading in coastal locations, the demand for skilled workers has to be met through internal migration, further increasing the population growth along the coast.

The impacts of trade on migration destinations of skilled and unskilled migrants are even more striking. In the short run, unskilled workers are much more likely to move to the coastal locations than skilled ones, driven by the export boom in the unskilled-intensive industries. In the long run, however, skilled workers are much more likely to migrate toward the coastal cities than unskilled ones. Put differently, the distance elasticity of skill composition switches signs as time goes by. In the context of China, the skill composition of migrants depends on two counteracting forces. In the short run, comparative advantage in unskilled intensive industries tends to increase demand for unskilled workers, attracting unskilled migration towards the coastal locations. In the long run, however, the export boom enables faster capital accumulation and, subsequently, higher demand for skilled workers via capital-skill complementarity. Skill upgrading again works to narrow the gap between the long- and the short-run impacts: without skill upgrading, the migration flow towards coastal locations is even more dominated by the skilled workers in the long run, as the higher demand for them cannot be met with locally trained unskilled workers anymore. The spatial impact on skill premium follows a similar pattern over time. In the short run, trade reduces skill premiums more in coastal locations due to their comparative advantage in unskilled-intensive industries. In the long run, the skill premiums in the coastal locations increase instead due to capital-skill complementarity.[2]

Lastly, we demonstrate that policies designed to mitigate the spatial impacts of trade would be ineffective if they disregard temporal variations. To highlight this point, we implement a hypothetical policy that subsidizes workers in the less developed inland regions to prevent trade-induced population loss. This hypothetical policy resembles real-world policies such as China's "Western Development Policies" or Japan's "Local Revitalization Policy" that aim to prevent population loss in less developed regions. We find that to revert 50 percent of the trade-induced population loss in the less-developed provinces within ten years after the trade shock would require the central government to spend 0.017 percent of the GDP each year as wage subsidies. However, given the growing attractiveness of the coastal locations, the subsidy rate above would be ineffective in the long run. To achieve the same goal of halving the trade-induced population loss in a steady state, the wage subsidy would need to increase by 65 percent to 0.028 percent of the GDP instead.

Other place-based policies suffer similar issues if the policymaker ignores the temporal variations of spatial responses to trade shocks. No subsidy would be needed in the short run

[2]Note that the long-term result is still consistent with the prediction from the Stopler-Samuelson theorem. In the long run, the trade-induced changes in factor endowments reshaped the pattern of comparative advantage so that the coastal locations started to specialize in capital- and skill-intensive industries. As a result, the abundant factor in the long run (the skilled workers) benefits from trade through higher skill premiums.

to prevent trade-induced investment loss in the inland locations. However, the faster capital accumulation in the coastal locations implies that a capital-return subsidy equivalent to 0.028 percent of GDP would be needed to revert 50 percent of the trade-induced investment loss in the long run. Policies that aim to promote skill upgrading need to be adjusted as well. In the short run, the policy should spend more in the coastal areas as the unskilled workers, attracted by the boom in the unskilled intensive industries, would hesitate to upgrade their skills in the short run. However, in the long run, the policy emphasis should shift towards inland areas. This transition occurs as the lack of physical investment in the inland regions reduces the incentives for skill acquisition in those areas. Overall, our findings suggest that given the substantial gap in the impacts of trade between the short run and long run, policy designs need to be flexible and responsive to temporal changes.

This paper mainly speaks to a broad literature investigating the distributional impact of trade (Feenstra and Hanson, 1996; Goldberg and Pavcnik, 2007; Helpman et al., 2010; Autor et al., 2013; Burstein and Vogel, 2017; Autor et al., 2021). Many papers in this literature consider distributional impact across skill types, assume fixed endowment 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 trade shocks differently across space within a country, so regional policies must be adjusted accordingly.

Our work is also related 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 physical capital formation and human capital acquisition across space and time. 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.

Finally, we also contribute to the literature studying China's spatial economy (Fan, 2019; Tombe and Zhu, 2019; Ma and Tang, 2020, 2024). Our work is closest to Fan (2019), which considers the spatial impacts of trade in a static model with exogenously determined capital stock and skill types. 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 responses to trade shocks across space and time horizons.

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

2.1 The Environment

The model endogenizes skill formation and incorporates capital-skill complementarity into a general equilibrium dynamic spatial similar to Kleinman et al. (2023). The economy has N geographically segmented locations indexed by i and J sectors indexed by j. Time is discrete and indexed by $t = 0, 1, 2, \cdots, \infty$. Two types of agents, workers and landlords, reside in each location.

2.2 Workers

Workers differ in skill levels: skilled or unskilled. Regardless of skill level, workers inelastically supply one unit of labor each period and earn income accordingly. Workers do not save, so they consume all their income each period. At the end of each period, all workers decide where to migrate, and unskilled workers decide whether to upgrade their skills.

The worker's flow utility depends on their consumption bundle:

$$c = \prod_{j=1}^{J} \left(\frac{c^j}{\gamma^j} \right)^{\gamma^j}.$$

In the expression above, the expenditure share on goods produced by industry j is γ^j and $\sum_j \gamma^j = 1$. The industry-level consumption, c^j, is a constant elasticity of substitution (CES) aggregator over N varieties available in j:

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

where θ is the elasticity of substitution among varieties available in industry j.

After production and consumption in the current location, a worker decides where to live in the next period. The migration decision depends on three elements: 1) the expected lifetime utility from living in any of the J locations, 2) an idiosyncratic preference shock that follows an extreme value distribution towards each destination denoted as ε_{nt}, and 3) the skill-specific bilateral migration costs. We denote the migration costs as $\kappa^d_{ni,t}$ for a worker with skill $d \in \{l, s\}$ to migrate from i to n at time t. We use superscript l to denote unskilled workers and s to denote skilled workers. Standard properties on bilateral migration cost $\kappa^d_{ni,t}$ apply: (1) $\kappa^d_{ni,t} > 0$ for $n \neq i$, (2) $\kappa^d_{ii,t} = 0$, and (3) $\kappa^d_{ni,t} \leq \kappa^d_{nj,t} + \kappa^d_{ji,t}$ for any third location j.

In addition to the migration decision, an unskilled worker decides whether to acquire skills subject to switching costs of κ^l_s in the unit of utility, following Chang et al. (2022). Skilled workers cannot downgrade to unskilled, and we normalize the costs of staying as skilled workers to zero.[3] As it will be clear later, the skill-upgrading decision depends on comparing the option value of being a skilled worker in location i in the next period against the cost of upgrading. The option value of a skilled worker, in turn, reflects not only the skill premium at location i but also the option value of migrating to other locations as a skilled worker starting from location i in the future. Lastly, we introduce an i.i.d exogenous exit shock so that each individual has $\xi \leq 1$ 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.

In summary, a worker with current skill level d living in location i at time t solves the

[3]Specifically, we assume (1) $\kappa^d_d = 0$ for $d \in \{l, s\}$, (2) $\kappa^s_l = \infty$ and (3) $\kappa^l_s < \infty$.

following recursive problem:

$$V_{it}^{d} = \ln b_{it} + \ln \frac{w_{it}^{d}}{p_{it}} + \max_{\{n,e\}}\{\xi\beta\mathbb{E}V_{nt+1}^{e} - \kappa_{ni,t}^{d} - \kappa_{e}^{d} + \rho\varepsilon_{nt}^{e}\},$$

where V_{it}^{d} is the value of skill-type d at location i at time t, and w_{it}^{d} is the skill-specific wage rate. b_{it} is the amenity, $p_{it} = \prod_{j=1}^{J} \left(p_{it}^{j}\right)^{\gamma^{j}}$ is aggregate price index at location i, β is the discount rate, and ρ controls the dispersion of mobility shocks. The individuals choose the future location, n, and the skill type, e, simultaneously subject to the expected future value and frictions. The term $\mathbb{E}V_{n,t+1}^{e}$ is the expected value of being type-e at location n in the next period, where the expectation is taken over realizations of future shocks. The idiosyncratic preference shocks ε_{nt}^{e} are i.i.d across types of workers, locations, and time, following the Gumbel distribution with the cumulative distribution function (CDF): $F(\varepsilon) = e^{e^{(-\varepsilon-\bar{\gamma})}}$, where $\bar{\gamma}$ is the Euler-Mascheroni constant. Define $v_{it}^{d} = \mathbb{E}V_{it}^{d}$ as the expected lifetime utility. Applying standard properties of the Gumbel distribution gives:

$$v_{it}^{d} = \ln b_{it} + \ln \frac{w_{it}^{d}}{p_{it}} + \rho \ln \sum_{e=l}^{s} \sum_{g=1}^{N} \exp\left[(\xi\beta v_{gt+1}^{e} - \kappa_{gi,t}^{d} - \kappa_{e}^{d})/\rho\right]. \tag{1}$$

We can also compute the fraction of type d workers in location i and time t that migrate to g and become type e in $t+1$ as

$$D_{gi,t}^{ed} = \frac{\exp\left[(\xi\beta v_{gt+1}^{e} - \kappa_{gi,t}^{d} - \kappa_{e}^{d})/\rho\right]}{\sum_{d=l}^{s}\sum_{n=1}^{N}\exp\left[(\xi\beta v_{nt+1}^{d} - \kappa_{ni,t}^{d} - \kappa_{e}^{d})/\rho\right]}, \tag{2}$$

where $1/\rho$ in Equation (2) captures the migration elasticity. Finally, the supply of unskilled workers and skilled workers in each location evolves as follows:

$$L_{it+1}^{l} = \xi \sum_{n=1}^{N} D_{in}^{ll} L_{nt}^{l} + \left(L_{it}^{l} + L_{it}^{s}\right)(1-\xi), \tag{3}$$

and

$$L_{it+1}^{s} = \xi \left(\sum_{n=1}^{N} D_{in}^{ss} L_{nt}^{s} + \sum_{n=1}^{N} D_{in}^{sl} L_{nt}^{l}\right). \tag{4}$$

The supply of unskilled workers in location i is the combination of unskilled migration inflows with the newborn population, and the supply of skilled workers contains the inflows of skilled workers and local unskilled workers who upgrade their skills.

2.3 Landlords

We closely follow Kleinman et al. (2023) in modeling landlords. Landlords are immobile and have access to the financial market. With an initial endowment of capital stock, the landlords optimally choose the sequences of consumption and investments to maximize their lifetime utility. In the baseline model, the landlords can only invest in their local markets. In the extension, we show that the results are robust if we allow cross-location investments. 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:

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

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 capital accumulation equation can thus be characterized as:

$$k_{it+1} = \xi\beta \left(1 - \delta + \frac{r_{it}}{p_{it}}\right) k_{it}. \tag{5}$$

The equation above highlights the key mechanism through which trade shocks affect capital accumulation. A location that benefits from a positive demand shock sees higher real returns to capital stock, r_{it}/p_{it}. Higher returns encourage investment and lead to faster capital accumulation in the following periods.

2.4 Production

Firms at each location i and industry j specialize in one variety and operate in a perfectly competitive market, using unskilled workers (l^j_{it}), skilled workers (s^j_{it}), and capital (k^j_{it}) as inputs. The production function in location i and industry j at time t features a nested CES functional form as:

$$y^j_{it} = z_{it}\left[(\mu^j)^{\frac{1}{\sigma}}(z^{-\psi}_{it}l^j_{it})^{\frac{\sigma-1}{\sigma}} + (1-\mu^j)^{\frac{1}{\sigma}}(z^{\psi}_{it}h^j_{it})^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}},$$

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

$$h^j_{it} = \left[(\lambda^j)^{\frac{1}{\eta}}(k^j_{it})^{\frac{\eta-1}{\eta}} + (1-\lambda^j)^{\frac{1}{\eta}}(s^j_{it})^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}.$$

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. We assume $\sigma > \eta$ so that capital is more complementary with skilled workers than unskilled ones. This is to capture capital-skill complementarity for all industries.[4] Following Burstein and Vogel (2017), the elasticity ψ disciplines the strength of skilled-biased productivity and is assumed to satisfy $\psi(\sigma-1) > 0$.

Both unskilled and skilled workers are perfectly mobile across sectors within a location. The production structure implies that the unit cost of production for a variety in the industry j and location i, denoted as c^j_{it}, is given by:

$$c^j_{it} = \frac{1}{z_{it}}\left[\mu^j(z^{\psi}_{it}w^l_{it})^{1-\sigma} + (1-\mu^j)(z^{-\psi}_{it}w^h_{it})^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{6}$$

[4]For more details, see Duffy et al. (2004).

In the above expression, w^l_{it} is the wage rate of an unskilled worker, and w^h_{it} is the unit cost of an equipped skilled worker, which can further be expressed as a function of skilled wage, w^s_{it}, and the rental price of capital, r_{it}:

$$w^{hj}_{it} = \left[\lambda^j (r_{it})^{1-\eta} + (1-\lambda^j)(w^s_{it})^{1-\eta}\right]^{\frac{1}{1-\eta}}. \tag{7}$$

Combining solutions from the profit maximization problem and zero profit condition, we can further obtain income shares of unskilled labor (ϕ^{lj}_{it}), skilled labor (ϕ^{sj}_{it}), and capital (ϕ^{kj}_{it}) for industry j respectively:

$$\phi^{lj}_{it} = \left[1 + z_{it}^{2\psi(\sigma-1)} \frac{1-\mu^j}{\mu^j} \left(\frac{w^l_{it}}{w^h_{it}}\right)^{\sigma-1}\right]^{-1}, \tag{8}$$

$$\phi^{sj}_{it} = \left[1 + z_{it}^{-2\psi(\sigma-1)} \frac{\mu^j}{1-\mu^j} \left(\frac{w^h_{it}}{w^l_{it}}\right)^{\sigma-1}\right]^{-1} \left[1 + \frac{\lambda^j}{1-\lambda^j} \left(\frac{w^s_{it}}{r_{it}}\right)^{\eta-1}\right]^{-1}, \tag{9}$$

$$\phi^{kj}_{it} = \left[1 + z_{it}^{-2\psi(\sigma-1)} \frac{\mu^j}{1-\mu^j} \left(\frac{w^h_{it}}{w^l_{it}}\right)^{\sigma-1}\right]^{-1} \left[1 + \frac{1-\lambda^j}{\lambda^j} \left(\frac{r_{it}}{w^s_{it}}\right)^{\eta-1}\right]^{-1}. \tag{10}$$

We assume standard iceberg trade costs between locations. In any industry j, the price of a variety in location n imported from location i ($p^j_{ni,t}$) is

$$p^j_{ni,t} = \tau_{ni,t} c^j_{it}.$$

Lastly, as shown in the Appendix, the price index in location n and industry j, denoted as $p^j_{n,t}$, satisfies:

$$\begin{aligned}(p^j_{nt})^{1-\theta} = \sum_{i=1}^{I} (\frac{\tau_{ni,t}}{z_{it}})^{1-\theta} \Big[& \mu^j z_{it}^{\psi(1-\sigma)} (w^l_{it})^{1-\sigma} \\ & + (1-\mu^j) z_{it}^{-\psi(1-\sigma)} \left[\lambda^j (r_{it})^{1-\eta} + (1-\lambda^j)(w^s_{it})^{1-\eta}\right]^{\frac{1-\sigma}{1-\eta}} \Big]^{1-\theta}\end{aligned} \tag{11}$$

2.5 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 $L_{it}^l + L_{it}^s$:

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

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

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

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.

2.6 Equilibrium

We define the dynamic equilibrium of the economy below.

Definition 1. Dynamic Equilibrium. Given initial conditions $\{L_{i0}^l, L_{i0}^s, k_{i0}\}$ in each location, the dynamic equilibrium contains a sequence of location-specific prices $\{w_{it}^l, w_{it}^s, r_{it}\}_{t=0}^{\infty}$, quantities $\{L_{it}^l, L_{it}^s, k_{it}\}_{t=0}^{\infty}$ and value functions $\{v_{it}^l, v_{it}^s\}_{t=0}^{\infty}$, 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 capital and population is characterized as in equations (3), (4), and (5).

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

$$w_{it}^{l} = \frac{\sum_{j=1}^{J} \phi_{it}^{lj} X_{it}^{j}}{L_{it}^{l}} \tag{12}$$

$$w_{it}^{s} = \frac{\sum_{j=1}^{J} \phi_{it}^{sj} X_{it}^{j}}{L_{it}^{s}} \tag{13}$$

$$r_{it} = \frac{\sum_{j=1}^{J} \phi_{it}^{kj} X_{it}^{j}}{k_{it}} \tag{14}$$

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

5. Trade balance condition holds in all locations:

$$X_{it}^{j} = \gamma_j \sum_{n=1}^{N} \pi_{ni,t} X_{nt} = \gamma_j \sum_{n=1}^{N} \pi_{ni,t} \sum_{s=1}^{J} X_{nt}^{s}, \tag{15}$$

where $\pi_{ni,t}$ denotes the trade share between origin i and destination n at time t defined in equation (B.3) in the Appendix B.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:

Definition 2. Steady State. A steady state of the economy is an equilibrium in which the endogenous variables are constant over time: $\{w_i^{l*}, w_i^{s*}, r_i^{*}, v_i^{l*}, v_i^{s*}, L_i^{l*}, L_i^{s*}, k_i^{*}\}$.

3 Quantification

This section presents the details regarding the quantification of the model. 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 one year, with the initial year in 2000. 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 in which we have access to all the needed data, as explained later. The prefectures in our sample are representative: 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 map the 82 industries observed in China's 2002 Industrial Classification for National Economic Activities into four broad sectors by skill intensities and tradability: the skill- and unskill-intensive manufacturing sectors and the skill- and unskill-intensive service sectors.[5] To estimate the skill intensity at the industry level, we follow Fan (2019) and use the income share of skilled workers in each industry from the *2005 One Percent Population Survey.* We rank industries by skill intensity separately for manufacturing and service sectors and then group the industries above the median skill intensity into the skilled sectors and those below into the unskilled sectors. Tables C.1 and C.2 in the Appendix provide the detailed mapping between industries and the four sectors in the paper.

3.1 Initial Conditions

Population The initial distribution of the population by location and skill type comes from the 2000 Census. We 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 in-

[5] Out of the 82 industries, 29 are manufacturing, and 53 are service industries

vestment 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 A.1 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, and the sectoral trade flow between China and each country for 2000-2006 from the World Input/Output Database (WIOD).

3.2 Geography

Trade Costs Products from the manufacturing sectors are tradable across locations, and those from the service industries are non-tradable. We use the estimated trade costs between Chinese prefectures from Ma and Tang (2024). The trade costs in that paper are based on freight infrastructure on road and rail networks each year during our sample period. Within China, trade costs do not vary across sectors.

We modify the methods in Ma and Tang (2024) to estimate the trade costs between Chinese prefectures and the ROW. Start with the 27 port cities in China identified in Ma and Tang (2024), we assume that all port cities face the same trade cost with the ROW in a given sector, denoted as "$\tau^j_{ROW,t}$", to be estimated later. Conditional 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 location i determined by the τ matrix within China. We allow the trade costs between China and ROW to 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:

$$\widehat{\tau^j_{\mathrm{ROW},t}} = \left(\frac{\widehat{S^j}_{(\mathrm{CN},\mathrm{ROW}),t} \times \widehat{S^j}_{(\mathrm{ROW},\mathrm{CN}),t}}{\widehat{S^j}_{(\mathrm{ROW},\mathrm{ROW}),t} \times \widehat{S^j}_{(\mathrm{CN},\mathrm{CN}),t}} \right)^{-\frac{1}{2\theta}}, \tag{16}$$

where $\widehat{S^j}_{(\cdot),t}$ is the changes in trade flow in sector j between year t and the initial year. With the trade elasticity parameter θ and the observed flows, we calculate the *changes* in trade costs for each sector year by year between 2000 and 2006. 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 precisely matching the observed trade costs in that year.[6] Our method allows us to match the trade shares between China and ROW observed along the transition path.[7] With the estimated $\tau^j_{ROW,t}$, we have complete trade costs matrices across all locations in all sample years.

Migration Costs Workers 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.

Our estimation procedure relies on two data sources: 1) the *2005 One Percent Population Survey*, and 2) the passenger transportation network from Ma and Tang (2024). The population survey allows us to compute the share of migrants in prefecture g with hukou from prefecture i for skill type d as a fraction of the population in location i, denoted as $\bar{D}^d_{gi,t}$. In Appendix C.4, we establish the following relation between $\bar{D}^d_{gi,t}$ and our model-predicted

[6]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.

[7]One can also invert the model along the transition path to match the trade shares to the data for all years after 2000. However, simulating the transition path is much more computationally expensive than solving the initial equilibrium. Nevertheless, we present the results using the inversion methods as a robustness check in Section 4.4.

migration share $D^d_{gi,t}$:

$$\bar{D}^d_{gi,t} = \frac{D^d_{gi,t}}{1 - D^d_{gg,t}},$$

where $D^d_{gi,t}$ is the model-consistent migration probability that depends on option values and the migration costs:

$$D^d_{gi,t} = \frac{\exp\left[(\beta v^d_{gt+1} - \kappa^d_{gi,t})/\rho\right]}{\sum_{n=1}^{N} \exp\left[(\beta v^d_{nt+1} - \kappa^d_{ni,t})/\rho\right]}.$$

To simplify notation, we drop the time subscript henceforth. Double differencing the relation above removes the option value from the expression and leads to:

$$\frac{\bar{D}^d_{gi}}{\bar{D}^d_{ii}} \frac{\bar{D}^d_{ig}}{\bar{D}^d_{gg}} = \frac{D^d_{gi}}{D^d_{ii}} \frac{D^d_{ig}}{D^d_{gg}} = \exp\left[-\frac{1}{\rho}(\kappa^d_{gi} + \kappa^d_{ig})\right]. \tag{17}$$

We further assume that the bilateral migration cost is the sum of the bilateral travel cost and the entry barrier of the destination location:

$$\kappa^d_{gi} = \kappa^d_g + \bar{\kappa}_{gi},$$

where $\bar{\kappa}_{gi} = \bar{\kappa}_{ig}$ is symmetric travel cost between location i and g that depends on the passenger travel infrastructure, and κ^d_g is type-specific entry barrier for entering location g. We interpret the entry barriers as policy restrictions such as the hukou registration. Conditional on the symmetric travel costs from Ma and Tang (2024), estimating migration costs is equivalent to estimating entry barriers for all locations. Taking stock, the estimation equation becomes:

$$\begin{aligned} \frac{\bar{D}^d_{gi}}{\bar{D}^d_{ii}} \frac{\bar{D}^d_{ig}}{\bar{D}^d_{gg}} &= \exp\left[-\frac{1}{\rho}(\kappa^d_g + \kappa^d_i + 2\bar{\kappa}_{gi})\right] \\ \frac{\bar{D}^d_{gi}}{\bar{D}^d_{ii}} \frac{\bar{D}^d_{ig}}{\bar{D}^d_{gg}} \exp\left(\frac{2\bar{\kappa}_{gi}}{\rho}\right) &= \exp\left[-\frac{1}{\rho}\left(\kappa^d_g + \kappa^d_i\right)\right]. \end{aligned} \tag{18}$$

Given travel costs and migration elasticity parameter ρ, we estimate entry barriers κ^d_g for each location and skill type using Poisson regression based upon equation (18).

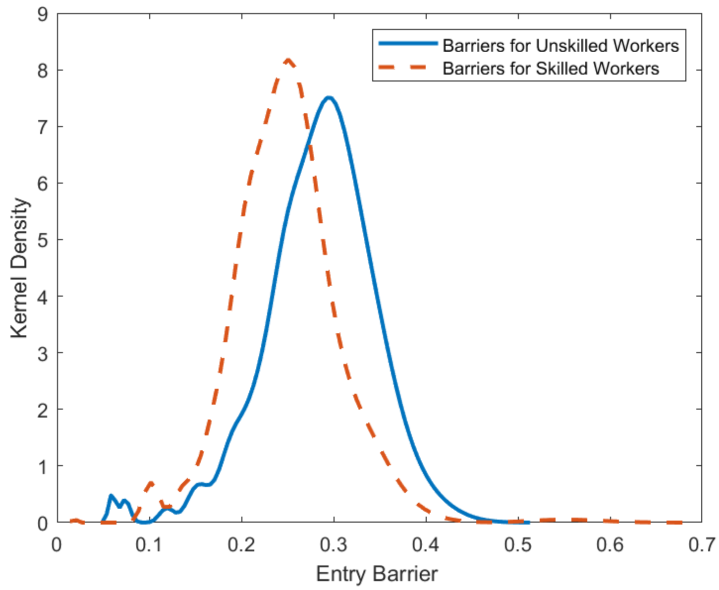


Figure 1: Distribution of Entry Barriers

Notes: This figure shows the histogram of the estimated entry barriers for unskilled and skilled workers. Entry barriers are estimated using PPML and normalized by the average lifetime utility of an unskilled worker.

Our estimation reveals that the migration frictions are substantial and, on average, higher for unskilled workers than for skilled ones. Figure 1 presents the histogram of entry barriers for unskilled and skilled workers across locations. In the figure, we normalize the entry barriers by the average lifetime utility of unskilled workers. The migration costs are formidable, equivalent to 28.7 percent for skilled workers or 42.1 percent for unskilled workers' lifetime utility. The higher migration costs unskilled workers face come from the discriminative hukou policy.

3.3 Parameterization

We discipline all the other parameters in one of the three ways. Some of the parameters were externally determined based on the estimates in the literature; Some parameters come from inverting the model in the initial static equilibrium. Lastly, the parameters affecting population distribution were calibrated along the transition path. In the rest of the section, we briefly discuss the quantification strategy of these parameters.

Pre-determined Parameters We choose trade elasticity $\theta = 5$ from Costinot and Rodríguez-Clare (2014). We assume an annual discount rate of $\beta = 0.97$, which is consistent with an annual interest rate of around 3 percent. We set the migration elasticity parameter $\rho = 3\beta$, 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 conventional values in the macroeconomic literature: we take the elasticities of substitution $\sigma = 1.67$ and $\eta = 0.67$ from Krusell et al. (2000). We choose skill-biased productivity parameter $\psi = 0.5$ following Burstein and Vogel (2017). Finally, according to the World Bank, the average annual mortality rate in China during 2000-2020 is 0.7 percent, suggesting an annual survival rate of $\xi = 0.993$.

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	0.97	-	Annual discount factor
θ	5	Costinot and Rodríguez-Clare (2014)	Trade elasticity
ρ	3β	Kleinman et al. (2023)	Inverse of migration elasticity
σ	1.67	Krusell et al. (2000)	EoS between l and h
η	0.67	Krusell et al. (2000)	EoS between s and k
ψ	0.5	Burstein and Vogel (2017)	Skill-biased productivity parameter
δ	0.1	Zhang et al. (2004)	Capital depreciation rate
τ_{ni}	-	Ma and Tang (2024)	Bilateral trade cost
γ^j	-	China 2002 IO table	Sectoral consumption share
ξ	0.993	World Bank	Annual 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 A subset of parameters, $\{\bar{z}_i, \mu^j, \lambda^j, \tau^j_{ROW,2000}\}$, are calibrated so that the initial static equilibrium matched the observed economic conditions in 2000. As is common in the dynamic spatial models, we do not need to assume that the model in 2000 is in a steady state. Instead, we only need to assume that the initial static equilibrium is on a transition path toward a future steady state.

The exogenous component of prefecture-level fundamental productivity, $\bar{z}_i$, 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 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 U.S. Input-Output Table in 2007 was used to obtain capital income shares.

Table 2 shows the calibrated results of weights on unskilled workers and capital in production function for China and the ROW. Unsurprisingly, unskilled sectors put more weight on unskilled workers than skilled workers. Moreover, in China, capital takes up higher weights in unskilled manufacturing sectors ($\lambda^j = 0.91$) than in skilled ones ($\lambda^j = 0.86$). This pattern reflects the fact that in the 2000s, capital-intensive industries in China, such as primary metal, were also more reliant on unskilled workers than skilled ones. On the contrary, the skilled sector is more capital-intensive than the unskilled sector in the ROW. These estimation results subsequently imply the pattern of comparative advantage in the quantitative analysis presented later. Considering that 1) the ROW is relatively more abundant in skilled workers and capital in the data, and 2) the skilled sector is capital-intensive in the ROW's production function, as per our estimation, the ROW specializes in the skilled sector when trading with China.

Lastly, as discussed above, we invert the model at the initial static equilibrium to back out the initial trade costs between ROW and Chinese port prefectures in levels, $\tau^j_{ROW,2000}$.

Amenities and Skill Upgrading Costs The last group of parameters is calibrated on the transition path, conditional on the abovementioned parameters. These parameters are the skill upgrading cost $\{\kappa^l_s\}$ and location-specific amenities $\{\bar{b}_i\}$. Specifically, $\{\kappa^l_s\}$ is chosen to match the aggregate skill ratio of 0.36 in the year 2010, as indicated by the population Census in China that year. Our calibrated skill upgrading cost is 48 percent of the average

Table 2: Calibrated Production Function

Panel (a): China

Weights	Unskilled Manu.	Skilled Manu.	Unskilled Service	Skilled Service
μ^j	0.33	0.17	0.27	0.04
λ^j	0.91	0.86	0.85	0.75

Panel (b): ROW

Weights	Unskilled Manu.	Skilled Manu.	Unskilled Service	Skilled Service
μ^j	0.29	0.16	0.27	0.07
λ^j	0.80	0.89	0.82	0.89

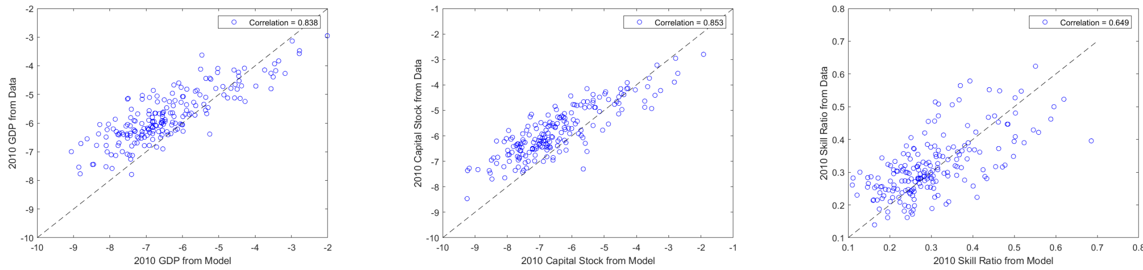
Notes: This table reports the results of production weights in four sectors for China and the ROW. The weights are calibrated in the initial static equilibrium by targeting sector-level factor income shares. $\mu \in [0, 1]$ is the weight on unskilled workers and $\lambda \in [0, 1]$ is the weight on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

lifetime utility among unskilled workers in the initial period. The skill-upgrading costs are slightly higher than the average migration costs among unskilled workers at 42 percent of lifetime utility. The high upgrading costs reflect 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.

Model Fit The quantification strategy described above aligns reasonably well with the untargeted data moments. Figure 2 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 2: 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 Analysis

This section discusses quantitative results. We start by assessing the basic patterns of skill premium in the baseline model and then move to counterfactual results. Our main counterfactual exercise reverses the trade liberalization after China joined WTO by assuming that the trade barriers between China and the ROW stayed at the same level in 2000. We infer the spatial impact of trade liberalization by comparing the counterfactual results to the baseline and show that the spatial impacts depend on the time horizon of the analysis.

To understand the source behind the time variation of the spatial impacts, we implement two other sets of counterfactual simulations in which we assume away capital accumulation or skill formation. In this "no capital accumulation" simulation, 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. We adopt an infinite skill upgrading cost in the "no skill upgrading" counterfactual so that no unskilled workers choose to upgrade their skills.

4.1 Benchmark Results

Before presenting the counterfactual simulations, we first discuss two features of the baseline transition path that highlight the key mechanisms that drive the time trends and spatial distribution of skill premium in the model.

The first feature is that capital accumulation and skill upgrading drive the aggregate skill premium in opposite directions along the transition path. Figure 3 presents the evolution of skill premium predicted by the model in the baseline case and two counterfactual cases without capital accumulation or skill upgrading. Relative to the baseline case, shutting down capital accumulation leads to a much lower skill premium due to capital-skill complementarity, as seen in the red dashed line. On the other hand, shutting down skill upgrading significantly increases the overall skill premium since skilled workers are in short supply, as shown in the yellow dotted line. The realized aggregate skill premium, shown as the solid blue line in the middle, results from the trade-offs between the two counteracting forces.

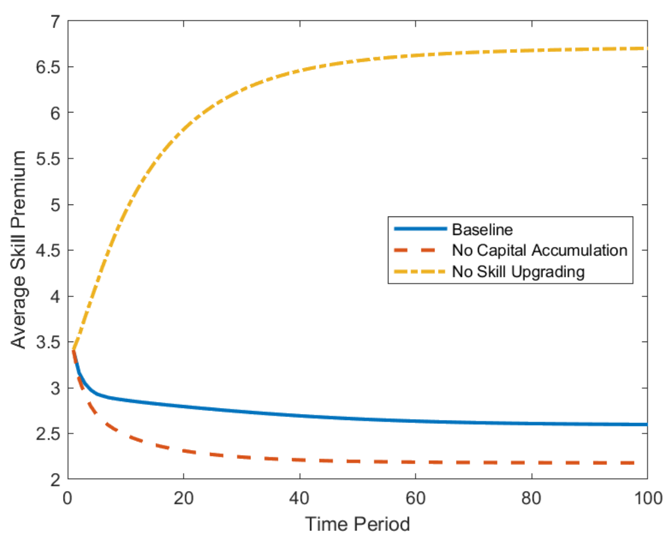
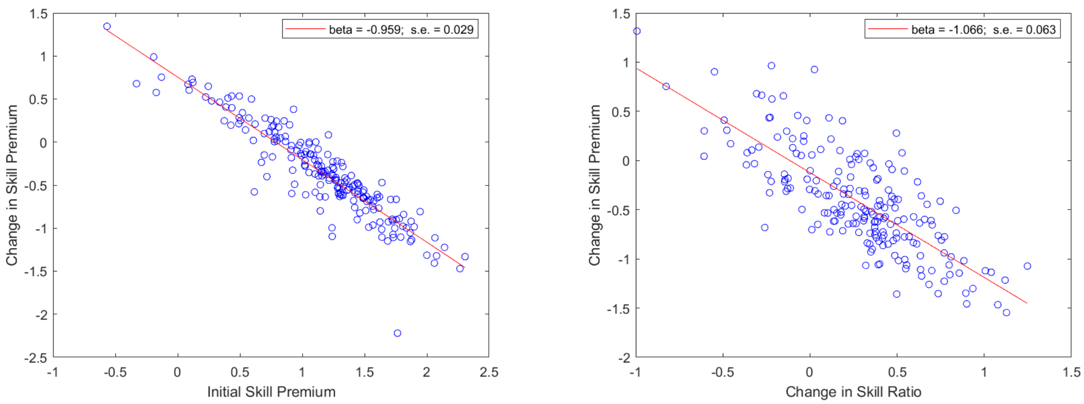


Figure 3: Average Skill Premium over Time

Notes: This figure shows the time path of population-weighted average skill premium over the transition path. The solid blue line is based on baseline simulation, the dashed red line is based on simulation without capital accumulation, and the dotted yellow line is based on a counterfactual simulation without skill upgrading.

The second feature of the baseline simulation is the spatial convergence of skill premiums. Panel (a) in Figure 4 presents the β-convergence graph of skill premium across 196 prefectures by plotting the logarithm of changes in skill premiums between the initial period and the steady state against the logarithm of the initial levels. The figure suggests a strong convergence in skill premiums, as skill premiums grow faster in locations with initially lower skill premiums. The growth rate of skill premiums in those locations with initially high skill premiums is even negative. The β convergence coefficient is also substantial and significantly negative at -0.96.



(a) β-Convergence of Skill Premium

(b) Skill Premium and Skill Ratio

Figure 4: Convergence of Skill Premium Across Space

Notes: The left panel shows the β-convergence of skill premium across prefectures between the initial and the steady state. The right panel shows the change in skill premium against the change in skill ratio in the steady state. All variables are in logarithm. Each dot represents a prefecture. The straight lines are the best linear fits.

The spatial convergence of skill premium is driven by internal migration. Panel (b) in Figure 4 shows a strong negative relationship between the changes in skill ratio and skill premium: cities with declining skill premiums also have a large inflow of skilled workers over time. Not surprisingly, these locations also have the highest initial skill premiums, thus attracting more skilled workers and increasing the relative supply of skills. This spatial movement of skilled workers over time reduces the large disparity of spatial skill premiums.

4.2 The Spatial Impacts of Trade

In the rest of this section, 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.

Aggregate Impacts The aggregate welfare gain of WTO accession is 0.49 percent.[8] The unskilled workers gain slightly more at 0.53 percent, while the skilled workers gain 0.34 percent. The higher gains accrued to the unskilled workers come from China having a comparative advantage in the unskilled sector – an expected prediction from the Stopler-Samuelson effect. Furthermore, distance to the coast is a strong predictor for the regional welfare gains: those with below-median distance to the ROW gain 0.47 percent, while those with above-median distance only gain 0.39 percent in welfare.

Trade liberalization also reduces the aggregate skill premium in the long run. At the steady state, the skill premium on average decreases by 0.17 percent due to trade liberalization. The decline is again driven by the Stoper-Samuelson effect, as China holds a comparative advantage in the unskilled-intensive sectors. Capital accumulation helps to alleviate the decline due to capital-skill complementarity: the skill premium would fall by 0.27 percent without capital accumulation. Endogenous skill upgrading also alleviates the decline in the skill premium. As the trade shocks benefit the unskilled manufacturing sector in China, skill upgrading is thus less favorable, and therefore, the unskilled workers delay the upgrading decision for unskilled workers. In the equilibrium, the relatively reduced supply of skilled workers resulted in a smaller decrease in skill premiums.

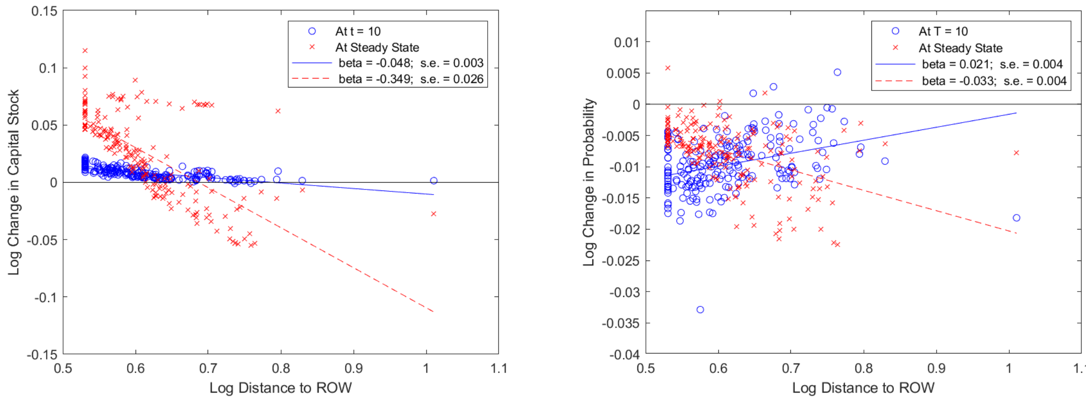
[8] We use population-weighted changes in v_{it}^d in the initial period to measure the welfare impact. This measures the discounted present value of the gains from trade upon impact. Our welfare gain is consistent with quantitative trade and urban literature estimates. For example, in Tombe and Zhu (2019), the gains from external trade costs without intermediate inputs are 0.3 percent. In Fan et al. (2021), the Chinese welfare gain is 1.5 percent.

4.2.1 The Spatial Impacts of Trade along the Transition

The spatial impacts of trade vary along the transition path. In the following, we first show its impacts on capital accumulation and skill acquisition in each prefecture. We then move to trade's impact on population movements, skill ratio, and skill premium between the short and long run.

Capial Accumulation Trade liberalization induces faster capital accumulation in the coastal prefectures than in the inland ones, and the effects are stronger in the long run. Panel (a) in Figure 5 plots the impacts of trade on capital stock growth in each prefecture against its distance to the ROW at various time horizons. The blue dots indicate the changes at $t = 10$, and the red stars represent the changes at the steady state. Regardless of the time horizon, prefectures closer to the world market accumulate capital stock faster, as indicated by the negative slopes in both the short- and the long-run. The slope of the linear fits, which we denote as $\zeta^k(t)$, adopts a natural interpretation as the **distance elasticities of capital growth** at time t. This elasticity provides an intuitive measure of the relative advantage of the coastal locations in capital accumulation: on average, changing the distance to the ROW by 1 percent changes its capital growth rate by $\zeta^k(t)$ percent at period t. At $t = 10$, the distance elasticity equals -0.048. In the steady state, the distance elasticity increases by 6 folds to -0.349 in absolute values. This effect is intuitive: locations closer to the world market are better positioned to benefit from trade liberalization. The positive demand shock, in turn, leads to higher returns to capital and thus faster accumulation, as indicated in Equation (5).

Skill Upgrading Similarly, the short and long-run impacts of trade on skill acquisition are drastically different. Panel (b) in Figure 5 illustrates the impact of trade on skill upgrading decisions in a similar manner as compared to Panel (a) discussed above. Instead of plotting the capital growth rate, Panel (b) plots the change in the fraction of the population who upgrade their skills against the distance to the ROW for each prefecture. In the short run (blue circles), unskilled individuals in coastal cities are *less likely* to acquire skills than inland cities, as suggested by the positive distance elasticity in the figure. The reluctance in



Capital Accumulation　　　　Skill Acquisition

Figure 5: Impacts of Trade on Factor Accumulation

Notes: These figures show the effects of trade liberalization on capital accumulation and skill acquisition across Chinese prefectures in the short run and long run from baseline simulation. Each dot represents a prefecture. For each prefecture i, we measure skill acquisition using the "skill upgrading probability", which is defined as the ratio of total number of skilled workers in the next period originating from i over location i's total number of unskilled workers in the current period. The blue dots come from the cross section at period 10, and the red dots are from the steady state. The straight lines are linear fits.

skill acquisition is expected: comparative advantage in unskilled-intensive industries rewards unskilled workers and reduces the value of upgrading. The delay is also consistent with the empirical pattern documented in Li (2018).

Over time, however, the trade-induced capital accumulation in coastal locations reverses the pattern: we instead observe a negative distance elasticity (-0.033) at the steady state. As higher capital stock increases the relative marginal product of skilled workers, the coastal locations start to gain the comparative advantage in the capital- and skill-intensive industries. Ultimately, the local unskilled workers in these locations, attracted by the higher skill premium there, are *more likely* to upgrade than their counterparts in the inland areas.

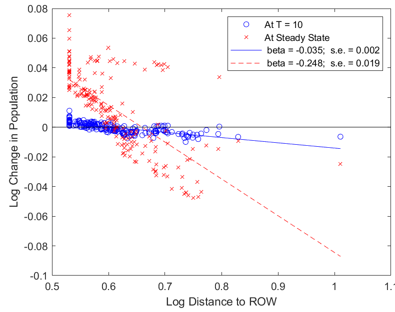
The differential impacts of trade on capital accumulation and skill acquisition are the fundamental driving forces behind the spatial impacts of trade that we will discuss in the rest of the section. As the coastal locations stand to accumulate capital stock faster, in the long run, these locations start to gain comparative advantage in capital-intensive industries and become attractive to skilled workers; the changes in comparative advantage, in turn, drive the gap between the short and the long-run response of migration and skill premium to the trade shocks that we will discuss in detail later.

Population We start with the impacts of trade on internal migration. As well documented in the literature (Tombe and Zhu, 2019; Fan, 2019; Ma and Tang, 2020), trade liberalization drives the population movement toward the coastal locations. We find similar impacts of trade on internal migration. Panel (a) in Figure 6 plots the impacts of trade on population changes against the distance to ROW for each prefecture. Both in the short and the long run, coastal cities attract populations from other inland ones after the trade liberalization. As measured by the slope of the linear fit, the **distance elasticity of population changes**, denoted as $\zeta^L(t)$, is -0.035 when $t = 10$.

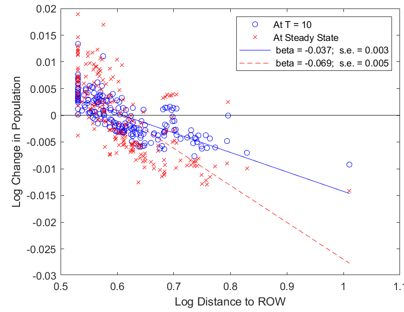
The long-run impacts of trade on population concentration are much stronger: the absolute value of the distance elasticity of population change increases from -0.035 to -0.248 at the steady state. In other words, researchers would underestimate the impacts of international trade on population movement by a factor of $0.248/0.035 \approx 7$, or $0.248 - 0.035 = 21.3$ percentage points, if they were to rely on a static model that does not differentiate across time horizons. Later in this section, we will discuss that from a policy designing perspective, ignoring the variations along the transition path would similarly lead to inefficient and underfunded policies in the long run.

The gaps between trade's long- and short-run impacts on population are mainly driven by capital formation. To shed light on the mechanism, we present two other sets of results in Panels (b) and (c) of Figure 6. In Panel (b), we show the results from counterfactual simulations without capital accumulation, and in Panel (c), without skill upgrading. In a world without capital accumulation, the gap between the long- and the short-run elasticity shrinks considerably to $0.069 - 0.037 = 3.2$ percentage points. By this measure, capital accumulation is responsible for $1 - \frac{0.069-0.037}{0.248-0.035} = 85$ percent of the observed gap between the long- and short-run distance elasticity.

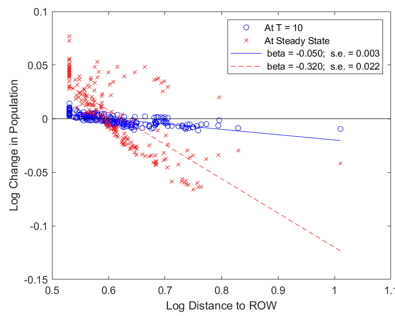
On the other hand, shutting down skill upgrading would lead to an even higher long- and short-run gap in the distance elasticity. As shown in Panel (c) of the same figure, without skill upgrading, the gap in distance elasticity is $0.32 - 0.05 = 27$ percentage points, higher than the gap of 21.3 percentage points in the baseline model. The effect of skill upgrading is not



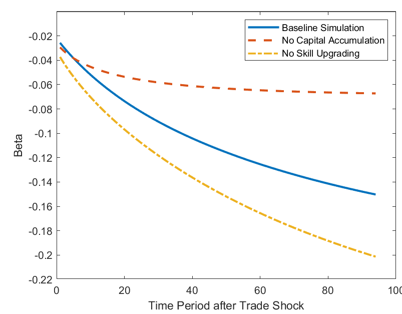
(a) Baseline



(b) No Capital Accumulation



(c) No Skill Upgrading



(d) Distance Elasticity of Population Change, $\zeta^L(t)$

Figure 6: Impacts of Trade on Population

Notes: These figures show the effects of trade liberalization on population changes across Chinese prefectures. Panels (a) to (c) show the trade-induced population changes in the baseline model, the model without capital accumulation, and the model without skill upgrading, respectively. Each dot represents a prefecture. The blue dots come from the cross section at period 10, and the red dots are from the steady state. The straight lines are linear fits. Panel (d) shows the distance elasticity of population change, $\zeta^L(t)$, as a function of time.

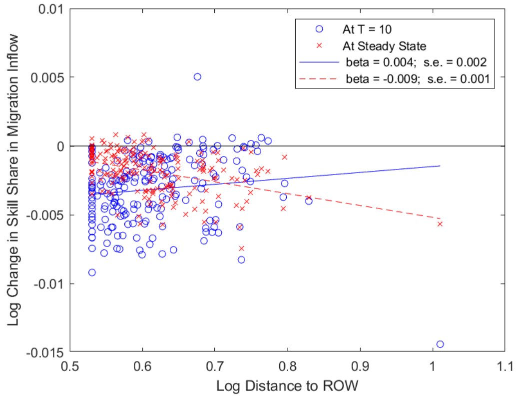
surprising. While capital accumulation increases the relative demand for skilled workers in coastal locations, skill upgrading offsets the effect by increasing the supply of skilled workers in these locations. In a model with capital accumulation but no skill upgrading, the relative abundance of coastal capital stock would increase the attractiveness of these locations to the skilled workers and eventually lead to a higher concentration of coastal population.

Lastly, Panel (d) in Figure 6 summarizes these results and presents $\zeta^L(t)$ as a function of time. As shown in this figure, the distance elasticity of population movement is always negative, indicating that trade always induces internal migration towards the coast. However, the $\zeta^L(t)$ is not constant over time, and the role of distance strengthens as time goes by. In the model without capital accumulation (the red dashed line), the $\zeta^L(t)$ is flatter, highlighting the role of capital stock in driving the temporal changes in $\zeta^L(t)$. On the other hand, in the model without skill upgrading (the yellow dotted line), the long- and short-run gap in distance elasticity is even more pronounced.

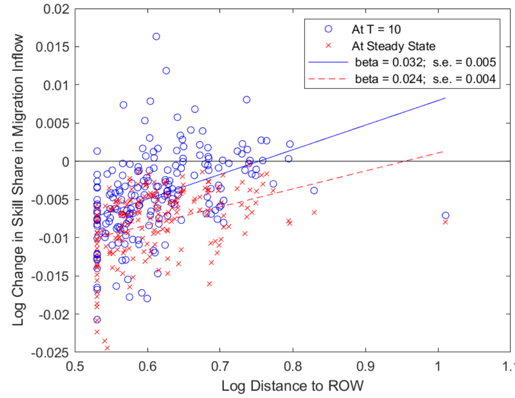
Skill ratio The gaps between the long- and the short-run impacts of trade on the skill composition of migrants are even more striking. Similar to Figure 6, Figure 7 shows how each location's inflow skill ratio, defined as the fraction of skilled workers in the inflow population in each prefecture, responds to the trade liberalization in both the short run and long run. As shown in Panel (a), the distance elasticity of the skill ratio, $\zeta^{\mathrm{sr}}(t)$, is positive at 0.004 in the short run but turns negative at -0.009 in the long run. In other words, unskilled workers are more likely to migrate towards coastal locations in the short run, whereas skilled populations are more inclined to do so in the long run.[9]

How skilled and unskilled workers migrate in response to trade depends on two counteracting forces in our context. On the one hand, comparative advantage in unskilled-intensive industries tends to attract unskilled workers to the coastal locations to exploit the export boom. On the other hand, capital-skill complementarity draws skilled workers to locations that enjoy an abundance of capital stock. In the short run, the impacts of initial conditions dominate, resulting in a positive distance elasticity. However, in the long run, the advantage of coastal locations in capital accumulation takes precedence, leading to the observed changes in the skill composition of the inflow population. Panels (b) and (c) further highlight these points by shutting down capital accumulation and skill upgrading separately. Without capital accumulation, the effects of comparative advantage always dominate, and therefore $\zeta^{\mathrm{sr}}(t)$ is always positive, as seen in Panel (b) and the red dashed line in Panel (d) of the same figure. In this case, the unskilled workers are always drawn to the coastal locations due to China's comparative advantage in the unskilled-intensive sectors. Without skill upgrading, the migration flow towards coastal locations is even more dominated by skilled workers, with a $\zeta^{\mathrm{sr}}(t) = -0.047$. This higher distance elasticity is attributed to coastal locations relying more on internal migration to meet the increased demand for skilled workers resulting from capital accumulation, especially when local unskilled workers cannot upgrade their skills.

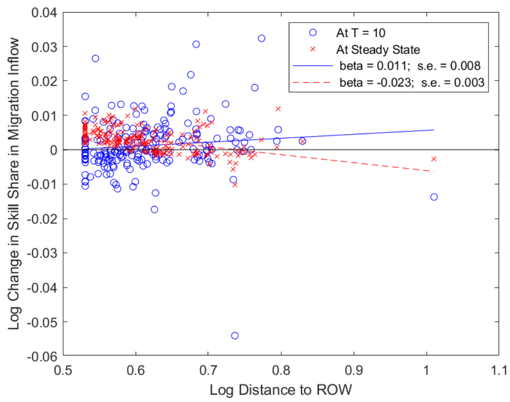
[9] Figure A.1 in the Appendix presents the skill ratio in the total workforce, instead of inflow population, within a prefecture and the results are similar.



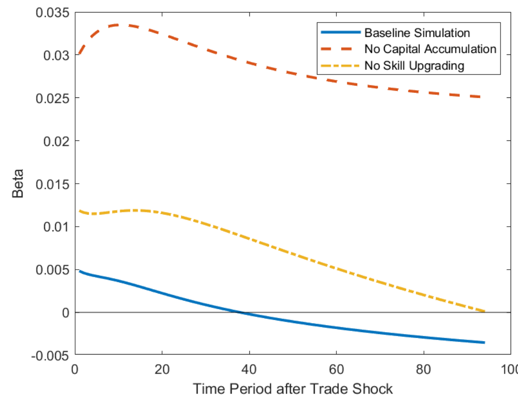
(a) Baseline



(b) No Capital Accumulation



(c) No Skill Upgrading



(d) Distance Elasticity of Skill Ratio, $\zeta^{\mathrm{sr}}(t)$

Figure 7: Impacts of Trade on Skill Ratio in Migration Inflow

Notes: This figure shows the effects of trade liberalization on the skill ratio in migration inflow for each Chinese prefecture. The skill ratio in migration inflow for each prefecture is calculated as the ratio of skilled workers inflow to total workforce inflow, excluding those upgrading their skill levels. Panels (a) to (c) show the trade-induced skill ratio changes in the baseline model, the model without capital accumulation, and the model without skill upgrading, respectively. Each dot represents a prefecture. The blue dots come from the cross section at period 10, and the red dots are from the steady state. The straight lines are linear fits. Panel (d) shows the distance elasticity of the inflow skill ratio, $\zeta^{\mathrm{sr}}(t)$, as a function of time.

Skill Premium The spatial impacts of trade on the skill premium follow a similar pattern as discussed above. In the short run, trade reduces skill premiums more in coastal locations due to the Stopler-Samuelson effects. However, in the long run, skill premiums in coastal locations tend to increase instead.

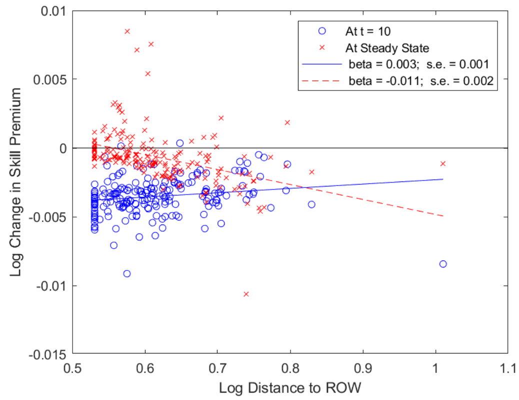
Figure 8 presents the impacts of trade on skill premium similarly to the previous ones. In the short run, the changes in skill premium are negative in all cities, a finding consistent with the Stoper-Samuelson effect. The distance elasticity of skill premium, $\zeta^{\mathrm{sp}}(t)$, is positive, suggesting coastal cities experience the largest decline in skill premium due to trade shocks. However, in the long run, skill premiums rise in many coastal cities, and the distance elasticity becomes negative: locations closer to the world market enjoy a higher trade-induced skill premium. It's important to note that the long-run impacts of trade on spatial skill premiums align with the predictions of the Stopler-Samuelson effect as well. As coastal locations accumulate capital stock and attract skilled migrants, they progressively become capital- and skill-abundant. Consequently, in the long run, the abundant factor – the skilled workers – gains from trade, a prediction that remains consistent with the Stopler-Samuelson theorem.

Similar to the previous findings, without capital accumulation, the coastal locations would have retained a comparative advantage in the unskilled-intensive industries, and the distance elasticity of skill premium would always be positive, as seen in Panel (b) and the red dashed line in Panel (d) of the same figure. In the absence of skill upgrading, the reversal of distance elasticity would be solely driven by capital accumulation. Nevertheless, the skill premium would consistently decline due to trade in this scenario, as China would continue to be an unskilled-abundant economy.

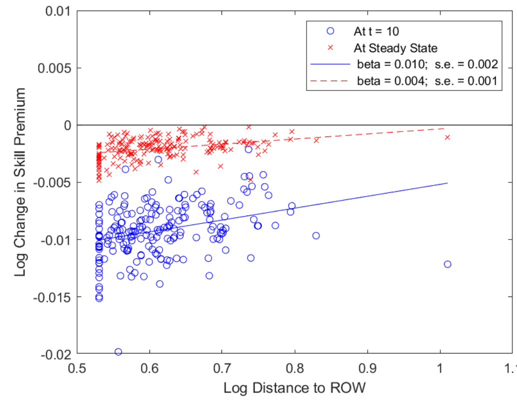
4.3 Place-Based Policies

As the spatial impacts of trade vary over time, policymakers should also take into account this intertemporal variation when designing policies to alleviate the negative effects of trade within a country. To emphasize this point, we conduct a hypothetical policy experiment demonstrating that myopic policies would be ineffective in the long run.

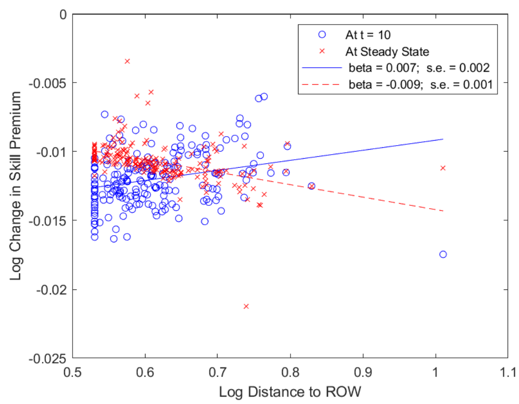
Our policy experiment subsidizes workers in less-developed locations in order to prevent emigration. This hypothetical policy resembles real-world place-based policies such as China's "Western Development Policies" or the "Local Revitalization Policies" in Japan. In the context of China, recall that in the baseline exercise, international trade draws the population away from inland prefectures toward the coastal ones. The population loss in inland areas could hinder economic development through various channels, including agglomeration externality, home market size, and tax base effects. This trade-induced population movement may further worsen the already increasing spatial inequality. From the central government's



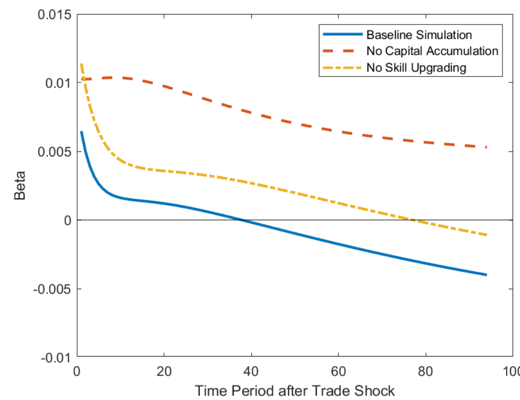
(a) Baseline



(b) No Capital Accumulation



(c) No Skill Upgrading



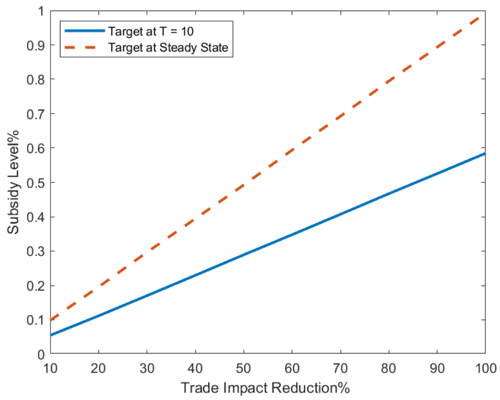
(d) Distance Elasticity of Skill Premium, $\zeta^{\mathrm{sp}}(t)$

Figure 8: Impacts of Trade on Skill Premium

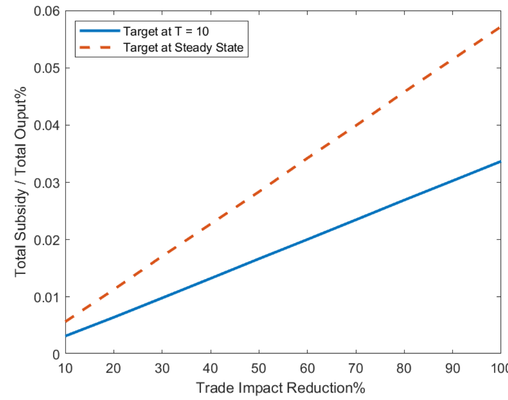
Notes: This figure shows the effects of trade liberalization on skill premium across Chinese prefectures. Panel (a)-(c) each shows the trade-induced skill premium changes in the baseline model, in the model without capital accumulation, and in the model without skill upgrading, respectively. Panel (d) shows the distance elasticity of skill premium change over time. Each dot represents a prefecture. The blue dots come from the cross section at period 10, and the red dots are from the steady state. The straight lines are linear fits. Panel (d) shows the distance elasticity of skill premium, $\zeta^{\mathrm{sp}}(t)$, as a function of time.

perspective, implementing a place-based policy that hedges against trade-induced migration could be desirable. In our hypothetical policy experiment, we implement a proportional wage subsidy of x percent for all workers in the targeted provinces, aiming to reverse the population loss caused by trade partially.[10]

[10] The targeted provinces in the hypothetical policy are the same as in the "Western Development Policy", which includes Inner Mongolia, Shannxi, Ningxia, Gansu, Xinjiang, Qinghai, Chongqing, Sichuan, Guizhou, Yunnan, and Guangxi. We broadly interpret the fiscal transfers implemented in the real world as wage subsidies that guide workers to these targeted locations. We avoid implementing a tax to fund these policies and assume that the funding comes from government reserves. This assumption is for the sake of simplicity



(a) Subsidy Level



(b) Total Amount of Subsidy as a Share of GDP

Figure 9: Wage Subsidy to Reduce Trade-Induced Emigration

Notes: This figure shows the cost of wage subsidies to reduce trade-induced population loss in the western and central provinces. The blue line shows the subsidy required for reducing population loss in period ten, and the red line shows that in the steady state. Subsidy level is the percentage increase in all workers' wages in selected western prefectures; total subsidy / total output is defined as the total amount of annual subsidy spent by the government as a share of total output in period seven when the trade shock ends.

A myopic subsidy that only aims to reverse the short-run population movement would be gravely underfunded in the long run. For example, we find that in order to revert 50 percent of the trade-induced population loss in the less-developed provinces in the short-run at $t = 10$, the central government would need to subsidize the income of all the workers in these provinces by 0.3 percent each year, which costs around 0.017 percent of the GDP on average. However, given the growing attraction of coastal locations, the subsidy policy would be insufficient to keep the population in the inland locations. To achieve long-run population retention at a steady state, the subsidy needs to be 65 percent higher at 0.028 percent of GDP each year. Figure 9 shows that the subsidy policies persistently differ over time across different policy targets. If the government aims to reverse the trade-induced migration by 20 percent, the long-term policy would cost 77 percent more than the one aimed in the short term. Conversely, if the goal is to revert trade-induced migration completely, the short-term policy would be underfunded by 70 percent in the long term.

Other place-based policies would suffer similar issues if the policymaker overlooked the

and is innocuous. As demonstrated later in this section, the subsidy policies cost, at most, around 0.03% of GDP.

intertemporal variation of spatial responses to trade shocks. For instance, our baseline results indicate that trade diminishes the incentives for landlords to invest in inland locations. If the government intends to counteract the negative impacts of trade on capital accumulation in inland provinces, a myopic policymaker might take no action, as the trade shock actually encourages mild capital accumulation in the short run. However, a government aiming for a long-run steady state would implement a 0.4 percent subsidy, incurring an average cost equivalent to around 0.028 percent of annual GDP. Likewise, a policy promoting skill acquisition must adjust its targeted locations over time. In the short run, the policy should invest more in coastal locations, as unskilled workers might hesitate to upgrade due to the high demand for unskilled worker. However, in the long run, such policies should shift toward inland locations because of the lack of capital accumulation and the emigration of skilled workers in the inland prefectures.

4.4 Robustness Check

This section summarizes the robustness of our results with respect to various model specifications and quantification. Table 3 summarizes these results. Panel (a) of the table replicates the baseline results discussed above for easy reference. Panels (b) to (e) report the robustness results discussed in this section.

Cross-location investment In the baseline economy, landlords are limited to investing only in their own locations and cannot access capital markets outside their hometowns. Although this assumption simplifies the model solution, it magnifies the role of the initial capital stock. In principle, landlords in remote locations could invest in the thriving coastal areas and directly benefit from the trade shock. In equilibrium, the returns from investments in coastal locations could contribute to economic growth in the inland areas, mitigating the impact of trade on spatial inequality.

In the first robustness check, we allow the landlords to invest freely in all the locations and show that our main results still hold. To achieve this, we modify landlords' investment decisions following Kleinman et al. (2023). In each period, the landlords make cross-investment decisions to allocate their own capital to different locations. We assume that capital is

Table 3: Distance Elasticities under Different Robustness Checks

	Full Model			W/o Capital Accu.			W/o Skill Upgrading		
	SR	LR	Δ	SR	LR	Δ	SR	LR	Δ
(a) Baseline									
Pop	-0.035	-0.248	-0.212	-0.037	-0.069	-0.033	-0.050	-0.320	-0.270
Skill Rat.	0.004	-0.009	-0.013	0.032	0.024	-0.009	0.011	-0.023	-0.035
Skill Pre.	0.003	-0.011	-0.014	0.010	0.004	-0.006	0.007	-0.009	-0.016
(b) Cross-Location Investment									
Pop	-0.010	-0.045	-0.035	-0.009	-0.013	-0.003	-0.011	-0.043	-0.032
Skill Rat.	-0.000	-0.007	-0.006	0.003	0.000	-0.002	0.003	-0.006	-0.009
Skill Pre.	-0.001	-0.003	-0.003	0.004	0.001	-0.004	0.000	-0.002	-0.002
(c) Calibrated Trade Cost									
Pop	-0.053	-0.432	-0.379	-0.058	-0.123	-0.065	-0.074	-0.540	-0.466
Skill Rat.	0.010	-0.016	-0.026	0.054	0.039	-0.015	0.022	-0.041	-0.063
Skill Pre.	0.011	-0.019	-0.031	0.018	0.007	-0.010	0.019	-0.015	-0.034
(d) No Skill-Productivity Complementarity									
Pop	-0.027	-0.141	-0.114	-0.027	-0.050	-0.022	-0.041	-0.196	-0.155
Skill Rat.	0.002	-0.002	-0.004	0.028	0.019	-0.009	0.014	-0.006	-0.020
Skill Pre.	0.003	-0.004	-0.007	0.010	0.005	-0.005	0.008	-0.004	-0.012
(e) Same Technology									
Pop	-0.040	-0.279	-0.238	-0.035	-0.069	-0.034	-0.046	-0.314	-0.268
Skill Rat.	-0.013	-0.021	-0.007	-0.001	-0.000	0.001	-0.022	-0.042	-0.020
Skill Pre.	-0.008	-0.017	-0.009	-0.003	-0.000	0.003	-0.010	-0.012	-0.001

Notes: This table reports the distance elasticity of population, skill ratio, and skill premium under different robustness checks. Short-run elasticities are from prefecture-level cross-section regressions at period 10, and long-run elasticities are from those at steady state. The column "△" reports the changes in elasticities from the short run to the long run.

freely mobile, and the cross-location investment is subject to a location-specific idiosyncratic investment shock. Appendix D.1 fully describes the extended model.

We recompute the distance elasticities of the population, skill composition of the migration flow, and skill premium over time for the extended model. Panel (b) in Table 3 presents the results. The distance elasticity of population in absolute terms increases from 0.01 in the short run to 0.05 in the long run, while it barely changes in the absence of capital accumulation. Like before, trade liberalization, in the long run, induces faster capital accumulation in coastal cities and improves their attractiveness for workers. The distance elasticities of skill migration and skill premium again are negative and increase in magnitude from the short run to the long run, suggesting the opposite impacts of trade liberalization between the short and long run. This pattern is, once again, attributed to capital accumulation. Notably, when capital accumulation is excluded, these elasticities show minimal changes over time and are even insignificant in the long run.

It is worth noting that when cross-location investment is allowed, the long-run elasticities become smaller in magnitude than benchmark results. For example, the long-run distance elasticity of population is 0.05 with cross-location investment, contrasting with the benchmark result of 0.25. This result is driven by the fact that cross-location investment induces a more even spatial distribution of capital demand. Comparing the spatial distribution of capital stocks in the baseline against that in the cross-location investment setup reveals that locations with relatively larger capital stocks in the baseline now experience lower capital stocks (capital inflows), while those with smaller capital stocks in the baseline now have larger capital inflows. Consequently, cross-location investment weakens the advantage of coastal locations and counteracts the positive effect of capital accumulation led by trade liberalization. Nevertheless, our main results remain robust.

Calibrated Trade Costs In the baseline results, we used the methods of Head and Ries (2001) to infer the trade costs between ROW and China. As a robustness check, we adopt an alternative quantification strategy and calibrate the trade costs along the transition path and the amenity parameters. Unlike the baseline approach, this alternative approach allows us to match the observed trade shares in the data exactly. Appendix D.2 provides more details on how to invert the model to recover the unobserved trade costs along the transition path.

We conduct the same counterfactual analysis as before to quantify the impacts of trade with or without capital accumulation. Panel (c) in Table 3 shows the results. From the short to long run, we again find a large increase in distance elasticity of population (from -0.05 to -0.43) and reversed signs of elasticities of skill migration (from 0.01 to -0.02) and skill premium (from 0.01 to -0.02). As suggested in the same panel, capital accumulation drives these large changes over time.

No skill-productivity complementarity In the baseline model, we introduce skill-productivity complementarity through the parameter ψ: productive locations are more complementary in using skilled workers. This parameter helps align the spatial distribution of skill premium with the data. However, there may be concerns about the extent to which this parameter influences our main results. Given that many coastal cities in China have relatively higher productivity, they are more attractive for skilled workers due to skilled-biased productivity. To address this concern, we repeat our main analysis assuming $\psi = 0$, therefore shutting down the complementarity between skill and productivity.

Table 3 Panel (d) shows the simulation results. While taking away skill-productivity complementarity weakens the role of capital accumulation, the distance elasticities are still statistically significant, and their changes over time are consistent with our benchmark results.

Same technology In the baseline economy, we allow production technologies to differ between China and the ROW. Consequently, both Ricardian and Heckscher-Ohlin forces are at play in determining the pattern of comparative advantage. To isolate one force from the other, in this section, we equalize China's and the ROW's production functions, which are determined by the income shares of each factor from China's data. By doing so, the trade pattern is solely determined by Heckscher-Ohlin forces through relative factor endowments. We recalibrate the production function along with other parameters in equilibrium. Panel (e) in Table 3 presents the results, and once again, we observe a persistent gap between the short and long-run distance elasticities.

5 Conclusion

In this study, we develop a dynamic spatial framework to understand better how trade liberalization's spatial impact changes over time. 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 the spatial impacts of trade crucially depend on the time horizon. In the short run, the spatial impacts are primarily driven by initial conditions. However, in the long run, the factor endowments in each location become functions of the trade shocks themselves. Consequently, the long-term spatial impacts of trade could differ significantly from the short-term effects. Understanding the intertemporal variations in spatial impacts is crucial for policy design. Policies directing migration flow could be underfunded if one overlooks the time variation, and policies aimed at encouraging skill acquisition could be misdirected if one ignores the long-term changes.

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.

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.

_ , _ , _ , and Kaveh Majlesi, "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *American Economic Review*, October 2020, *110* (10), 3139–83.

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.

Burstein, Ariel and Jonathan Vogel, "International trade, technology, and the skill premium," *Journal of Political Economy*, 2017, *125* (5), 1356–1412.

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.

Chang, Pao-Li, Yi-Fan Chen, Wen-Tai Hsu, and Xin Yi, "Labor Market Implications of Taiwan's Accession to the WTO: A Dynamic Quantitative Analysis," Technical Report, Singapore Management University, School of Economics 2022.

Colantone, Italo and Piero Stanig, "The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe," *American Journal of Political Science*, 2018, *62* (4), 936–953.

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.

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.

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

Fan, Haichao, Xiang Gao, and Lina Zhang, "How China's accession to the WTO affects global welfare?," *China Economic Review*, 2021, *69*, 101688.

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.

Feigenbaum, James J. and Andrew B. Hall, "How Legislators Respond to Localized Economic Shocks: Evidence from Chinese Import Competition," *The Journal of Politics*, 2015, *77* (4), 1012–1030.

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.

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.

Li, Bingjing, "Export expansion, skill acquisition and industry specialization: evidence from China," *Journal of International Economics*, 2018, *114*, 346–361.

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.

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.

Online Appendix

A Additional Tables and Figures

Table A.1: List of Countries in the ROW

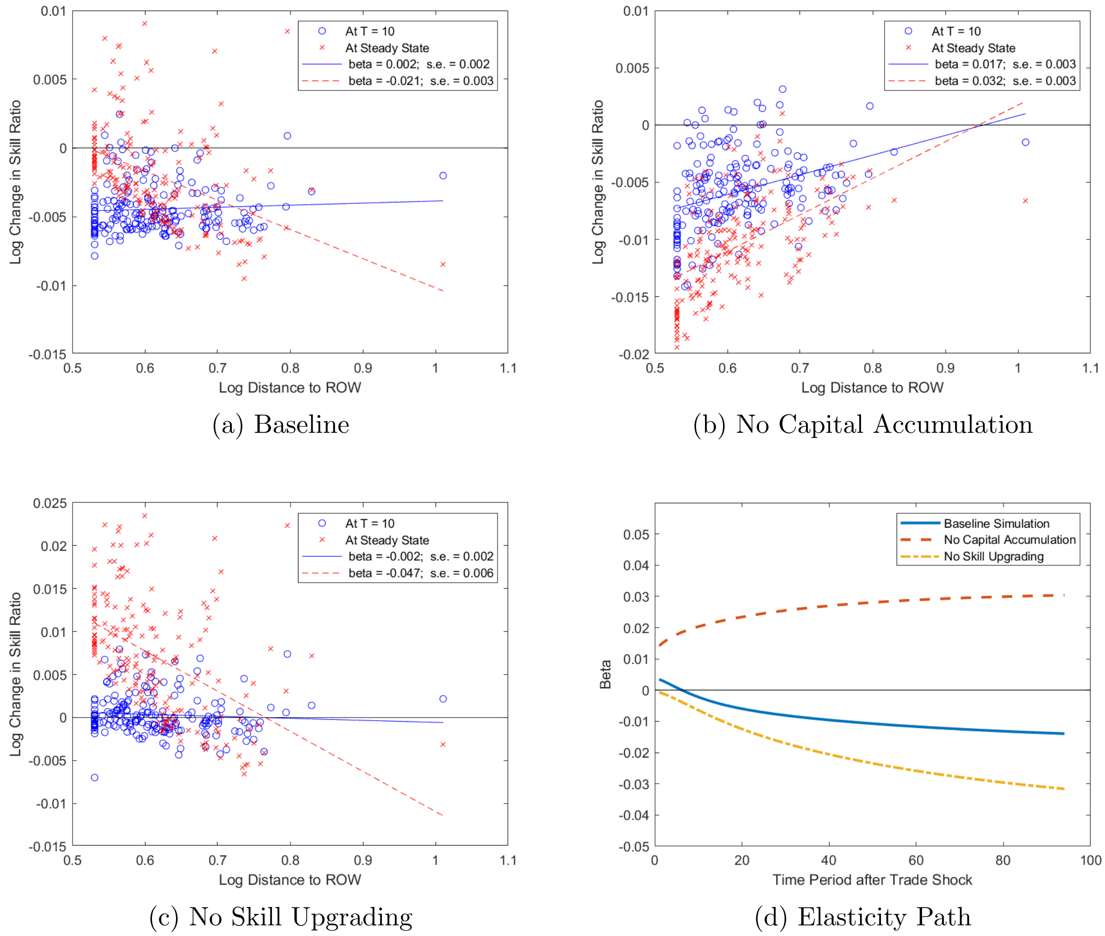
Australia	Belgium	Canada	Costa Rica	Czech Republic
Denmark	Estonia	Finland	France	Germany
Greece	Hungary	Ireland	Italy	Japan
Korea	Latvia	Lithuania	Luxembourg	Mexico
Netherlands	New Zealand	Poland	Portugal	Slovak Republic
Slovenia	Spain	Sweden	Switzerland	Türkiye
United Kingdom	United States			

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

Table A.2: 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



(a) Baseline

(b) No Capital Accumulation

(c) No Skill Upgrading

(d) Elasticity Path

Figure A.1: Impacts of Trade on Skill Ratio

Notes: This figure shows the effects of trade liberalization on skill ratio across Chinese prefectures. Panel (a)-(c) each shows the trade-induced skill ratio changes in the baseline model, the model without capital accumulation, and the model without skill upgrading, respectively. Panel (d) shows the distance elasticity of skill ratio change over time. Each dot represents a prefecture. The blue dots come from the cross section at period 10, and the red dots are from the steady state. The straight lines are linear fits.

B Details of the Model

B.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

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

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

$$p^j_{it} = \left[\sum_{n=1}^{N} \left(p^j_{in,t}\right)^{-\theta}\right]^{-\frac{1}{\theta}}. \tag{B.2}$$

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

$$\pi^j_{in,t} = \frac{(p^j_{in,t})^{-\theta}}{\sum_{m=1}^{N}(p^j_{im,t})^{-\theta}}. \tag{B.3}$$

B.2 Firm's Problem

In this part, we drop the time notation for brevity. Following Parro (2013), the problem of a producer in sector j at location i is given by:

$$\min_{l,s,k} \; w^l_i l^j_i + w^s_i s^j_i + r_i k^j_i,$$

subject to:

$$z_i \left[(\mu^j)^{\frac{1}{\sigma}}(z_i^{-\psi} l^j_i)^{\frac{\sigma-1}{\sigma}} + (1-\mu^j)^{\frac{1}{\sigma}}(z_i^{\psi} h^j_i)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \geq q^j_i \tag{B.4}$$

$$h^j_i = \left[(\lambda^j)^{\frac{1}{\eta}}(k^j_i)^{\frac{\eta-1}{\eta}} + (1-\lambda^j)^{\frac{1}{\eta}}(s^j_i)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}. \tag{B.5}$$

First order conditions for s^j_i and k^j_i yield:

$$k^j_i = \frac{\lambda_j}{1-\lambda_j}\left(\frac{w^s_i}{r_i}\right)^{\eta} s^j_i. \tag{B.6}$$

Using this expression to replace k^j_i in equation (B.5) and define the price w^{hj}_i for composite input h^j_i such that $w^{hj}_i h^j_i = r_i k^j_i + w^s_i s^j_i$, we obtain:

$$w^{hj}_i = \left[\lambda^j (r_{it})^{1-\eta} + (1-\lambda^j)(w^s_{it})^{1-\eta}\right]^{\frac{1}{1-\eta}}. \tag{B.7}$$

Similarly, first order conditions for l_i^j and h_i^j give:

$$l_i^j = \frac{\mu_j}{1-\mu_j}\left(\frac{w_i^{hj}}{w_i^l}\right)^{\sigma} z_i^{-2\psi(\sigma-1)} h_i^j. \tag{B.8}$$

Using equation (B.8) and define the unit cost c_i^j for the variety y_i^j such that $c_i^j y_i^j = w_i^l l_i^j + w_i^{hj} h_i^j$, we obtain:

$$c_i^j = \frac{1}{z_i}\left[\mu^j (z_i^{\psi} w_i^l)^{1-\sigma} + (1-\mu^j)(z_i^{-\psi} w_i^h)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{B.9}$$

B.3 Numerical Algorithm for Solving Steady State

We first write down the corresponding equilibrium conditions in the steady state. The population flow conditions (1)(2)(3)(4) become:

$$v_i^{d*} = \ln b_i^* + \ln \frac{w_i^{d*}}{p_i^*} + \rho \ln \sum_{e=l}^{s} \sum_{g=1}^{N} \exp\left[(\xi\beta v_g^{e*} - \kappa_{gi}^d - \kappa_e^d)/\rho\right], \tag{B.10}$$

$$D_{ig}^{ed*} = \frac{\exp\left[(\xi\beta v_g^{e*} - \kappa_{gi}^d - \kappa_e^d)/\rho\right]}{\sum_{e=l}^{s}\sum_{n=1}^{N} \exp\left[(\xi\beta v_n^{d*} - \kappa_{ni}^d - \kappa_e^d)/\rho\right]}, \tag{B.11}$$

$$L_i^{l*} = \xi \sum_{n=1}^{N} D_{in}^{ll*} L_n^{l*} + \left(L_i^{l*} + L_i^{s*}\right)(1-\xi), \tag{B.12}$$

and

$$L_i^{s*} = \xi\left(\sum_{n=1}^{N} D_{in}^{ss*} L_n^{s*} + \sum_{n=1}^{N} D_{in}^{sl*} L_n^{l*}\right). \tag{B.13}$$

The market clearing conditions (12)(13)(14)(15) become:

$$X_i^{j*} = \sum_{n=1}^{N} S_{ni}^{j*}\left[\gamma_j \sum_{m=1}^{J} X_n^{m*}\right], \tag{B.14}$$

$$w_i^{l*} = \frac{\sum_{j=1}^{J} \phi_i^{lj*} X_i^{j*}}{L_i^{l*}} \tag{B.15}$$

$$w_i^{s*} = \frac{\sum_{j=1}^{J} \phi_i^{sj*} X_i^{j*}}{L_i^{s*}} \tag{B.16}$$

$$r_i^{*} = \frac{\sum_{j=1}^{J} \phi_i^{kj*} X_i^{j*}}{k_i^{*}}, \tag{B.17}$$

with steady-state factor income shares given by:

$$\phi_i^{lj*} = \left[1 + (z_{it}^*)^{2\psi(\sigma-1)} \, \frac{1-\mu^j}{\mu^j} (\frac{w_i^{l*}}{w_i^{h*}})^{\sigma-1}\right]^{-1} \tag{B.18}$$

$$\phi_i^{sj*} = \left[1 + (z_{it}^*)^{-2\psi(\sigma-1)} \, \frac{\mu^j}{1-\mu^j} (\frac{w_i^{h*}}{w_i^{l*}})^{\sigma-1}\right]^{-1} \left[1 + \frac{\lambda^j}{1-\lambda^j} (\frac{w_i^{s*}}{r_i^*})^{\eta-1}\right]^{-1} \tag{B.19}$$

$$\phi_i^{kj*} = \left[1 + (z_{it}^*)^{-2\psi(\sigma-1)} \, \frac{\mu^j}{1-\mu^j} (\frac{w_i^{h*}}{w_i^{l*}})^{\sigma-1}\right]^{-1} \left[1 + \frac{1-\lambda^j}{\lambda^j} (\frac{r_i^*}{w_i^{s*}})^{\eta-1}\right]^{-1}. \tag{B.20}$$

The trade share adopts the following expression:

$$S_{ni}^{j*} = \frac{(p_{ni}^{j*})^{-\theta}}{\sum_{g=1}^{N} (p_{nm}^{j*})^{-\theta}}, \tag{B.21}$$

where

$$p_{ni}^{j*} = \frac{\tau_{ni}}{z_i^*}\Bigg[\mu^j \, (z_{it}^*)^{\psi(1-\sigma)} \, (w_i^{l*})^{1-\sigma} + \\ (1-\mu^j) \, (z_{it}^*)^{-\psi(1-\sigma)} \, \left[\lambda^j (r_i^*)^{1-\eta} + (1-\lambda^j)(w_i^{s*})^{1-\eta}\right]^{\frac{1-\sigma}{1-\eta}}\Bigg]^{\frac{1}{1-\sigma}}, \tag{B.22}$$

The capital accumulation condition becomes:

$$k_i^* = \xi\beta(1 - \delta + \frac{r_i^*}{p_i^*})k_i^*, \tag{B.23}$$

with

$$p_i^* = \prod_{j=1}^{J} \left[\sum_{n=1}^{N} (p_{in}^{j*})^{-\theta} \right]^{-\frac{\gamma_j}{\theta}}. \tag{B.24}$$

Given these conditions, the algorithm is as follows.

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

(2) Given $\{v_i^{l(0)}, v_i^{s(0)}\}_{i=1}^N$, compute migration shares $\{D_{ig}^{ed}\}_{i=1}^N$ for any skill type pair $\{e, d\}$ by (B.11), and then solve new labor allocations by (B.12) and (B.13) to obtain $\{l_i^{(1)}, s_i^{(1)}\}_{i=1}^N$.

(3) Given $\{l_i^{(1)}, s_i^{(1)}, k_i^{(0)}\}_{i=1}^N$, solve factor prices $\{w_i^l, w_i^s, w_i^k\}_{i=1}^N$ from markets clearing conditions as follows:

- (a) set an initial guess of factor prices $\{w_i^l, w_i^s, w_i^k\}_{i=1}^N$,
- (b) compute factor incomes shares $\{\phi_i^{jl}, \phi_i^{js}, \phi_i^{jk}\}_{i=1}^N$ from (B.18), (B.19), (B.20),
- (c) compute prices $\{p_{ni}\}_{n=1,i=1}^{N,N}$ and trade shares $\{S_{ni}\}_{i=1,n=1}^{N,N}$ from (B.22) and (B.21),
- (d) solve total output X_i^{j*} by (B.14)
- (e) obtain new factor prices $\{w_i^l, w_i^s, w_i^k\}_{i=1}^N$ by (B.15), (B.16), (B.17),
- (f) iterate until factor prices converge.

(4) Use $\{w_i^l, w_i^s, w_i^k\}_{i=1}^N$ to compute price index $\{p_i\}_{i=1}^N$ and solve new captial $\{k_i^{(1)}\}_{i=1}^N$ by (B.23).

(5) Given $\{v_i^{l(0)}, v_i^{s(0)}, w_i^{l(1)}, w_i^{s(1)}, p_i\}_{i=1}^N$, solve new value functions $\{v_i^{l(1)}, v_i^{s(1)}\}_{i=1}^N$ by (B.10).

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

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

B.4 Numerical Algorithm for Solving Path Equilibrium

Given the initial allocations of labor and capital, $\{l_{i0}, s_{i0}, k_{i0}\}$, 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 $\{v_{it}^{l(0)}, v_{it}^{s(0)}\}_{i=1,t=1}^{N,T}$, where $v_{iT}^{l(0)}$ and $v_{iT}^{s(0)}$ are approximated by steady-state level of value functions.

(2) Given $\{v_{it}^{l(0)}, v_{it}^{s(0)}\}_{i=1,t=1}^{N,T}$, solve migration shares $\{D_{ig,t}^{ed}\}_{i=1,t=1}^{N,T}$ for any skill type pair $\{e, d\}$ from (2).

(3) Use $\{D_{ig,t}^{ed}\}_{i=1,t=1}^{N,T}$ and $\{l_{i0}, s_{i0}\}$ to solve $\{l_{it}, s_{it}\}_{i=1,t=1}^{N,T}$ by (3) and (4).

(4) For each time period t, use current state variables $\{l_{it}, s_{it}, k_{it}^{(0)}\}_{i=1}^{N}$ to solve factor prices $\{w_{it}^{l}, w_{it}^{s}, w_{it}^{k}\}_{i=1}^{N}$:

- (a) set an initial guess of factor prices $\{w_{it}^{l}, w_{it}^{s}, w_{it}^{k}\}_{i=1}^{N}$,
- (b) compute factor incomes shares $\{\phi_{it}^{jl}, \phi_{it}^{js}, \phi_{it}^{jk}\}_{i=1}^{N}$ from (8), (9), (10),
- (c) compute trade shares $\{S_{ni,t}^{j}\}_{i=1}^{N}$ from (11) and (B.3),
- (d) solve total output by (15),
- (e) solve new factor prices by (12), (13), (14),
- (f) iterate until factor prices converge.

(5) Use solved factor prices $\{w_{it}^{l}, w_{it}^{s}, w_{it}^{k}\}_{i=1,t=1}^{N,T}$ to compute $\{p_{ni,t}\}_{i=1,t=1}^{N,T}$ by (11). Then obtain price index $\{p_{nt}\}_{n=1,t=1}^{N,T}$ by (B.2) and solve new captial allocations sequence $\{k_{i}^{(1)}\}_{i=1,t=1}^{N,T}$ from k_{i0} and (5).

(6) Set $\{v_{iT}^{l(1)}, v_{iT}^{s(1)}\} = \{v_{iT}^{l(0)}, v_{iT}^{s(0)}\}$. Given $\{v_{it}^{l(0)}, v_{it}^{s(0)}, w_{it}^{l}, w_{it}^{s}, p_{i}\}_{i=1,t=1}^{N,T}$ and $\{v_{iT}^{l(1)}, v_{iT}^{s(1)}\}$, solve new value functions $\{v_{it}^{l(1)}, v_{it}^{s(1)}\}_{i=1,t=1}^{N,T-1}$ backward by (1).

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

(8) Repeat steps (2)-(7) until value functions $\{v_{it}^{l}, v_{it}^{s}\}_{i=1,t=1}^{N,T}$ converge.

C Details of Data and Quantification

C.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.

C.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 **2007 Benchmark Input-Output Acount** of the U.S. provides 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 NAICS 2007 code. The labor income share is then divided between skilled and unskilled workers using the results from IPUMS USA.

C.3 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 C.1 and C.2 show 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 2006 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 NAICS 2007 code. Table C.1 and C.2 show 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.

Table C.1: Manufacturing Sectors

Panel A: Unskilled Manufacturing				Panel B: Skilled Manufacturing			
IO	Description	GB 2002	NACIS	IO	Description	GB 2002	NACIS
06	Food&Tabacco	C13-16	311-312	11	Petroleum	C25	324
07	Textile	C17	313	12	Chemicals	C26-30	325-326
08	Clothing	C18-19	314-316	14	Primary metals	C32-33	327
09	Wood&Furniture	C20-21	321,337	16	Machinery	C35-36	333
10	Paper&Printing	C22-24	322-323	17	Transportation equipment	C37	336
13	Nonmetallic mineral products	C31	327	18	Electrical equipment	C39	335
15	Manufactures of metal	C34	332	19	Telecommunication equipment	C40	334
21	Other manufactures	C42	339	20	Instruments	C41	334

Notes: This table shows the composition of the skilled and unskilled manufacturing sectors. Column "IO" shows the industry number 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; "NACIS" refers to the NACIS 2007 code. Both "GB 2002" and "NACIS" are manually matched with the IO industry by description.

Table C.2: Service Sectors

Panel A: Unskilled Service				Panel B: Skilled Service			
IO	Description	GB 2002	NACIS	IO	Description	GB 2002	NACIS
24	Natural gas	D45	2212	23	Electric power	D44	2211
25	Water	D46	2213	29	Computer service	G60-62	513-514
26	Construction	E47-50	23	32	Finance	J68-71	521-525
27	Transportation	F51-58	481-487,493	33	Real estate	K72	531
28	Postal	F59	491	34	Rental&Business	L73-74	532-561
30	Wholesale&Retail	H63,H65	42,441,445,452	36	Scientific service	M75	5412
31	Accommodation&Food	I66-67	721-722	37	Technical services	M76-N79	5412
38	Other services	N81-O83	81	39	Education	P84	61
				40	Health&Welfare	Q85-87	622-624
				41	Culture, Sports & Entertainment	R88-92	511-512,711-713
				42	Government	S93-97	G

Notes: This table shows the composition of the skilled and unskilled manufacturing sectors. Column "IO" shows the industry number 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; "NACIS" refers to the NACIS 2007 code. Both "GB 2002" and "NACIS" are manually matched with the IO industry by description.

C.4 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

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

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:

$$\bar{D}^d_{gi} = \frac{D^d_{gi}}{1 - D^d_{gg}}, \tag{C.2}$$

where D^d_{gi} is defined in the model as

$$D^d_{gi} = \frac{\exp\left[(\beta v^d_g - \kappa^d_{gi})/\rho\right]}{\sum_{n=1}^{N} \exp\left[(\beta v^d_n - \kappa^d_{ni})/\rho\right]}. \tag{C.3}$$

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

$$\frac{\bar{D}^d_{gi}}{\bar{D}^d_{ii}} \frac{\bar{D}^d_{ig}}{\bar{D}^d_{gg}} = \frac{D^d_{gi}}{D^d_{ii}} \frac{D^d_{ig}}{D^d_{gg}} = \exp\left[-\frac{1}{\rho}(\kappa^d_{gi} + \kappa^d_{ig})\right], \tag{C.4}$$

where we use the result

$$\frac{D^d_{gi}}{D^d_{ii}} = \frac{\exp\left[(\beta v^d_g - \kappa^d_{gi})/\rho\right]}{\exp\left[(\beta v^d_i)/\rho\right]} = \exp\left\{\left[\beta\left(v^d_g - v^d_i\right) - \kappa_{gi}\right]/\rho\right\}. \tag{C.5}$$

D Details of Robustness Checks

D.1 Cross-location Investment

We modify landlords' investment decisions following Kleinman et al. (2023). At the beginning of period t, landlords in each location can invest their capital in all locations. Then after each location'retrun on capital is determined in the temporal equilibrium, the landlords collect capital income from locations where they invest. At the end of period t, the landlords make intertemporal investment and saving decesion as in the baseline model.

The rate of return to capital in the host location n, denoted as $r_{n,t}$, comes from independently across locations in each period from a Frechet distribution $F_{n,t}(r) = e^{-(r/\alpha_{nt})^{-\nu}}$, where $\alpha_{nt} > 0$ controls the average rate of return in location n. For simplicity, we assume the Frechet distribution is identical across locations with a unit scale parameter, that is, $\alpha_{nt} = 1$ for all n and t.

Given these assumptions and the property of Frechet distribution, the capital inflow in each location i, K_{it}, is characterized as

$$K_{it} = \frac{(r_{it})^{\upsilon}}{\sum_{n=1}^{N}(r_{nt})^{\upsilon}} K_t, \tag{D.1}$$

where K_t is the aggregate capital and r_{it} is return of capital to capital in location i.

It can be found that capital inflow in each location solely increases with its own rate of return to capital. The realized return on capital, R_t, is the same across locations and given by:

$$R_t = \Gamma\left(\frac{\upsilon - 1}{\upsilon}\right)\left[\sum_{n=1}^{N}(r_{nt})^{\upsilon}\right]^{\frac{1}{\upsilon}} \tag{D.2}$$

In each period t, a landlord in location n has the following budget constraint:

$$R_t k_{it} = p_{it}(c_{it}^{k} + k_{it+1} - (1-\delta)k_{it}),$$

The optimal investment decision is thus

$$k_{it+1} = \xi\beta\left(1 - \delta + \frac{R_t}{p_{it}}\right)k_{it}. \tag{D.3}$$

The aggregate capital is then a sum of the total capital stock in all locations: $K_{t+1} = \sum_{n=1}^{N} k_{nt+1}$.

Now we are ready to close the model and describe the market clearing conditions. With cross-location investment, each location's total income is no longer equal to its total output, as its total output depends on the capital inflow but total income depends on the capital outflow. Let I_i denote location i's total income and X_{it}^j denote its sales from industry j. The good market clearing condition then is

$$X_{it}^j = \sum_{n=1}^{N} S_{nit}^j\left(\gamma_j I_{nt}\right), \tag{D.4}$$

where $I_{nt} = w_{nt}^l l_{nt} + w_{nt}^s s_{nt} + R_t k_{nt}$. The equation above implies all locations' expenditure on industry j's goods produced by location i equal its total sales.

Finally, the rate of return to capital in each location is given by the zero profit condition

$$r_{it} = \frac{\sum_{j=1}^{J}\phi_{it}^{kj}X_i^j}{k_{it}}. \tag{D.5}$$

D.2 Alternative Trade Costs

This section describes the model inversion to recover the trade costs between China and the ROW. We expect that the growth of the trade-to-GDP ratio between 2000 and 2006 in data corresponds to reductions in trade costs between China and the ROW. Therefore, we calibrate the trade costs in 2006 so that the model-generated trade-to-output ratio for China is consistent with the observed trade-to-GDP ratio in 2006. The calibration result shows that, from 2000 to 2006, the trade cost declined by 16 percent in the unskilled manufacturing sector and 17 percent in the skilled manufacturing sector. For comparison, the directly estimated results following Head and Ries (2001) show a 10 percent decline in the unskilled and a 14 percent decline in the skilled manufacturing sector.