The Long and Short-Run Spatial Impacts of Trade

Lin Ma Yunlong Song Yang Tang *

January 15, 2024

Abstract

We study the long and short-run spatial impacts of trade liberalization in a model with internal migration, skill formation, and capital accumulation. We show that the spatial impacts of trade on migration flows and skill composition depend critically on the time horizon. The time-varying impacts come from capital accumulation: trade liberalization allows faster capital accumulation in coastal locations, shifting comparative advantage across locations in the long run. In an unskill-abundance country such as China, unskilled workers are attracted to the coastal locations due to trade liberalization in the short run; in the long run, however, the skilled workers are more likely to move to coastal locations due to the faster accumulation of capital stock in those locations and capital-skill complementarity.

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 Dimitrijie Ruzic, Davin Chor, B. Ravikumar, Daniel Xu, Liugang Sheng, and participants in xxxx

1 Introduction

International trade leaves profound distributional impacts across regions within a country. While a country gains from trade on average, workers in a particular region could either 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), US labor markets specializing in industries that bear the brunt of Chinese import competition tend to perform poorly. On the other hand, Chinese labor markets closer to the coasts and thus with better access to the world market tend to perform better than the inland ones (Fan, 2019). Understanding the spatial impacts of trade underpins the policy discussions that could either alleviate the localized negative impacts of globalization or facilitate a better distribution of the positive impacts through factors or goods movements within the country.

We contribute to this literature by highlighting that the spatial impacts of trade vary substantially over the time horizon. The dependency on time horizon comes from the endogenous and gradual responses of factor accumulation to trade shocks. Consider the case of China upon joining the World Trade Organization (WTO) in 2001. As China has comparative advantages in unskilled-intensive manufacturing industries at the time, the export boom should mainly draw the unskilled migrants towards the coastal locations in the short run. In the long run, however, the prosperity brought in by the export boom allows for faster accumulation of physical capital — the better infrastructure, buildings, and machinery at the booming locations. As higher capital stocks are generally more complementary with the productivity of the skilled workers than the unskilled ones, the coastal locations might become more attractive to skilled workers in the long run. In this particular example, the answers to a key question concerning the spatial impacts of trade — how skilled and unskilled workers migrate differently in response to trade shocks — are drastically different depending on the time horizon.

To systematically study the spatial impacts of trade over a long time horizon, we develop a dynamic spatial model with endogenous production factor accumulation in this paper. Our

[1] In the context of the US, 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.

framework highlights how the accumulation and distribution of production factors across space — physical capital and human capital — respond 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 (Caliendo et al., 2019; Kleinman et al., 2023). We also model type-specific migration frictions and capital-skill complementarity to highlight the interaction between production factors and geography. In our model, the supply of capital stock in each location at each period is determined by the forward-looking investment decision of the landlord as in Kleinman et al. (2023). Each location's supply of skills at each period is determined by the previous period's migrant inflow of skilled workers and unskilled workers who choose to upgrade their skill type. In the short run, the mobility frictions and the gradual nature of capital accumulation and skill acquisition imply that the initial factor endowments mainly determine the pattern of comparative advantages across countries and regions. Over time, however, the distribution of production factors shifts across space due to endogenous capital accumulation, skill acquisition, and migration in response to trade shocks. As a result, the model has the potential to deliver spatial impacts of trade that are dependent on the time horizon.

We quantify our model in the context of China, as China experienced great trade liberalization in the early 2000s and also features large spatial inequalities. We consider a world economy with 196 prefecture-level cities in China and an additional location representing the rest of the world (ROW) and set the year 2000 as the initial year. We also construct four sectors by skill intensity. Using rich datasets, we can uncover locational fundamentals and also several parameters. Conditional on migration elasticity, we fully estimate spatial migration frictions for both unskilled workers and skilled workers. The estimation result suggests that, on average, unskilled workers face about 9.6 percent larger migration frictions than skilled workers. We also find that, from the year 2000 to 2006, the trade cost between China and the ROW for the unskilled manufacturing sector dropped by around 16.5 percent, and that for the skilled manufacturing sector dropped by around 17.5 percent.

In quantitative analysis, we first study how skill premium is determined in our model and its behavior in the long run across space. With capital-skill complementarity, capital accumulation is important as it increases the average skill premium in China from 2.21 to

2.67 in the long run. Without skill upgrading, the average skill premium is much higher than that in the baseline case (6.53 compared to 2.67). This is quite intuitive since the total supply of skill in China is fixed if unskilled workers are not allowed to upgrade their skill type but capital still accumulates over time. We also compare the model-generated skill premium with that observed in data for the year 2005, and find that both capital accumulation and skill upgrading are important in explaining the spatial variations in skill premium. Furthermore, we find that, both in the short run and in the long run, skill premiums always tend to converge across space. This spatial convergence is mainly driven by spatial migration over time, as cities with initially high skill premiums attract skilled workers in other locations. This movement of skilled workers increases the relative supply of skill and drops skill premiums in those cities.

Next, we proceed by examining the spatial impacts of China's early 2000s trade liberalization and try to understand how capital accumulation interacts with the impacts. As China is relatively abundant in unskilled workers, it specializes more in unskilled sector in international trade. We first find that our model without capital accumulation suggests that further opening trade induces coastal cities to attract workers from inland cities, but attract less skilled workers, and also have a larger decrease in skill premium. This result is consistent with the Stolper-Samuelson theorem. However, with capital accumulation, coastal cities benefit much more from trade liberalization. Those locations become much more attractive for all workers and particularly for skilled workers, since capital accumulates faster and more there. Furthermore, trade liberalization is more likely to increase skill premiums in coastal cities but decline those in inland cities. This result contrasts starkly with the case without capital accumulation, justifying the role of capital accumulation in investigating spatial impacts of trade.

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; Parro, 2013; Burstein and Vogel, 2017). Many of those papers consider distributional impact across countries and feature little internal geography within a country. Our paper studies how spatial skill premiums are interrelated with internal trade, internal migration, and capital accumulation after trade shock. We also contribute to this strand of literature by highlighting the role of capital accumulation and skill formation and showing that trade's distributional impact depends critically on the time horizon. The results also suggest that the rapid capital accumulation in those countries can partially explain the increase in skill premium in developing countries.

Our work also relates to the quantitative spatial literature (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, Liu and Redding (2023). Relative to Kleinman et al. (2023), we consider dynamic migration decisions and skill upgrading decisions for multiple types of workers. We also include capital-skill complementarity to study skill premium better.

Finally, our paper also contributes to the literature studying China's spatial economy (Fan, 2019; Tombe and Zhu, 2019; Ma and Tang, 2020, 2022). While Fan (2019) considers the distributional impact of trade with capital-skill complementarity, we highlight the role of capital accumulation and dynamic migration.

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, Liu and Redding (2023). The economy has N geographically segmented cities indexed by i and J sectors indexed by j. Time is discrete and indexed by t. 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 their entire income is spent on consumption. At the end of each period, all workers decide where to migrate, and unskilled workers decide whether to upgrade their skill levels.

The worker's utility flow depends on his consumption and the residing city's amenities:

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

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 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 earning wages and consuming goods in the current city, a worker needs to decide which city to live in the next period. The migration decision depends on the expected lifetime utility from living in any of the J cities, skill-specific bilateral migration costs denoted as $\kappa_{ni,t}^d$ 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 $\kappa_{ni,t}^d$ apply: (1) $\kappa_{ni,t}^d > 0$ for $n \neq i$, (2) $\kappa_{ii,t}^d = 0$ and (3) $\kappa_{ni,t}^d \leq \kappa_{nj,t}^d + \kappa_{ji,t}^d$ for any third location j, and an idiosyncratic preference shock towards each destination city, denoted as ε_{nt}.

In addition to the migration decision, an unskilled worker decides whether to upgrade to a skilled worker, subject to switching costs of κ_s^l in utility. Skilled workers cannot downgrade to unskilled again, and we normalize the costs of staying as a skilled worker to zero.[2] As it will be clear later, the skill-upgrading decision depends on comparing the option value of

[2] We essentially assume (1) $\kappa_d^d = 0$ for $d \in \{l, s\}$, (2) $\kappa_l^s = \infty$ and (3) $\kappa_s^l < \infty$.

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 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, and ρ is the inverse migration elasticity. 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 city i and time t that migrate to city

g and become type e in $t+1$ as

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

$1/\rho$ in Equation (2) captures the migration elasticity. Finally, the supplies of unskilled workers and skilled workers in each location evolve according to

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

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 unskilled workers who upgrade their skills.

2.3 Landlords

We closely follow Kleinman et al. (2023) in characterizing 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. 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

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

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:

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

2.4 Production

Firms at each location i and industry j specialize in one variety, 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:

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

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.[3] 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.

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

The production structure implies that the unit cost of production for a variety in the industry j and location i (c^j_{it}) is given by

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

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.

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

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

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

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

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

Combining equations xx, xx, and xx, price index in location n and industry j,$p^j_{n,t}$, satisfies

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

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 $L^l_{it} + L^s_{it}$:

$$b_{it} = \bar{b}_{it}(L^l_{it} + L^s_{it})^{\alpha_b},$$

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

$$z_{it} = \bar{z}_{it}(L^l_{it} + L^s_{it})^{\alpha_z},$$

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 $\{L^l_{i0}, L^s_{i0}, k_{i0}\}$ in each location, the dynamic equilibrium contains a sequence of location-specific prices $\{w^l_{it}, w^s_{it}, r_{it}\}^{\infty}_{t=0}$, quantities $\{L^l_{it}, L^s_{it}, k_{it}\}^{\infty}_{t=0}$ and value functions $\{v^l_{it}, v^s_{it}\}^{\infty}_{t=0}$, such that the following holds.

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

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

3. The evolution of capital and population is characterized as in equation (3)(4)(5).

4. labor market for unskilled and skilled workers, and capital market clear in each location.

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

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

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

where X_{it}^{j} denotes total income earned in city 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 city i and destination city n at time t defined in equation 28 in the appendix A.1.

The steady state of the economy is reached when all the exogenous fundamentals of the economy become constant over time, and all the endogenous variables are steady accordingly. We formally define the steady state of the economy as:

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 on calibrating and estimating the key parameters of the model.

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.

The model contains four sectors: the skill- and unskill-intensive manufacturing and services sectors. We map the 82 industries observed in China's 2002 Industrial Classification for National Economic Activities into four broad sectors by their skill intensities[4]. 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 sector and those below into the unskilled sector. Table 5 and Table 6 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 start by using data on investment from *China City Statistical Yearbooks* from 1994 to 2000 to construct a panel dataset of capital stocks. Specifically, the capital stock in location i at time t is given by

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

where I_{it} is the real investment observed in the data, and K_{it} is the sequence of capital stock inferred using the perpetual inventory method. We compute real investment as Nominal Investment$_{it}$ $\times$ Investment Deflator$_{it}$, where the nominal investment is proxied using "gross fixed capital formation" from the *China City Statistical Yearbooks*; the investment deflators also come from the same source. To infer the initial capital stock, we adopt the

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

standard approach as in Young (2003) and assume capital stock in 1994 is equal to real investment at that year divided by the depreciation rate.

Rest of the World The ROW is an aggregate of 32 OECD countries. In Appendix xx, we list all 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).

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, Liu and Redding (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, Wu and Zhang (2004)	Capital depreciation rate
τ_{ni}	-	Ma and Tang (2020)	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 several parameters in the model. They are either from the results of current literature or from data.

3.2 Geography

Trade Costs Products from the manufacturing sectors are tradable across locations and those from the service industries are non-tradable. We directly use the estimated trade costs from Ma and Tang (2022) between Chinese prefectures. The trade costs in that paper are based on freight infrastructure on road and rail networks and usage in each year during our sample period. Within China, trade costs do not vary across sectors.

We modify the methods in Ma and Tang (2022) to estimate the trade costs between Chinese prefectures and the ROW. Start with the 27 port cities in China identified in Ma and Tang (2022), we assume that all port cities face the same trade cost with the ROW in a given sector, denoted by $\tau^j_{ROW,t}$ 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 $\tau_{i,\text{port}_i} \cdot \tau^j_{ROW,t}$, where port_i is the nearest port to location i determined by the τ matrix within China. Note that 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. Rearranging equation (28), the changes in trade costs can be inferred as:

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

where $\widehat{S^j}_{(\cdot),t}$ is the changes in trade flow in sector j between year t and the initial year, 2000. Given trade elasticity θ and the observed flows, we compute the *changes* in trade costs for each year between 2000 and 2006 sector-by-sector. Lastly, we infer the initial levels of trade costs in 2000, $\tau^j_{ROW,2000}$, by inverting the model in the initial spatial equilibrium and exactly matching the observed trade costs in that year. Our approach here is different from Head and Ries (2001), who directly inferred the levels of trade costs in each year. This is because our model features a rich internal geography inside China, while Head and Ries (2001) abstracted away from such elements. As a result, the levels of $\{\tau^j_{ROW,t}\}$ inferred using the above methods 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. Our method allows us to match the trade shares between China and ROW observed along the transition path.[5] With the estimated $\tau^j_{ROW,t}$, we have complete trade costs matrices across all locations

[5] One can also inver the model along the transition path to match the trade shares to the data for all years

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 the data from two sources: 1) the *2005 One Percent Population Survey*, and 2) the passenger transportation data from Ma and Tang (2022). The population survey allows us to compute the number of migrants in city g with hukou from city i for skill type d, denoted as $\tilde{D}^d_{gi,t}$. We use $\bar{D}^d_{gi,t}$ to denote the fraction of $\tilde{D}^d_{gi,t}$ in the total type d population in origin city i: $\bar{D}^d_{gi,t} = \tilde{D}^d_{gi,t}/L^d_{it}$. We need information on the bilateral migration probability $D^d_{gi,t}$ to estimate migration frictions. In the appendix B.3, we establish the following relation between $\bar{D}^d_{gi,t}$ and our theoretical prediction of migration share $D^d_{gi,t}$:

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

where $D^d_{gi,t}$ is given by

$$D^d_{gi,t} = \frac{\exp\left[(\beta v^d_{gt+1} - \kappa^d_{gi,t})/\rho\right]}{\sum_{n=1}^{N} \exp\left[(\beta v^d_{nt+1} - \kappa^d_{ni,t})/\rho\right]}. \tag{17}$$

To simplify notation, we drop the time subscript henceforth. Double differencing the relation above gives:

$$\frac{\bar{D}^d_{gi}}{\bar{D}^d_{ii}} \frac{\bar{D}^d_{ig}}{\bar{D}^d_{gg}} = \frac{D^d_{gi}}{D^d_{ii}} \frac{D^d_{ig}}{D^d_{gg}} = \exp\left[-\frac{1}{\rho}(\kappa^d_{gi} + \kappa^d_{ig})\right], \tag{18}$$

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

$$\kappa^d_{gi} = \kappa^d_g + \bar{\kappa}_{gi}, \tag{19}$$

where $\bar{\kappa}_{gi} = \bar{\kappa}_{ig}$ is symmetric travel cost between location i and g, and κ^d_g is type-specific

after 2000. However, simulating the transition path is much more computationally expensive than solving the initial equilibrium.

entry barrier for entering location g. The travel costs are assumed to be determined by infrastructure and geography and, therefore, exogenously fixed. Given travel costs, estimating migration costs is equivalent to estimating entry barriers for all locations. Denote $\frac{\bar{D}_{gi}^d}{\bar{D}_{ii}^d}\frac{\bar{D}_{ig}^d}{\bar{D}_{gg}^d}$ by R_{gi}^d, then equation (18) becomes

$$R_{gi}^d = \exp\left[-\frac{1}{\rho}(\kappa_g^d + \kappa_i^d + 2\bar{\kappa}_{gi})\right]. \tag{20}$$

Given travel costs and migration elasticity parameter ρ, we estimate entry barriers for each location using Poisson regression based upon equation 20. Figure 1 shows our estimation results, represented as distributions of entry barriers across locations for unskilled workers and skilled workers. It suggests that unskilled workers face relatively larger entry barriers than skilled workers overall. In other words, skilled workers are more mobile across locations.

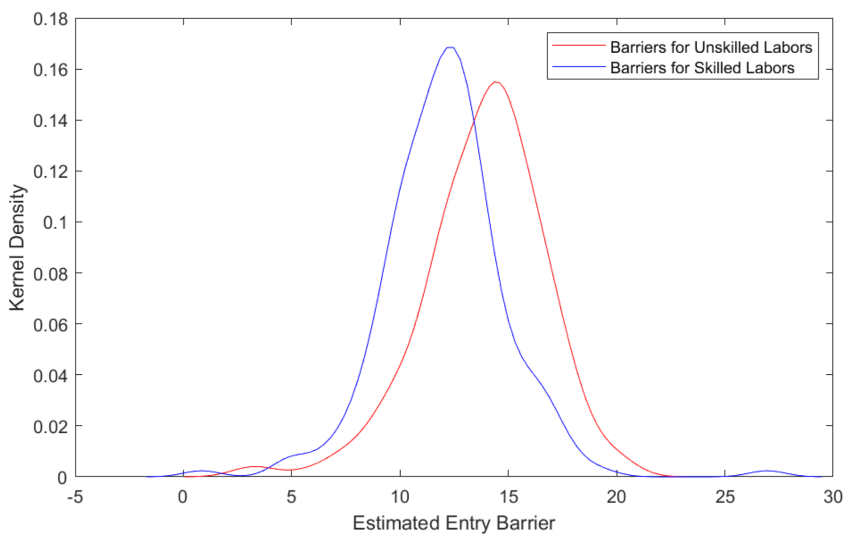


Figure 1: Distribution of Entry Barriers

Notes: this figure shows the distributions of estimated entry barriers for both unskilled workers and skilled workers. Entry barriers are estimated using PPML

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. Another group of parameters comes 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 externally choose values for some parameters based on the estimates in the literature. Specifically, we choose trade elasticity $\theta = 5$ from Costinot and Rodríguez-Clare (2014). We assume an annual discount rate of $\beta = 0.97$ and choose 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 $\xi = 0.993$.

Parameters Calibrated in the Initial Equilibrium A subset of parameters including $\{\bar{z}_i, \mu^j, \lambda^j, \}$ are calibrated so that the initial static equilibrium matched the observed economic conditions in 2000.[6] 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 the year 2000. We normalize the fundamental productivity in the $i = 1$ to $\bar{z}_1 = 1$.

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

[6]The trade costs between Chinese prefectures and the ROW discussed earlier, $\tau^j_{ROW,2000}$, are also calibrated at this stage.

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 IPUMS 1%. The U.S.'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 attach a higher weight to unskilled workers than skilled workers within the manufacturing and service sectors. Moreover, our estimation also shows that capital takes up higher weights in unskilled manufacturing sectors ($\lambda^j = 0.91$) in China 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. As the ROW is relatively more abundant in skilled workers and capital in the data, and 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.

Table 2: Calibrated Production Weights

Panel (a): Calibrated Production Weights for China by Sectors

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): Calibrated Production Weights for the ROW by Sectors

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

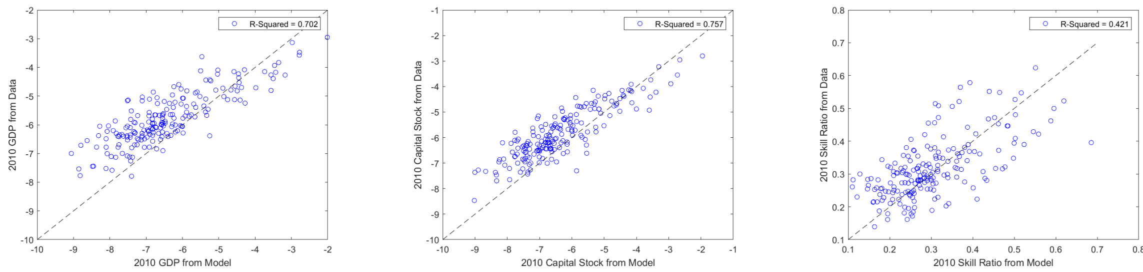
Notes: this table reports the results of production weights in four different sectors for China and the ROW. The weights are calibrated in the initial static equilibrium by targeting sector-level factor income shares. $\mu \in [0, 1]$ is the weight of unskilled workers and $\lambda \in [0, 1]$ is the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

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 $\{\kappa_s^l\}$ and city-specific amenities $\{\bar{b}_i\}$. Specifically, $\{\kappa_s^l\}$ is chosen to match the national 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 lifetime utility among newborns in the initial period. Compared to the migration costs, we find that the skill-upgrading costs are slightly higher than the average migration costs among unskilled workers.

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.

3.4 Model Fit

The quantification strategy outlined above matches the untargeted data moments. Figure 2 compares the model-predicted spatial distribution of total output, capital stock, and skill ratio with the data counterpart, none of which is our calibration target. The model matches the data reasonably well with an R-squared value at around 0.5.



(a) 2010 GDP (b) 2010 Capital (c) 2010 Skill Ratio

Figure 2: Data Fitness from Model Simulation

Notes: these figures compare results from baseline model simulation with observed actual data. Each dot represents a prefecture in China, and the black dotted line is the 45-degree line. Variables in panel (a), (b) and (c) are logarithmic.

We have also compared the model-predicted spatial distribution of skill premium with that from the data. Table 3 summarizes the fitness of our baseline specification. Our baseline model can explain around 9 percent spatial variations of observed skill premium, indicated by R-squared equal to 0.091.

Table 3: Fitness of Skill Premium

	Dependent variable: 2005 Skill Premium in Data Explanatory variable: 2005 Skill Premium in Model		
	Beta	s.e.	R-Squared
Baseline	0.201	0.201	0.091
Without Capital Accumulation	0.138	0.138	0.046
Without Skill Upgrading	0.129	0.129	0.068
$\psi = 0$	0.317	0.317	0.091

Notes: this table reports the results of goodness-of-fit of different model specifications. Each row corresponds to an OLS regression of skill premium in data on skill premium in the model for the year 2005. In the second row, we set each location's capital stock as the initial capital stock and then constant over time to shut down capital accumulation; in the third row, we set the skill upgrading cost as infinity to shut down skill upgrading.

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 further decompose the time variation of the spatial impacts of trade, 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 year level over time. In other words, we change the landlords' optimal investment decision in equation (5) and force them to invest only to cover the depreciated capital exactly in each period. 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.

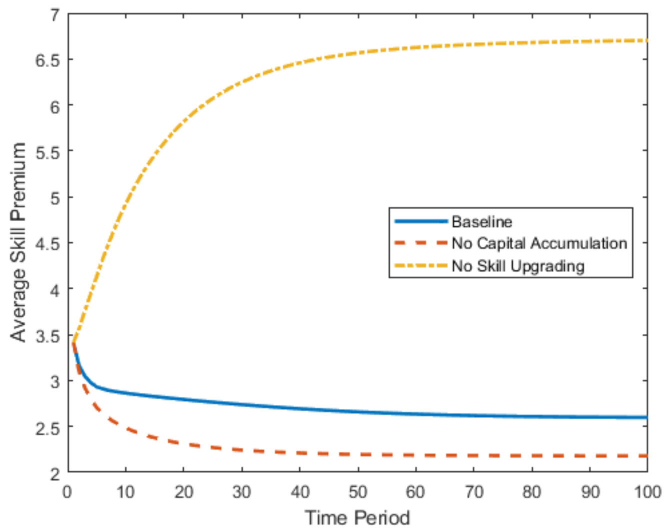
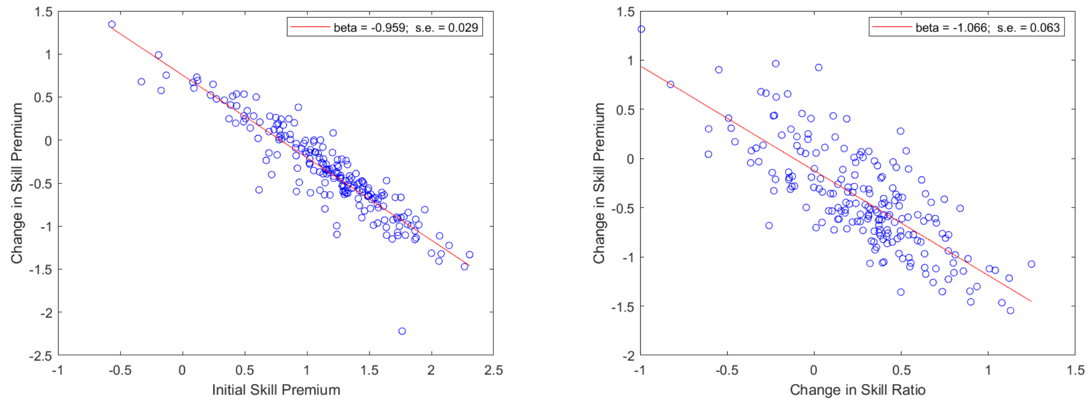


Figure 3: Average Skill Premium over Time

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.

In addition to their aggregate impacts, capital accumulation and skill upgrading also matter for variations of skill premiums across locations. Table 3 summarizes the model's explanatory power regarding spatial variations of skill premium. The baseline model can account for around 9.1 percent of the spatial variation of skill premium along the transition path in the year 2005, as measured by the R^2 when we linearly project the observed skill premium onto the model simulation. On the other hand, shutting down capital accumulation drastically reduces the explanatory power by more than half to only 4.1 percent, and shutting down skill upgrading reduces the explanatory power by one quarter to 6.8 percent.

The second feature of the baseline simulation is the spatial convergence of skill premiums. Figure 4 panel (a) 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 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. 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 skill ratio changes and skill premium changes: 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 relative spatial movement of skilled workers over time reduces the large disparity of spatial skill premiums. Moreover, the convergence of skill premiums does not depend on the above-mentioned determinants of skill premium in levels. In particular, regardless of capital accumulation or skill upgrading, the model yields similar β-convergence coefficients compared to panel (a) in figure 4.

4.2 The Aggregate and 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.

The aggregate welfare gain of WTO accession is 0.49 percent, which is in line with the findings in the trade and the urban literature (see, e.g. [:ADD CITATION:]).[7] 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, as coastal cities are closer to the ROW, they benefit more from the trade liberalization: the welfare gain is 0.47 percent among coastal prefectures and 0.39 percent among the inland ones.[8]

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.

4.2.1 Trade Liberalization in the Short and Long run

The spatial impacts of trade vary along the transition path. In the following, we compare its impacts on capital accumulation, population movements, skill ratio, premium, and spatial inequality between the short and long run.

[7] we use population-weighted utility changes in the initial period to measure the welfare impact

[8] We define "coastal cities" as those with a distance to the ROW less than the median distance among all cities to the ROW.

Capital accumulation Trade liberalization enriches the coastal prefectures and induces faster capital accumulation there than in the inland locations. 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 change changes during 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. Note that 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 demand shock, in turn, leads to higher returns to capital and thus faster accumulation, as indicated in Equation (5).

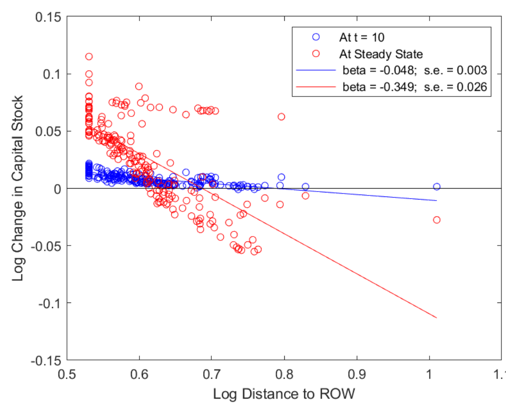


Figure 5: Impacts of Trade on Capital Accumulation

Notes: These figures show the effects of trade liberalization on capital accumulation across Chinese prefectures in the short run and long run from baseline simulation. Each dot represents a prefecture. The straight lines are linear fits.

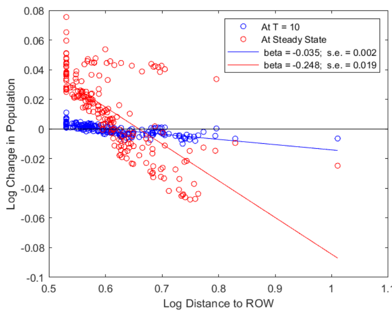
The differential impacts of trade on capital accumulation are the fundamental driving forces behind the spatial impacts of trade that we will discuss in the rest of the section. As the coastal locations stand to accumulate capital stock faster, in the long run, these locations start to gain comparative advantage in capital-intensive industries; 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, such as in Tombe and Zhu (2019) and Fan (2019), trade liberalization drives the population movement toward the coastal locations. We find similar impacts of trade on internal migration as well. Panel (a) in Figure 6 plots the impacts of trade on population changes against the distance from 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 $t = 10$.

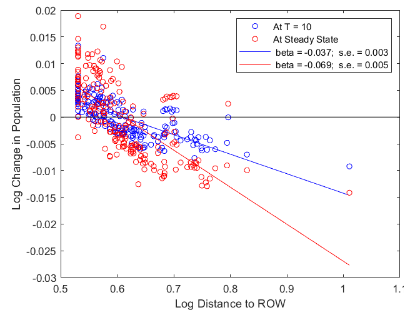
Strikingly, the long-run impacts of trade on population concentration are much stronger: the absolute value of the distance elasticity of population change increases from -0.035 to -0.248 at the steady state. In other words, researchers would underestimate the impacts of international trade on population 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.

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.

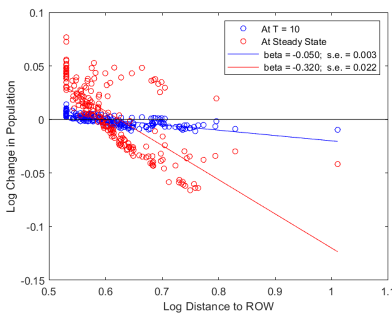
On the other hand, shutting down skill upgrading would lead to an even higher long- and short-run gap in the distance elasticity. As shown in Panel (c) of the same figure, without skill



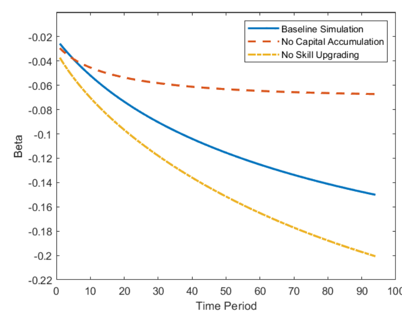
(a) Baseline



(b) No Capital Accumulation



(c) No Skill Upgrading



(d) Elasticity Path

Figure 6: Impacts of Trade on Population

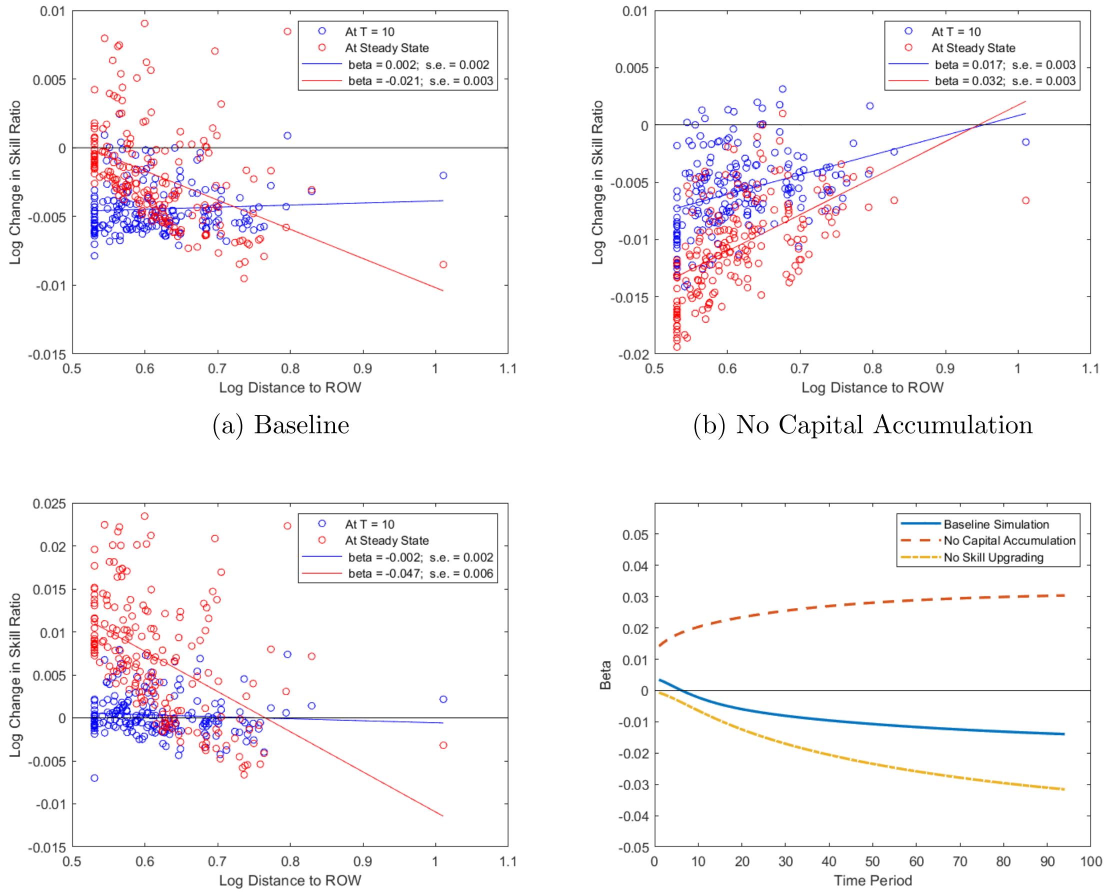
Notes: These figures show the effects of trade liberalization on total workers across Chinese prefectures. Panel (a)-(c) each shows the trade-induced population changes in the baseline model, in the model without capital accumulation, and in the model without skill upgrading, respectively. Panel (d) shows the distance elasticity of population 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.

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 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 skill ratio, defined as the fraction of skilled workers in the total workforce 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 nearly 0 in the short run but turns negative at 0.021 in the long run. In other words, while the skilled and the unskilled workers are equally likely to migrate toward the coastal locations in the short run, the skilled populations are much more likely to migrate to the coast than the unskilled ones.

Intuitively, the skill composition of migrants 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 in skilled workers in locations that enjoy an abundance of capital stock. The two forces roughly cancel out in the short run, leading to zero distance elasticity. In the long run, however, the advantage of the coastal locations in capital accumulation dominates, and subsequently, we observe the 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, respectively. 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 the skilled workers with a $\zeta^{\mathrm{sr}}(t) = -0.047$. The higher distance elasticity comes from the fact that the coastal locations rely more on internal migration to meet the higher demand for skilled workers from capital accumulation if the local unskilled workers cannot upgrade their skills.



(a) Baseline

(b) No Capital Accumulation

(c) No Skill Upgrading

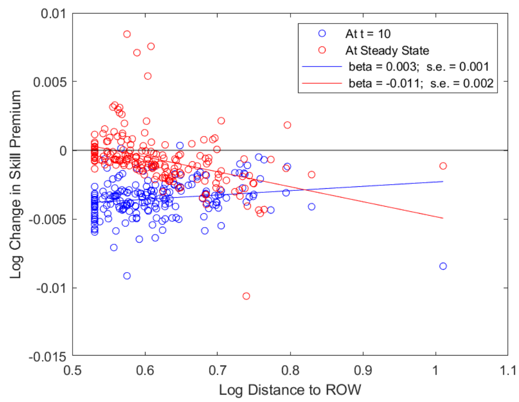
(d) Elasticity Path

Figure 7: Impacts of Trade on Skill Ratio

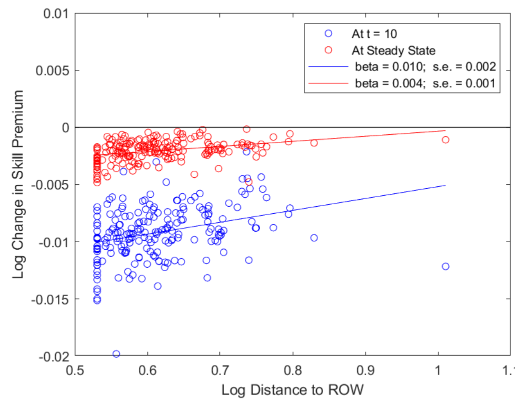
Notes: this figure shows the effects of trade liberalization on skill ratio across Chinese prefectures. Panel (a)-(c) each shows the trade-induced skill ratio changes in the baseline model, in the model without capital accumulation, and in the model without skill upgrading respectively. Panel (d) shows 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.

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

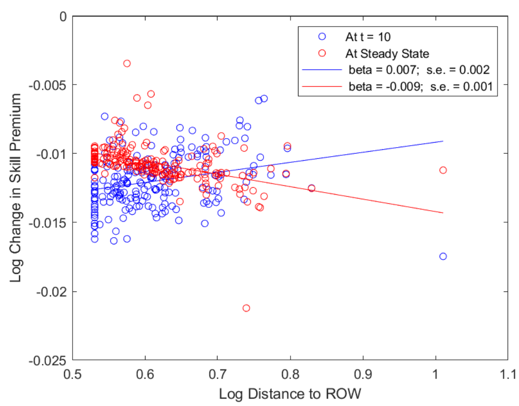
Figure 8 presents the impacts of trade on skill premium similarly to the previous ones. In the short run, the changes in skill premium are negative in all cities, a finding led by the Stoper-Samuelson effect. The distance elasticities of changes in skill premium are positive, suggesting coastal cities experience the largest decline in skill premium. However, in the long run, as panel (b) shows, changes in skill premium become positive in most coastal cities, and the distance elasticity becomes negative. This suggests that the effect led by capital accumulation dominates the Stoper-Samuelson effect across space. Overall, skill premiums increased in 65 cities, mainly coastal cities, and dropped the most in inland cities.



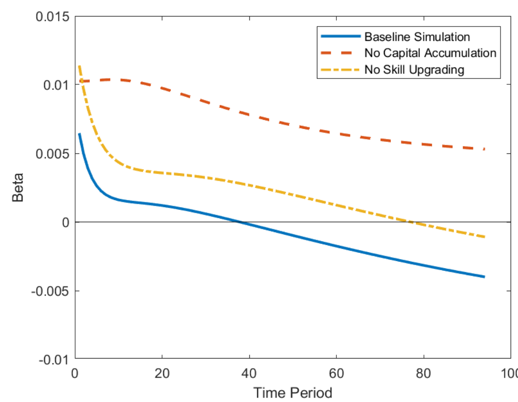
(a) Baseline



(b) No Capital Accumulation



(c) No Skill Upgrading

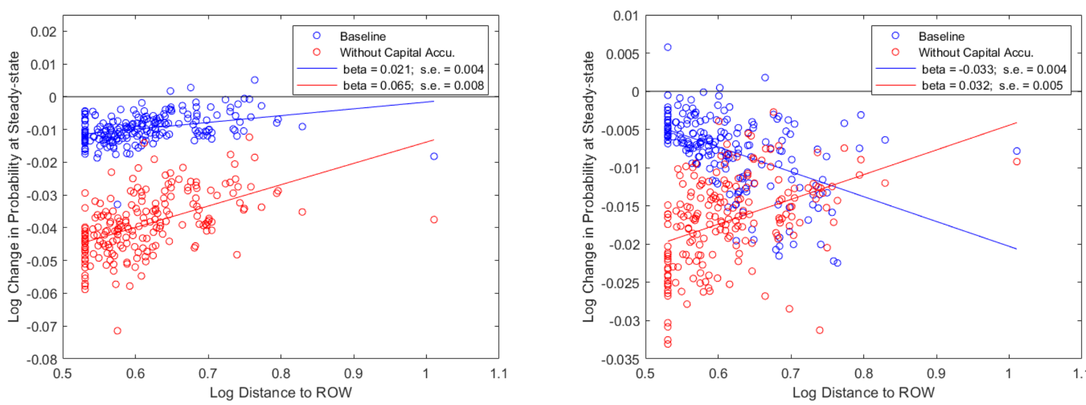


(d) Elasticity Path

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.

Skill Upgrading Figure 9 illustrates the impact of trade on skill upgrading decisions. It 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, disproportionately more unskilled individuals in coastal cities remain unskilled than inland cities, as suggested by the positive distance elasticity in the left panel. Over time, however, as capital accumulates faster in coastal cities due to the trade shock, this negative effect largely dissipates in these areas. In fact, the sign of distance elasticity reverses over time (from 0.021 to -0.033).



(a) At T=10

(b) At Steady State

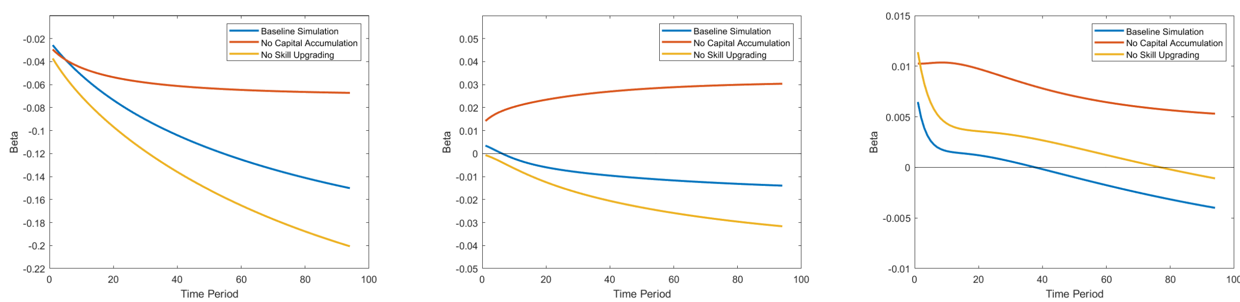
Figure 9: Impacts of Trade on Skill Upgrading Depends on Time Horizon and Capital Accumulation

Notes: this figure shows the effects of trade liberalization on the probability of skill upgrading across Chinese prefectures in the short run and long run. The skill upgrading probability for a typical prefecture i is defined as the ratio of total number of skilled workers in the next period originating from i over the total number of unskilled workers in the current period. Each dot represents a prefecture. The blue dots come from the counterfactual simulation with capital accumulation, and the red dots are based on the counterfactual simulation without capital accumulation. The straight lines are linear fits.

4.3 The Role of Capital Accumulation

We have found significant differences in the impacts of trade between the short and long run. In this part, we explore the role of capital accumulation in driving the pattern. To do so, we shut down the capital accumulation and re-simulate the trade impacts.

Capital accumulation plays an important role in diverting population to coastal cities. Our results in Figure 6 show that the distance elasticity is only 0.069 in the long run without capital accumulation. In particular, capital accumulation explains around 72 percent of the distance elasticity of the population change (from 0.069 to 0.248). Figure 7 shows that the distance elasticity of skill ratio barely changes from the short to the long run and is significantly positive in the economy without capital accumulation. This suggests that capital accumulation and capital-skill complementarity serve as the mechanism to drive skilled workers into coastal cities. Capital accumulation also significantly alleviates the decline in the skill premium. Figure 8 shows that skill premium drops in 195 cities out of 196 in China in the long run if we shut down capital accumulation, and those in coastal cities drop the most.



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

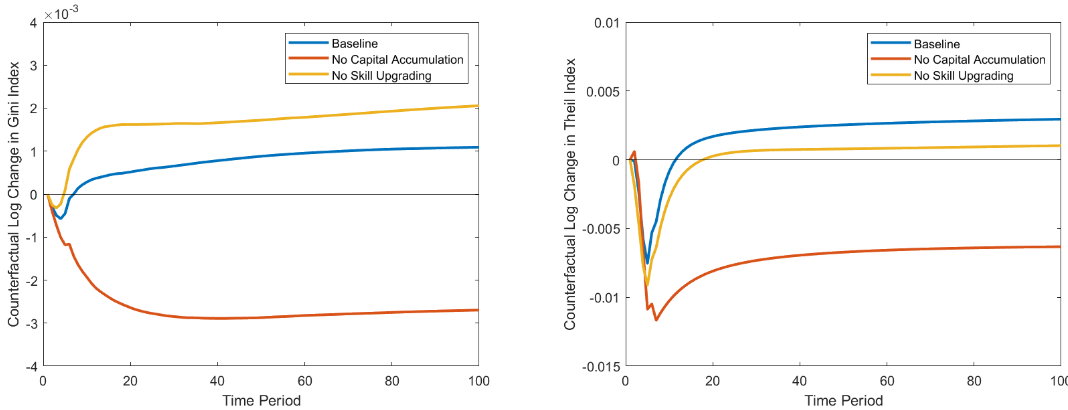
Figure 10: The Distance Elasticities over Time

Notes: this figure shows the time series of distance elasticities of population change, skill ratio change, and skill premium change respectively in each panel. In each panel, the elasticity is estimated for each time period by regression of changes in outcome variable on trade costs for 196 prefectures. The blue line is the elasticity result from the baseline quantification of trade impacts; the red line corresponds to simulation with constant capital stock in each location; the yellow line corresponds to simulation with infinite skill upgrading cost (no skill upgrading).

To keep track of the distance elasticities over time, Figure 10 plots the evolution of the distance elasticities for both the baseline and the economy without capital accumulation for 100 periods. Panel (a) presents the dynamics of distance elasticity of changes in population. As capital accumulates over time, the gap between the baseline and the economy without capital accumulation widens. Panel (b) in figure 10 shows the distance elasticity of skill ratio reverses the sign from baseline to the economy without capital accumulation, and their gap widens over time. Panel (c) shows that the distance elasticity of changes skill premium turns negative in about 40 periods in the baseline economy, while it remains positive in the economy without capital accumulation.

Spatial Inequality Existing literature finds that trade liberalization overall increases spatial inequality in China (Han et al. (2012), Fan (2019)). Our quantitative results instead find trade liberalization increases spatial inequality as a result of capital accumulation. As shown in figure 11, the trade shock slightly increases spatial inequality, measured in Gini index and Theil index over the spatial distribution of wages. In the steady state, the trade shock increases spatial Gini index by 0.14 percent and Theil index by 0.40 percent.

Comparing the baseline result with that without capital accumulation, we conclude that capital accumulation is central in driving up spatial inequality. If we shut down capital accumulation, trade liberalization actually decreases spatial inequality in China. With capital accumulation, trade liberalization benefits coastal cities more. Given that coastal cities are already relatively richer than inland cities, this enlarges spatial inequality in China. To see this in more detail, we follow the same definition of coastal cities as before and consider the welfare impacts of trade across space. The steady-state welfare gain from trade liberalization for unskilled workers is 0.40 percent in coastal cities, compared to 0.33 percent in inland cities; the welfare gain for skilled workers is 0.30 percent in coastal cities, versus 0.21 percent in inland cities.



(a) Change in Spatial Gini Index

(b) Change in Spatial Theil Index

Figure 11: Impacts of Trade on Spatial Inequality

Notes: This figure shows the effects of trade liberalization on spatial inequality at the prefecture level. It plots over time the log change of the index from baseline simulation to the counterfactual without trade liberalization. The Gini index and Theil index are calculated based on wages across prefectures.

[:later] We also evaluate how trade's impact on population mobility depends on skill upgrading. Figure 10 suggests that the elasticity is even higher if we do not allow workers to upgrade their skill levels. Over time, Similar to before, the gap is even larger without skill upgrading, suggesting that skill upgrading counteracts the effects of capital accumulation.

4.4 Robustness Check

Cross-location investment In the baseline economy, landlords can only invest in their own locations. This assumption may amplify the role of initial capital stock in driving our results, as the landlords in the remote locations cannot directly invest in the booming coastal ones. In this section, we relax the assumption by allowing landlords to invest in all the locations. We find that our main results still hold.

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. The rate of return to capital in host location n, $r_{n,t}$, is drawn 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

$$K_{it} = \frac{(r_{it})^{\upsilon}}{\sum_{n=1}^{N}(r_{nt})^{\upsilon}} K_t, \tag{21}$$

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:

$$R_t = \Gamma\left(\frac{\upsilon - 1}{\upsilon}\right)\left[\sum_{n=1}^{N}(r_{nt})^{\upsilon}\right]^{\frac{1}{\upsilon}} \tag{22}$$

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

$$k_{it+1} = \xi\beta\left(1 - \delta + \frac{R_t}{p_{it}}\right)k_{it}. \tag{23}$$

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

When the landlord can invest in other locations, each location's total income is no longer equal to its total output, as total output depends on the capital inflow and 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

$$X_{it}^{j} = \sum_{n=1}^{N} S_{nit}^{j}\left(\gamma_j I_{nt}\right), \tag{24}$$

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

$$r_{it} = \frac{\sum_{j=1}^{J} \phi_{it}^{kj} X_i^j}{k_{it}}. \tag{25}$$

We recompute the distance elasticities of population, skill ratio, and skill premium over time for the extended model. Table xx presents the results. The distance elasticity of population in absolute terms increases from 0.010 in the short run to 0.049 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 ratio 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 again driven by capital accumulation. When we shut down capital accumulation, those elasticities barely change 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 0.049 with cross-location investment, compared to 0.248 in the benchmark. 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, we find that locations with relatively larger capital stocks in the baseline now have lower capital stocks (capital inflows), and those with smaller capital stocks in the baseline now have larger capital inflows. Therefore, cross-location investment weakens coastal locations' advantage and counteracts the positive effect on capital accumulation led by trade liberalization. Nevertheless, our main results are still robust.

Alternative trade costs We next show that our main results are robust to an alternative measure of reduction in international trade costs. Instead of using the reduced-form estimates of trade costs between China's port cities with the ROW, we calibrate these trade costs in the transition path. We aim to calibrate the trade costs in the year 2006 so that the model-generated trade-to-output ratio for China is consistent with the trade-to-GDP ratio observed in the data for the year 2006. 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. 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.

We conduct the same counterfactual analysis as before to quantify the impacts of trade with or without capital accumulation. Panel (a) in table 4 shows the results. From the short to long run, we again find a large increase in distance elasticity of population (from -0.053 to -0.432) and reversed signs of elasticities of skill ratio (from 0.006 to -0.038) and skill premium (from 0.011 to -0.019). These large changes over time are driven by capital accumulation, as suggested in the same panel. The changes in elasticity are also slightly larger than those in the benchmark because the estimated trade shocks are larger, and thus, capital accumulation plays a more significant role. Larger trade shocks bring more advantages to coastal cities in the long run, increasing the long-run elasticity.

No Skill-productivity complementarity In the baseline model, we introduce skill-productivity complementarity: highly productive locations are more efficient in employing skilled workers. We introduce it mainly to match the data counterpart, but this assumption could potentially influence our main results. As many coastal cities in China have relatively higher productivity, they are more attractive for skilled workers due to skilled-biased productivity. In view of this concern, we re-simulate the results assuming no skill-productivity complementarity. We simply assume the skilled-biased productivity elasticity $\psi = 0$.

Table 4 panel (b) shows the simulation results. While taking away skill-productivity complementarity weakens the role of capital accumulation in trade impacts, the distance elasticities are still statistically significant and their changes over time are consistent with our benchmark results.

No Skill Upgrading Skill upgrading is one of our key mechanisms shaping skill premium and the spatial impacts of trade. As a robustness test, we consider an alternative scenario in which we assume the skill upgrading cost now is infinity. As a result, there is no skill upgrading, and the total supply of skills is fixed over time. Overall, assuming no skill upgrading strengthens the effects of capital accumulation. The long-run distance elasticities of population and of skill ratio are larger than those obtained from the benchmark economy. Intuitively, as skill supply now becomes scarce, locations with higher capital stocks provide higher marginal returns for skilled workers than the benchmark counterpart, thus attracting more skilled workers.

Same Production Function In the baseline economy, we allow for different production technologies for China and the ROW, so the trade pattern is pinned down by both technological differences and relative factor endowments between China and ROW. To isolate one force from the other, in this part, we allow China and the ROW to have the same production function, determined by income shares of each factor from China's data. In this way, the trade pattern is solely determined by relative factor endowments. We recalibrate the production function together with other parameters in the equilibrium.

Panel (d) in table 4 reports the results. The distance elasticities increase significantly from the short to the long run. These changes are largely due to capital accumulation.

5 Conclusion

In this study, we develop a rich dynamic spatial framework to study skill premium and the spatial impact of trade liberalization. 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 economy. In quantitative analysis, we find that the spatial impacts of trade crucially depend on the time horizon and capital accumulation. Most importantly, we find that, given the trade shock, in the long run capital accumulation drives much more workers towards coastal cities, particularly skilled workers, and it reverses the Stolper-Samuelson effect across space so that skill premiums in coastal cities increase but those in inland cities decrease.

Table 4: Distance Elasticities under Different Robustness Checks

	Full Model			W/o Capital Accu.			W/o Skill Upgrading		
	SR	LR	Change	SR	LR	Change	SR	LR	Change
Baseline									
Population	-0.035	-0.248	6.001	-0.037	-0.069	0.897	-0.050	-0.320	5.410
Skill Ratio	0.002	-0.021	-15.032	0.017	0.032	0.855	-0.002	-0.047	18.858
Skill Premium	0.003	-0.011	-4.454	0.010	0.004	-0.573	0.007	-0.009	-2.240
Mobile Capital									
Population	-0.010	-0.045	3.564	-0.009	-0.013	0.346	-0.011	-0.043	2.989
Skill Ratio	-0.004	-0.011	1.643	0.004	0.007	0.586	-0.003	-0.014	3.274
Skill Premium	-0.001	-0.003	3.677	0.004	0.001	-0.867	0.000	-0.002	-12.242
Calibrated Trade Cost									
Population	-0.053	-0.432	7.202	-0.058	-0.114	0.972	-0.074	-0.540	6.291
Skill Ratio	0.006	-0.038	-7.856	0.026	0.049	0.893	-0.001	-0.082	70.337
Skill Premium	0.011	-0.019	-2.662	0.018	0.012	-0.309	0.019	-0.015	-1.831
No Skill-Productivity Complementarity									
Population	-0.027	-0.141	4.262	-0.027	-0.050	0.816	-0.041	-0.196	3.770
Skill Ratio	0.004	-0.004	-1.926	0.018	0.029	0.625	0.002	-0.016	-8.353
Skill Premium	0.003	-0.004	-2.348	0.010	0.005	-0.546	0.008	-0.004	-1.545
Same Technology									
Population	-0.040	-0.279	5.891	-0.035	-0.069	0.983	-0.046	-0.314	5.813
Skill Ratio	-0.022	-0.043	0.949	-0.010	0.004	-1.421	-0.020	-0.085	3.284
Skill Premium	-0.008	-0.017	1.041	-0.003	-0.000	-0.911	-0.010	-0.012	0.130

Notes: This table reports 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 'Change' reports the changes in elasticities from short run to long run, calculated as $\frac{elasticity\ in\ LR}{elasticity\ in\ SR} - 1$.

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.

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," 2022.

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.

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," 1996.

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

Han, Jun, Runjuan Liu, and Junsen Zhang, "Globalization and wage inequality: Evidence from urban China," *Journal of international Economics*, 2012, *87* (2), 288–297.

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.

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," 2022.

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.

Wu, Jidong, Li Ning, and Shi Peijun, "Benchmark wealth capital stock estimations across China's 344 prefectures: 1978 to 2012," *China Economic Review*, 2014, *31*, 288–302.

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*, 2004, *10* (1), 35–44.

Online Appendix

A Details of the Model

A.1 Price and Trade

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

$$p_{it} = \prod_{j=1}^{J} \left(p^j_{it}\right)^{\gamma^j}, \tag{26}$$

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

$$p^j_{it} = \left[\sum_{n=1}^{N} \left(p^j_{in,t}\right)^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{27}$$

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

$$\pi^j_{in,t} = \frac{(p^j_{in,t})^{-\theta}}{\sum_{m=1}^{N}(p^j_{im,t})^{-\theta}}. \tag{28}$$

A.2 Firm's Problem

For firm's problem we drop the time notation for brevity. The problem of a variety producer in sector j at location i is given by

$$\min_{l,s,k} \; w^l_i l^j_i + w^s_i s^j_i + r_i k^j_i \tag{29}$$

subject to

$$z_i \left[(\mu^j)^{\frac{1}{\sigma}} (z_i^{-\psi} l_i^j)^{\frac{\sigma-1}{\sigma}} + (1-\mu^j)^{\frac{1}{\sigma}} (z_i^{\psi} h_i^j)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \geq q_i^j \tag{30}$$

$$h_i^j = \left[(\lambda^j)^{\frac{1}{\eta}} (k_i^j)^{\frac{\eta-1}{\eta}} + (1-\lambda^j)^{\frac{1}{\eta}} (s_i^j)^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}. \tag{31}$$

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

$$k_i^j = \frac{\lambda_j}{1-\lambda_j} \left(\frac{w_i^s}{r_i} \right)^{\eta} s_i^j. \tag{32}$$

Using this expression to replace k_i^j in equation (31) and define the price w_i^{hj} for composite input h_i^j such that $w_i^{hj} h_i^j = r_i k_i^j + w_i^s s_i^j$, we obtain

$$w_i^{hj} = \left[\lambda^j (r_{it})^{1-\eta} + (1-\lambda^j)(w_{it}^s)^{1-\eta} \right]^{\frac{1}{1-\eta}}. \tag{33}$$

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

$$l_i^j = \frac{\mu_j}{1-\mu_j} \left(\frac{w_i^{hj}}{w_i^l} \right)^{\sigma} z_i^{-2\psi(\sigma-1)} h_i^j \tag{34}$$

Using equation (34) to replace l_i^j in equation (30) 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

$$c_i^j = \frac{1}{z_i} \left[\mu^j (z_i^{\psi} w_i^l)^{1-\sigma} + (1-\mu^j)(z_i^{-\psi} w_i^h)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{35}$$

A.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

$$v_i^{d*} = \ln b_i^* + \ln \frac{w_i^{d*}}{p_i^*} + \rho \ln \sum_{g=1}^{N} \exp\left[(\xi \beta v_g^{d*} - \kappa_{gi}^d - \kappa_e^d)/\rho \right] \tag{36}$$

$$D_{ig}^{ed*} = \frac{\exp\left[(\xi\beta v_g^{d*} - \kappa_{gi}^d - \kappa_e^d)/\rho\right]}{\sum_{n=1}^{N}\exp\left[(\xi\beta v_n^{d*} - \kappa_{ni}^d - \kappa_e^d)/\rho\right]} \tag{37}$$

$$L_i^{l*} = \xi\sum_{n=1}^{N} D_{in}^{ll*}L_n^{l*} + \left(L_i^{l*} + L_i^{s*}\right)(1-\xi) \tag{38}$$

and

$$L_i^{s*} = \xi\left(\sum_{n=1}^{N} D_{in}^{ss*}L_n^{s*} + \sum_{n=1}^{N} D_{in}^{sl*}L_n^{l*}\right), \tag{39}$$

the market clearing conditions (24)(25)(13)(14) become

$$X_i^{j*} = \sum_{n=1}^{N} S_{ni}^{j*}\left[\gamma_j\sum_{m=1}^{J} X_n^{m*}\right], \tag{40}$$

$$w_i^{l*} = \frac{\sum_{j=1}^{J}\phi_i^{lj*}X_i^{j*}}{L_i^{l*}} \tag{41}$$

$$w_i^{s*} = \frac{\sum_{j=1}^{J}\phi_i^{sj*}X_i^{j*}}{L_i^{s*}} \tag{42}$$

$$r_i^{*} = \frac{\sum_{j=1}^{J}\phi_i^{kj*}X_i^{j*}}{k_i^{*}}, \tag{43}$$

with steady-state factor income shares given by

$$\phi_i^{lj*} = \left[1 + (z_{it}^*)^{2\psi(\sigma-1)}\,\frac{1-\mu^j}{\mu^j}(\frac{w_i^{l*}}{w_i^{h*}})^{\sigma-1}\right]^{-1} \tag{44}$$

$$\phi_i^{sj*} = \left[1 + (z_{it}^*)^{-2\psi(\sigma-1)}\,\frac{\mu^j}{1-\mu^j}(\frac{w_i^{h*}}{w_i^{l*}})^{\sigma-1}\right]^{-1}\left[1 + \frac{\lambda^j}{1-\lambda^j}(\frac{w_i^{s*}}{r_i^{*}})^{\eta-1}\right]^{-1} \tag{45}$$

$$\phi_i^{kj*} = \left[1 + (z_{it}^*)^{-2\psi(\sigma-1)}\,\frac{\mu^j}{1-\mu^j}(\frac{w_i^{h*}}{w_i^{l*}})^{\sigma-1}\right]^{-1}\left[1 + \frac{1-\lambda^j}{\lambda^j}(\frac{r_i^{*}}{w_i^{s*}})^{\eta-1}\right]^{-1}, \tag{46}$$

and trade share given by

$$S_{ni}^{j*} = \frac{(p_{ni}^{j*})^{-\theta}}{\sum_{g=1}^{N}(p_{nm}^{j*})^{-\theta}}, \tag{47}$$

where

$$p_{ni}^{j*} = \frac{\tau_{ni}}{z_i^*}\left[\mu^j\left(z_{it}^*\right)^{\psi(1-\sigma)}(w_i^{l*})^{1-\sigma} + (1-\mu^j)\left(z_{it}^*\right)^{-\psi(1-\sigma)}\left[\lambda^j(r_i^*)^{1-\eta} + (1-\lambda^j)(w_i^{s*})^{1-\eta}\right]^{\frac{1-\sigma}{1-\eta}}\right]^{\frac{1}{1-\sigma}}, \tag{48}$$

and the capital accumulation condition becomes

$$k_i^* = \xi\beta(1-\delta+\frac{r_i^*}{p_i^*})k_i^*, \text{ with } p_i^* = \prod_{j=1}^{J}\left[\sum_{n=1}^{N}(p_{in}^{j*})^{-\theta}\right]^{-\frac{\gamma_j}{\theta}}. \tag{49}$$

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 (37), and then solve new labor allocations by (38) and (39) 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 (44), (45), (46),
- (c) compute prices $\{p_{ni}\}_{n=1,i=1}^{N,N}$ and trade shares $\{S_{ni}\}_{i=1,n=1}^{N,N}$ from (48) and (47),
- (d) solve total output X_i^{j*} by (40)
- (e) obtain new factor prices $\{w_i^l, w_i^s, w_i^k\}_{i=1}^{N}$ by (41), (42), (43),
- (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 (49).

(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 (36).

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

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

A.4 Numerical Algorithm for Solving Path Equilibrium

Here we only solve a transition path of the economy with sufficiently long time periods T, although the equilibrium is defined over an infinite time horizon. Given the initial allocations of labor and capital, $\{l_{i0}, s_{i0}, k_{i0}\}$, the path equilibrium is uniquely determined and the algorithm is 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 (28),
- (d) solve total output by (24),
- (e) solve new factor prices by (25), (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 (27) 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 $\{v_{it}^{l(0)}, v_{it}^{s(0)}\}_{i=1,t=1}^{N,T}$ and get $\{v_{it}^{l(1)}, v_{it}^{s(1)}\}_{i=1,t=1}^{N,T}$.

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

B Details of Data and Quantification

B.1 Data Sources for China

1. The **2000 Census** and **2010 Census** in China. These datasets provide prefecture-level population and skill ratios at year 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 detailed classification of 96 industries at two-digit level. We exclude industries in agriculture and mining and the waste processing industry, resulting in a total number of 82 industries.

   We further classify these industries into four sectors based on industrial skill intensities: skilled manufacturing sector, unskilled manufacturing sector, skilled service sector, and unskilled service sector. Specifically, we compute skill intensity of each industry by taking the ratio of skilled workers' income to total income for each industry. Then we rank manufacturing and service industries seperately by skill intensity. We treat industries above the median skill intensity as the *skilled industries* and group them to define the skilled sector. Those below the median skill intensity are aggregated as unskilled sector.

3. The **One Percent Population Survey** in 2005. We use this dataset to obtain prefecture-level bilateral migrants stocks at year 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 at year 2005.

4. The **City Statistical Yearbooks** of China, from which we obtain prefecture-level GDPs at year 2000 and 2009, 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 at year 2000.

5. The **2002 China Input-Output Table**. This table provide final consumption and capital income shares in the value added for 42 industries at two digits. We exclude agricultural and mining industries and manually map the remaining 37 industries with the 82 industries in GB/T 4754-2002 classification (GB hereafter) so that the industry classification is consistent. Usually the 2002 IO table industries each contains multiple GB industries. Since we define skilled sectors based on GB system, for the 2002 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 that industry than unskilled GB one(s). Then we aggregate IO Table industries into four sectors by skill intensity and compute the corresponding sectoral capital income shares as the ratio of total sectoral capital income to sectoral vaule added.

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

1. The **OECD Statistics**. This database provides the initial population aged from 25 to 64 at year 2000 for 33 countries including China. We distribute this total population of China to prefectures using prefectural population shares calculated from 2000 China census. We aggregate the remaining contries' population as the population of the ROW. From the same database, we also observe country-level shares of unskilled workers 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 it's total workers.

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

3. The **World Input-Output Database (WIOD)**. We use the WIOD 2016 Release to obtain sectoral trade-to-GDP ratios of China from 2000 to 2006. The World Input Output Tables provide intercountry trade flows for 56 industries including 19 manufacturing industries. We only consider trade in manufacturing sector and trade flows between China and 32 countries included in the ROW. To obtain sectoral imports and exports of China, we map manually the 19 manufacturing industries with 16 Chinese manfacturing industries in 42 industry classfication and define skilled and unskilled sector.

   From the WIOD, we also use National IO tables of China 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.

4. The **IPUMS USA**. We use the one-percent sample of the U.S. 2000 Census from IPUMS USA to obtain sectoral skilled workers' income share. We define skilled worker as workers with education level at least 12 grade, i.e. high school graduates or college graduates. To obtain sectoral income shares consistent with China's sector, we match China's 42 IO table industries with NAICS 2007 code. Specifically, we match China's manufacturing with 3-digit NAICS code, utility industries with 4-digit NAICS code, and service industries with 2-digit NAICS code. 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.

5. The **2007 Benchmark Input-Output Acount** of the U.S. provides sectoral capital income share in the total value added. Again we match China's 42 industries with NAICS 2007 code. The labor income share is simply the income share other than from capital and is distributed to skilled workers and unskilled workers using result from IPUMS USA. We use these income shares of the U.S. to represent income shares of the ROW and to calibrate the ROW's production function.

B.3 Estimate Migration Costs

In this part, we ignore workers' skill upgrading choices. 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

$$\bar{D}^d_{gi,t} = \frac{L^d_{it} D^d_{gi,t} + \sum_{\tau=1}^{\infty} L^d_{it-\tau} D^d_{gi,t-\tau} (D^d_{gg,t-\tau})^{\tau}}{L^d_{it}}, \ d = l, s. \tag{50}$$

In the numerator on the left-hand side, the first term is migration flow from the origin i at period t, and the second term is period t's total migrants moving from location i previously and staying at the current destination location g. Assume that the migrant stocks are observed at a steady state, then this ratio becomes

$$\bar{D}^d_{gi} = \frac{D^d_{gi}}{1 - D^d_{gg}}, \tag{51}$$

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

$$D^d_{gi} = \frac{\exp\left[(\beta v^d_g - \kappa^d_{gi})/\rho\right]}{\sum_{n=1}^{N} \exp\left[(\beta v^d_n - \kappa^d_{ni})/\rho\right]}. \tag{52}$$

Therefore, double differencing the migrant stock share yields our main structural equation:

$$\frac{\bar{D}_{gi}}{\bar{D}_{ii}} \frac{\bar{D}_{ig}}{\bar{D}_{gg}} = \frac{D^d_{gi}}{D^d_{ii}} \frac{D^d_{ig}}{D^d_{gg}} = \exp\left[-\frac{1}{\rho}(\kappa^d_{gi} + \kappa^d_{ig})\right], \tag{53}$$

where we use the result

$$\frac{D^d_{gi}}{D^d_{ii}} = \frac{\exp\left[(\beta v^d_g - \kappa^d_{gi})/\rho\right]}{\exp\left[(\beta v^d_i)/\rho\right]} = \exp\left\{\left[\beta\left(v^d_g - v^d_i\right) - \kappa_{gi}\right]/\rho\right\}. \tag{54}$$

B.4 Additional Tables

Table 5: Manufacturing Sectors

Panel A: Unskilled Manufacturing		Panel B: Skilled Manufacturing	
Classification No.	Description	Classification No.	Description
C13	Food Processing	C15	Beverage
C14	Food Manufactures	C16	Tabacco Manufactures
C17	Textile	C25	Petroleum, Coke and Nuclear Fuel
C18	Clothing, Shoes and Hats	C26	Chemicals
C19	Leather, Hide and Feather Manufactures	C27	Medicinal and pharmaceutical Products
C20	Wood Processing	C28	Chemical Fiber Manufactures
C21	Furniture	C32	Iron and Steel
C22	Pulp and Paper	C33	Non-ferrous Metals
C23	Printing	C35	General Equipment
C24	Educational and Sporting Products	C36	Specialized Equipment
C29	Rubber Manufactures	C37	Transport Equipment
C30	Plastics Manufactures	C39	Electrical machinery, apparatus and appliances
C31	Non metallic mineral manufactures	C40	Telecommunication and Computer Manufactures
C34	Manufactures of Metal	C41	Instrument and Office Equipment
C42	Handicrafts and Other Manufactures		

Table 6: Service Sectors

Panel A: Unskilled Service		Panel B: Skilled Service	
Classification No.	Description	Classification No.	Description
D45	Gas production and supply	D44	Electricity and Heat Production and Supply
D46	Water production and supply	G60	Information Transmission Service
E47-E50	Construction	G61	Computer Service
F51-F59	Transportation	G62	Software Service
H63	Wholesale	J68-J71	Financial Service
H65	Retail	K72	Real Estate
I66-I67	Accommodation and Catering	L73-L74	Leasing and Business Services
N80	Environment Management	M75-M78	Scientific research, technical services and geological prospecting
N81	Public Facilities Mangement	N79	Water Management
O82-O83	Residential and Other Service	P84	Education
		Q85-Q87	Health, social security and social welfare
		R88-R92	Culture, sports and entertainment
		S93-S97	Public Administration and Social Organization

Table 7: 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			

Table 8: List of Port Cities

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