Internal Geography, Labor Mobility, and the Distributional Impacts of Trade*

Jingting Fan†
Pennsylvania State University

Abstract

This paper develops a spatial equilibrium model to quantify the distributional impacts of *international* trade in an economy with *intra-national* trade and migration costs. Focusing on China, I find that international trade increases both the between-region inequality among workers with similar skills and the within-region inequality between skilled and unskilled workers, with the former accounting for 75% of the overall inequality increase. Ignoring domestic spatial frictions will significantly underestimate trade's impact on the overall inequality and overestimate its impact on the aggregate skill premium. Domestic trade and Hukou reforms can improve welfare and alleviate trade-induced inequality, while at the same time reduce the share of international trade in the economy.

1 Introduction

In recent decades we have witnessed increasing integration of large developing countries, such as Brazil, China, India, and Mexico, into global trade. This trend has renewed the interest of policymakers and academics in understanding the aggregate and (especially) distributional effects of globalization.

*First Draft: 2014. I am grateful to Nuno Limao for his continuous guidance and support, and to Rafael Dix-Carneiro for extensive discussions and in-depth comments at various stages of this project. For helpful comments, I also thank Marisol Chatruc, Kerem Cosar, Lorenzo Caliendo, Pablo D'Erasmo, Jonathan Eaton, Wenlan Luo, Peter Morrow, Felipe Saffie, Lixin Tang, and participants at the seminar at the University of Maryland, 2014 Midwest International Trade Meetings, 2015 Urban Economic Association Annual Meetings, 2015 Georgetown Center for Economic Research Biennial Conference. I thank the University of Maryland Graduate School Summer Research Fellowship for financial support. All errors are my own.

†Kern Building, Pennsylvania State University, University Park, PA, 16802. Email address: jxf524@psu.edu

In developing countries with poor domestic infrastructure and limited inter-regional worker mobility, such distributional impacts might have an important geographic dimension. Consider workers living far away from a nation's ports. Because of the high *intra-national* trade costs, they might not benefit much from cheaper imported products, and tariff cuts at the border can exacerbate the intra-national inequality in living standards. Moreover, in a world with both skilled and unskilled workers, if one type of workers is more mobile and responds to trade liberalization by migrating to the coast, then the workers left behind might even lose from trade. These losses can be independent of regional sectoral specializations. This geographic margin in the distributional impacts of trade is not only plausible, but also empirically relevant.[1]

With a focus on China, this paper answers two questions arising naturally from the scenario discussed above. First, in the presence of *intra-national* trade and migration costs, how does *international* trade liberalization affect within-country inequality—including both the between-region inequality among workers of similar skill levels, and the within-region inequality between skilled and unskilled workers (the skill premium)? Second, many developing countries are investing in transportation infrastructure and launching structural reforms, with the aim of reducing the within-country spatial frictions. To what extent would these changes affect domestic welfare and our answer to the first question?

The coexistence of rapid trade growth and large spatial inequality makes China a useful setting for this study. As is well known, China has experienced rapid integration into world trade since its economic reform in 1978, and the process accelerated after its WTO accession in 2001. At the same time, China has had historically high intra-national trade costs and strict controls on worker migration through the Hukou system. Perhaps partially due to these spatial frictions, China's economic growth over the past decades has been uneven. Indeed, as shown in Figure 1, inter-regional inequality grew rapidly during the period of fast trade expansion in China.

To answer the questions posed, I develop and quantify a spatial equilibrium model of trade (Redding, 2016). Regions in the model represent Chinese cities and the rest of the world (RoW), and are connected through costly trade and migration. The model allows international trade to affect within-country inequality through both geographic and skill dimensions. Because of domestic trade costs, it is more costly for cities in the interior

[1] Limão and Venables (2001) and Coşar and Demir (2016) document that domestic transportation infrastructure affects a country's participation in international trade; Atkin and Donaldson (2015) estimates the intra-national trade costs to be 4-6 times larger in their sample of African countries than in the United States. Topalova (2010) shows that in India, trade liberalization hurt the poorest workers because of their limited inter-regional and inter-sectoral mobility. See also Kanbur and Venables (2005) for an excellent overview of the UNU-Wider project on "Spatial disparities in development," which analyzes evidence in over 50 developing countries, and concludes that international trade and the lack of infrastructure are two important factors in the increasing spatial disparities in many of these countries.

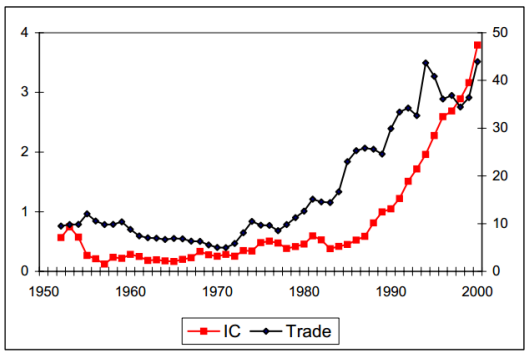


Figure 1: Trade and Growth Inequality in China
Source: Figure 5 in Kanbur and Zhang (2005). IC (left axis) is the measure of inland-coast inequality in China, Trade (right axis) is defined as (import+export)/GDP

to trade with the rest of the world, so international trade affects regions differently. To capture the effect of international trade on regional and aggregate skill premia, the model incorporates two channels emphasized in the literature: the factor content of trade dating back to Stolper and Samuleson, and trade in capital goods and capital-skill complementarity (Burstein et al., 2013; Parro, 2013). Due to domestic spatial frictions, these channels will also have differential impacts across regions. On the worker side, workers decide where to live according to potential utility in all destinations, which in turn depends on regional prices and wages. The endogenous migration of workers will prove important in shaping how trade affects skilled and unskilled workers from different parts of China.

I parameterize the model using rich micro and macro data from China, including domestic trade and migration information, international trade statistics, and the distribution of production across sectors and space. The model is able to match untargeted moments on heterogeneous changes in migration and skill premia across cities in response to trade liberalization. My estimation reveals large barriers to migration. Of course, not all of them are created by the Hukou system. I construct a city-level data set of partial Hukou reforms since 1997, and then exploit the over-time variation in Hukou openness to separate the component of migration costs created by the Hukou system—and thus amenable to policy reforms—from the component arising from workers' home bias or other policy distortions. I find that the Hukou system creates substantial mobility costs and can explain two-thirds of the higher migration costs in China compared to the U.S. In some counterfactual experiments, I will remove the remaining Hukou component of migration costs for all cities to analyze how a comprehensive Hukou reform affects welfare and international trade.

To examine the impacts of international trade, I shut down trade between the model China and the RoW. The average gains from trade are around 7.5%. These gains, however,

are distributed unevenly: skilled workers gain 13% on average, while unskilled workers gain only 6%. The average percentage wage difference between skilled and unskilled workers, or the aggregate skill premium, increases by 5%. The impacts also differ among workers with similar skills from different locations—regions on the coast reap most of the welfare gains while regions in the interior benefit little. Aggregate inequality, as measured by the Theil index, increases by 6.8% after the international trade liberalization. The geographic dimension—the increase in inequality between geographic regions—accounts for 75% of the increase in overall inequality, while the skill dimension—the increase in within-region inequality—accounts for the rest.

Consistent with existing reduced-form evidence (Han et al., 2012), the geographic dimension interacts with the skill dimension: skill premia increase more in coastal regions. In addition to capital-skill complementarity, which increases skill premia more in the coastal regions because these regions import more, two more forces are behind this result. First, because capital and other manufacturing industries use intermediate varieties more intensively, they tend to locate in regions with better access to suppliers. After the trade liberalization, coastal regions experience a larger increase in access to foreign suppliers and, as a result, become more specialized in capital and manufacturing industries. Because these industries are also more skill intensive, this change in specialization *within* China increases skill premia in the coast and decreases it in the interior. Second, because the estimated migration costs are lower for skilled workers, more skilled workers respond to trade liberalization by migrating from the hinterland to the coast. This channel decreases skill premia on the coast and increases skill premia in the interior, offsetting the previous forces. I show that all three channels are quantitatively important and all of them exist only because of domestic spatial frictions. Incorporating the internal geography of a country is thus relevant for the distributional impacts of trade along not only the geographic but also the skill dimensions.

I calibrate a similar model with free domestic trade and migration to match the same trade and production patterns, and then shut down international trade in this model. This experiment finds similar average welfare gains to the benchmark model, but has very different implications for measures of the aggregate inequality. The increase in the aggregate Theil index is 4.7%, significantly below the benchmark model (6.8%), in which inter-regional inequality drives up the aggregate inequality. On the other hand, the frictionless model predicts the aggregate skill premium to increase by 12%, much larger than the 5.5% prediction of the benchmark model. This comparison highlights the role of domestic spatial frictions for understanding the aggregate inequality changes after trade liberalization.

In recent years, China invested heavily in domestic infrastructure and has started re-

forms aiming at reducing barriers to domestic trade and migration. To quantify the effects of these reforms, I reduce trade and migration costs within China, and then first compare the post-reform economies with the calibrated benchmark economy. To make sure that the decreases in these frictions are attainable by policies, in the case of domestic trade, I use the U.S. as the benchmark for the post-reform level of "internal borders"; in the case of Hukou reforms, I use the estimated effect of Hukou openness on migration costs and assume that after the abolishment of Hukou system, the inward migration cost to all cities will be to reduced to a level corresponding to the highest Hukou openness. I find that both domestic trade and Hukou reforms generate large welfare gains, but they affect inequality differently. Lowering domestic trade costs leads to a modest increase in the aggregate inequality, primarily by increasing skill premia. In contrast, abolishing the Hukou system will reduce the aggregate inequality, mainly through decreasing inter-regional inequality.

To understand how these reforms interact with the distributional impacts of trade, I move the post-reform economies to autarky and calculate the changes. Compared to the benchmark model, the changes in inequality are smaller—these reforms indeed help spread the gains from trade more evenly. However, the post-reform economies participate less intensively in, and therefore benefit slightly less from, international trade. This happens for two reasons. First, as China becomes more integrated from the within through reforms, its economy expands relative to the RoW. China's terms-of-trade deteriorate and, as a result, it trade less intensively with the RoW. Second, lower domestic trade costs tend to divert trade between coastal regions and the RoW to interior regions, reducing the volume of international trade. Because the Hukou reform is not affected by the second channel, it is able to spread the gains from trade more evenly without sacrificing much the overall gains.

This paper is most closely related to a broad literature on the impacts of trade on inequality. While existing studies have analyzed this topic from different angles, the inequality between skill and unskilled workers is most emphasized (see, Goldberg and Pavcnik, 2007, for a review). This paper studies the distributional impacts of trade along spatial and skill dimensions jointly, and finds that the former plays a more important role for the overall inequality increase after trade liberalization. As both locations and skills are observable, understanding their relative importance can help better design inequality-alleviating policies. Further, by showing that models without domestic spatial frictions will significantly overestimate the effects of trade on the skill premium, this study complements existing quantitative work, such as Burstein et al. (2013) and Parro (2013), which focuses only on the skill premium.

A number of recent papers have also used a spatial equilibrium model to study trade

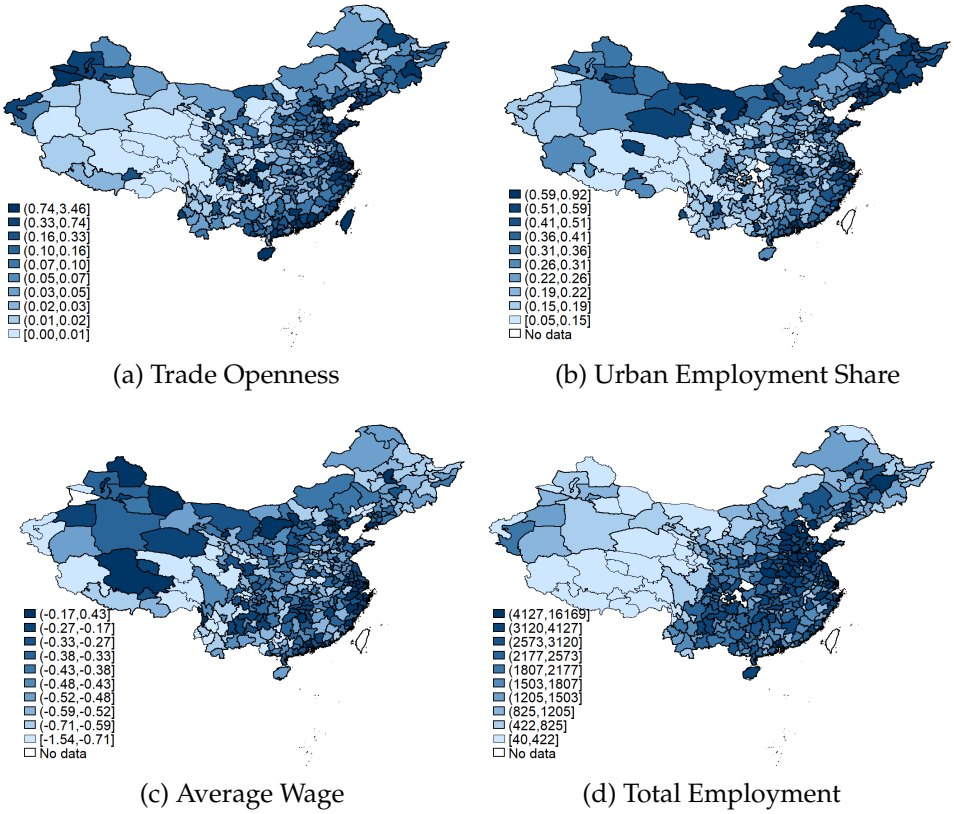
and/or inter-regional labor mobility (see, for example, Monte, 2016; Allen and Arkolakis, 2014; Ramondo et al., 2016; Caliendo et al., 2017; Bryan and Morten, 2018; Tombe and Zhu, 2017; Galle et al., 2017).[2] Most of these papers do not focus on how trade affect domestic inequality and moreover, do not differentiate between skill and unskilled workers. As a result, these papers are silent on the importance of skill and geographic dimensions, or the interactions between the two.[3] An additional difference of my paper is that it relates the gains from trade to city characteristics—their distance to port. This channel arises naturally in a setting with domestic trade costs and generalizes to developing countries lacking infrastructures, but has not been examined in these papers.

This paper also contributes to the literature on China's spatial economy (Poncet, 2005; Au and Henderson, 2006). The closest paper in this literature is Tombe and Zhu (2017). Relative to Tombe and Zhu (2017) the present paper differs in two important aspects. First and for most, the question is different. While their paper focuses on how trade and migration costs affect labor productivity in China without differentiating workers' skills (and therefore is silent on the skill premium), mine aims to understand how international trade affects the overall domestic inequality and the aggregate sill premium. Guided by this focus, my model is richer in that it incorporates skilled and unskilled workers and several ingredients emphasized in the skill premium literature, which prove important. Second, the structural estimation in this paper, combined with a newly constructed Hukou reform dataset, allows me to estimate the effect of Hukou reforms on migration costs, which then serves as input to the counterfactual experiments. This approach ensures that migration cost reductions in the reform scenario are reasonable. A by-product of this exercise is a new *prefecture-level* Hukou reform *panel*, which constitutes an independent contribution and will be useful for the research community on China's Hukou policy.[4] Finally, the exercise on the interaction between Hukou reforms and international trade also adds to the literature studying institutional reforms and the gains from trade (Kambourov, 2009).

[2]Also closely related are an economic geography literature that examines the interaction between international trade and domestic trade and specialization (Krugman and Elizondo, 1996; Venables and Limão, 2002; Cosar and Fajgelbaum, 2016), and a strand of empirical analysis on trade and inequality across regions (see, for example, Autor et al., 2013; Dix-Carneiro and Kovak, 2017).

[3]One exception is Galle et al. (2017), but in their paper, trade shocks have differential impacts across regions because workers from different regions have comparative advantage in different industries, rather than domestic trade costs—in fact, their benchmark model features frictionless domestic trade.

[4]The dataset spans between 1997 and 2010 and is constructed using a narrative approach. This approached has been used by Kinnan et al. (Forthcoming) and Sun et al. (2011), which measure Hukou reforms at the provincial level by *counting the number of reforms*. My dataset, on the other hand, is at the prefecture level, covers a longer period, and differentiates reforms based on their depth. See Section 2.2 and the supplementary note for more details.



(a) Trade Openness (b) Urban Employment Share

(c) Average Wage (d) Total Employment

Figure 2: City-Level Statistics

Source: Author's calculation based on 2005 City Statistics Yearbook (a), 2000 population census (b,d), and 2005 mini population census (c)

2 Spatial Economy and the Hukou System in China

This section provides basic facts about the economic geography of China and its Hukou system. These facts also motivate some of the model ingredients.

2.1 Economic Geography and Worker Mobility

Panels (a) of Figure (2) plots trade openness, defined as trade over GDP, for around 340 prefecture cities in China.[5] Border cities, especially those on the east coast, trade very intensively with the RoW. However, there is a steep decline in openness as the distance of a city to the coast increases. At the same time, as Panel (b) shows, cities along the east coast tend to have a larger urban sector. These spatial differences can be due to both intra-national trade costs and regional comparative advantages. The quantitative framework below incorporates both elements and will isolate the role of domestic trade costs through the lens of the model.

[5]The measure is winsorized at top 1%, 3.46. An outlying city has an openness measure of 43.

Table 1: Migration in China

Variable	Mean	Median	Std	N
Inter-Province Migrant Share	0.07	0.03	0.10	340
Intra-province Migrant Share	0.10	0.08	0.08	340
Skill Share in Employment	0.27	0.26	0.09	340
Skill Share in Inter-Provincial Migrants	0.32	0.29	0.16	340
Skill Share in Intra-Provincial Migrants	0.42	0.40	0.14	340

Notes: Source: authors' calculation based on the 2000 census. Sample includes all prefecture-level jurisdictions. Migration is defined based on the difference between the place of residence and the place of birth.

Panel (c) plots the log average wage relative to Beijing, net of differences in worker characteristics across cities.[6] The wage differ across cities by 40-70%. The southeastern coast tends to offer higher wages than the interior. (The exceptions are a few cities in the northeast, which are mostly natural resource cities with low population density.) Panel (d) plots the size of cities in terms of employment. Despite wages being higher along the southeastern coast, it is the central area of eastern China that exhibits a higher employment density, suggesting potentially significant barriers to migration. The development accounting literature has documented a significant wedge for workers relocating from rural to urban sectors (Gollin et al., 2013). Given the spatial differences in urban shares documented in Panel (b), if this wedge is not properly accounted for, I might incorrectly attribute the sectoral wedge to spatial migration costs. This consideration motivates a model with segmented rural and urban labor markets within each city.

Turning to the mobility of workers, while inter-city migration is far from enough to eliminate regional income differences, migration has been an important feature of the Chinese economy since the late 1980s. Table 1 summarizes the share of the inter- and intra-province migrant *stocks* in Chinese cities. On average, about 7% of workers in a city are born in other provinces, and about 10% are born in other counties within the same province. Skilled workers tend to be more mobile: in the median Chinese city, they account for 26% of employment, but 29% and 40% of inter- and intra- provincial migrants, respectively. As shown in the quantitative section, differential mobilities play a role in the spatial transmission of trade shocks.

[6]The regional average wages are measured as the regional fixed effects in an individual-level Mincer regression with worker characteristics—such as education, gender, and age—controlled for. Regression results are discussed in Online Appendix B.

2.2 The Hukou System and Reforms

The Hukou system that ties individuals to locations is one of the reasons why despite large spatial inequality, migration is not more prevalent. This subsection briefly discusses the history of Hukou and how it affects worker mobility. The supplementary note on the Hukou reform dataset provides more details.

First introduced in the 1950's, the original goal of Hukou was to manage individual mobility and occupation. In the era of a command economy, since most jobs were controlled by the state and foods rationed according to Hukou, Hukou could be strictly enforced. The boom in the private economy in the 80's and 90's made enforcement difficult. People started to move to cities for job opportunities. However, without official Hukou, migrants were ineligible for many local public goods, such as health care, schooling and social security. As a result, even though it was possible to find a job in the private sector, Hukou still greatly penalized migration.

Beginning in 1997, with the permission from the central government, some prefectures started gradual Hukou reforms and allowed qualified people from the rural area and other cities to obtain local Hukou. Implemented in only a small number of counties and with a high bar for qualification, reforms were initially very restricted. Over the years, reforms gradually expanded to more cities and a larger fraction of workers, but until today, Hukou remains restrictive in many Chinese cities.

Given this institutional background, it is important to quantify the migration barriers created by the Hukou system and the potential effects of its abolishment on inequality and trade. As detailed in the supplementary note, I construct a new prefecture-level panel of Hukou reforms spanning 1997-2010. Specifically, I collect news articles, official documents, and government regulations about Hukou policies at the local level, from keyword searches on two comprehensive databases. Following a set of criteria, I review these documents and hand code them into a score of 0-6 for each city based on the difficulty faced by migrants in obtaining Hukou.[7] A higher score means easier access to local Hukou. Using over-time variation in Hukou scores, I estimate that each additional point in this score translates into a 19% increase in inward migrants. In the quantitative section, I will back out the corresponding decrease in migration cost through an indirect-inference exercise. That parameter will determine how migration costs will change in the scenario of a complete abolishment of the Hukou system (with all cities having a score of 6).

[7] These criteria include, for example, whether renting/purchasing local properties and/or working in a city for a sustained period of time qualifies a worker for local Hukou; whether such qualification applies to just the surrounding counties and rural areas of a prefecture or it also applies to the central district. A score of 6 indicates complete openness. A score of 0 indicates strict control.

3 The Model

This section describes the spatial equilibrium model used in quantification.

3.1 Environment

There are $2N+1$ regions in the economy. These regions consist of rural and urban sectors of N Chinese cities, in total $2N$ regions, and one extra region that represents the rest of the world (RoW). I denote the set of all Chinese regions $\mathbf{G}$, and denote $\mathbf{R}$ and $\mathbf{U}$ the rural and urban subsets of $\mathbf{G}$: $\mathbf{G} = \mathbf{R} \bigcup \mathbf{U}$. I will use o, $d \in \mathbf{G}$ to refer to the origin and destination of domestic trade and migration flows.

3.2 Workers

Workers differ in levels of skill, $e \in \{h, l\}$, where h and l stand for high-skill and low-skill, respectively. Wage is their sole source of income. The wage of a type-e worker in region d depends on both the skill-specific wage rate for each labor unit, W_d^e, and the number of labor units a worker possesses—-or his productivity—in region d. I assume that the productivity of a worker in any region is a random draw from a given distribution, to be specified below.

Workers value the final consumption good and regional amenities, and choose to live in the region with the most desirable bundle of wages, prices, and amenities, taking into account migration costs. Consider worker i from region o, with a productivity draw $z_d(i)$ in d. Given prices and productivity draws, the indirect utility this worker would obtain by migrating to region d is assumed to take the following form:

$$V_{o,d}^e = \frac{B_d^e W_d^e z_d(i)}{P_d d_{od}^e},$$

In the above expression, P_d is the price of the final consumption good, B_d^e is the regional amenities, and d_{od}^e is the iceberg migration cost, which is allowed to be both skill-specific and source-destination specific.[8] This indirect utility function corresponds to a preference linear in the product of regional amenities and the quantity of the final consumption good. Worker i chooses d to maximize this indirect utility. Formally,

$$U_o^e = \max_{d \in \mathbf{G}} \{ \frac{B_d^e W_d^e z_d(i)}{P_d d_{od}^e} \}. \tag{1}$$

[8] As in other studies of spatial equilibrium models (Ahlfeldt et al., 2015), in calibration, amenities will act as the structural residuals that match exactly the distribution of employment in the data. I will keep them constant for all counterfactual experiments.

As in Ahlfeldt et al. (2015), I assume that $\mathbf{z} = (z_1, z_2, ...z_{2N})$ are generated from the Frechet distribution. To capture the individual-specific component in workers' productivity, I allow each worker's draws to be correlated across regions. Specifically, the vector of productivity draws for any given worker is generated from the following CDF:

$$F(\mathbf{z}) = \exp(-(\sum_{d \in \mathbf{G}} z_d(i)^{-\epsilon_e})^{1-\rho}). \tag{2}$$

where ρ controls the inter-regional correlation of productivity draws and ϵ_e controls their cross-sectional dispersion.[9] For ease of notation, let v_d^e be the amenity-adjusted real wage *rate* in region d: $v_d^e \equiv \frac{B_d^e W_d^e}{P_d}$. Then the probability that a worker from origin o moves to destination d is:

$$\pi_{od}^e = \frac{(\frac{v_d^e}{d_{od}})^{\epsilon_e}}{\sum_{g \in \mathbf{G}} (\frac{v_g^e}{d_{og}})^{\epsilon_e}} \tag{3}$$

Letting L_d^e be the *number of workers* with skill level e working in d, l_o^e be the number of workers born in o, and l_{od}^e be the number of workers moving from o to d, we have the following:

$$L_d^e = \sum_{o \in \mathbf{G}} l_{od}^e = \sum_{o \in \mathbf{G}} l_o^e \pi_{od}^e. \tag{4}$$

Because the model is static and migration is a once-for-life choice, L_d^e, l_o^e, and l_{od}^e should all be interpreted as stocks and will be mapped into corresponding stock variables in the data.

Due to the self-selection on productivity in migration, L_d^e is different from the supply of *effective labor units* in region d. Using properties of the Frechet distribution, in the appendix I derive the average labor efficiency of workers moving from o to d as:

$$\mathbb{E}(z_d^e | l_{od}^e) = (\frac{1}{\pi_{od}^e})^{\frac{1}{\epsilon_e}} \Gamma(1 - \frac{1}{\epsilon_e(1-\rho)}) \tag{5}$$

in which $\Gamma(\cdot)$ is the Gamma function. The negative relationship between the share of workers moving from o to d and the average labor efficiency of the migrants capture the selection effect. The intuition is that, if a higher fraction of workers from o choose to work in d, d must be especially attractive (with either high v_d^e or low d_{od}^e). This induces workers with bad z_d draws to migrate to d, lowering the average efficiency.

The total stock of effective labor units brought to d by workers from o is simply the product of l_{od}^e and $\mathbb{E}(z_d^e | l_{od}^e)$. Aggregating over migrants from all origins, the total supply

[9] I normalize the mean of the productivity distributions to be the same across regions. Differences in regional productivity enter the economy from the production side.

of effective labor units in d, denoted E_d^e, is given by:

$$E_d^e = \sum_{o \in \mathbf{G}} \mathbb{E}(z_d^e | l_{od}^e) l_o^e \pi_{od}^e. \tag{6}$$

Workers spend their income in d, so the consumption expenditures in region d is:

$$R_d = \sum_{e \in \{h,l\}} E_d^e W_d^e.$$

3.3 Production and Trade

The production side of the economy is a multi-sectoral version of Eaton and Kortum (2002), extended to incorporate input-output linkages (Caliendo and Parro, 2015) and capital-skill complementarity (Burstein et al., 2013; Parro, 2013). There are four production industries in the economy: agricultural (A), capital and equipment (K), other manufacturing (M), and services (S).

3.3.1 Intermediate Variety Production

Within industry $s \in \{A, M, K, S\}$, there is a continuum of intermediate varieties, indexed by $\omega \in \Omega_s$. Intermediate varieties are produced using industry final outputs and equipped composite labor, both of which are introduced below. To capture the segmentation between rural and urban labor markets, I assume intermediate variety producers in urban industries (M, K, and S) are located only in urban regions and hire equipped composite labor from urban labor markets; intermediate variety producers in the agricultural industry are located only in rural regions and hire equipped composite labor from rural labor markets. I abstract from such segmentation and capital-skill complementarity in the RoW as it is simply a statistical aggregation of countries.

The production function for intermediate variety ω in region d, industry s, is

$$y_d^s(\omega) = t_d^s(\omega) l_d^{s\,\gamma_s^L}(\omega) \prod_{s' \in \{A,M,S\}} m_d^{ss'\,\gamma_s^{s'}}(\omega)$$

$$s \in \{A\} \text{ if } d \in \mathbf{R} \cup \text{RoW}; \quad s \in \{M, K, S\} \text{ if } d \in \mathbf{U} \cup \text{RoW}.$$

In the production function, $t_d^s(\omega)$ is region d's efficiency in producing variety ω, $m_d^{ss'}(\omega)$ denotes the quantity of the final good of industry s', and $l_d^s(\omega)$ is the employment of equipped composite labor, which is made of skilled and unskilled labor, and capital equipments. The setup also implies that final good from the K industry enters this production only through l_d^s. $\gamma_s^{s'}$ is the share of each factor: $\gamma_s^L + \gamma_s^A + \gamma_s^M + \gamma_s^S = 1$.

Letting P_d^s be the price of final output of industry s in region d, and W_d the price for one unit of equipped composite labor in region d, the cost of production for ω is

$$\frac{(\frac{W_d}{\gamma_s^L})^{\gamma_s^L} \prod_{s' \in \{A,M,S\}} (\frac{P_d^{s'}}{\gamma_s^{s'}})^{\gamma_s^{s'}}}{t_d^s(\omega)} \equiv \frac{c_d^s}{t_d^s(\omega)}, \tag{7}$$

with c_d^s introduced to denote the cost of ω for a producer with unit productivity.

3.3.2 Industry Final Good Production

In each city and industry, there is a representative final good producer, which combines intermediate varieties of the same industry into final outputs, to be used for final consumption and production of intermediate varieties. I assume that industry final outputs are non-tradable across cities but freely tradable between the rural and urban regions *within* a city. Therefore, residents and intermediate variety producers in rural and urban regions of a city have the same access to industry final goods, despite their different specializations. The production technology for industry s, region d, is the following:

$$Q_d^s = [\int_{\omega \in \Omega_s} q_d^s(\omega)^{\frac{\sigma_s - 1}{\sigma_s}} d\omega]^{\frac{\sigma_s}{\sigma_s - 1}}, s \in \{A\} \text{ if } d \in \mathbf{R} \cup \text{RoW}; \quad s \in \{M, K, S\} \text{ if } d \in \mathbf{U} \cup \text{RoW}, \tag{8}$$

where is $q_d^s(\omega)$ is the quantity of variety ω used.

3.3.3 Trade in Intermediate Varieties

Intermediate varieties in A, M, and K industries are tradable both domestically and internationally; intermediate varieties in the service industry are non-tradable.[10] As in Eaton and Kortum (2002), I assume that $t_d^s(\omega)$ is generated independently across ω and d from the Frechet distribution with location parameter T_d^s and dispersion parameter θ:

$$F_d^s(t) = \exp(-\mathrm{T}_d^s t^{-\theta}).$$

This distribution implies that the share of region o's products in the intermediate varieties used in region d is

$$\delta_{do}^s = \frac{T_o^s (c_o^s \tau_{do})^{-\theta}}{\sum_{o'} T_{o'}^s (c_{o'}^s \tau_{do'})^{-\theta}}, \tag{9}$$

where the denominator sums over $\mathbf{U} \cup \text{RoW}$ if $s \in \{M, K, S\}$, that is, if s indexes an urban industry, and over $\mathbf{R} \cup \text{RoW}$ if $s \in \{A\}$. Familiar results under the Frechet distribution

[10] In the following, I assume trade costs are infinite for intermediate varieties in the service industry and proceed as if services were tradable.

also implies that the unit price for the industry final good corresponding to production function (8) is

$$P_d^s = [\Gamma(\frac{\theta + 1 - \sigma_s}{\theta})]^{\frac{1}{1-\sigma_s}} (\Psi_d^s)^{-\frac{1}{\theta}}, \tag{10}$$

where $\Psi_d^s = \sum_o T_o^s (c_o^s \tau_{do})^{-\theta}$. Again, the summation is taken over urban regions for urban industries and over rural regions for the agricultural industry.

3.3.4 Equipped Composite Labor Production

Equipped composite labor is produced by a representative producer in each region, from capital and the two types of labor. I incorporate capital-skill complementarity by specifying the production function of equipped composite labor in a nested CES form, with capital being complementary to high-skill labor, and substitutable to low-skill labor.

Formally, effective high-skill labor units, E_d^h, low-skill labor units, E_d^l, and capital and equipment, K_d, are combined into *equipped composite labor*, E_d, through the following technology:

$$E_d^{eh} = [(1-\eta_d^h)^{\frac{1}{\rho_{kh}}} (K_d)^{\frac{\rho_{kh}-1}{\rho_{kh}}} + (\eta_d^h)^{\frac{1}{\rho_{kh}}} (E_d^h)^{\frac{\rho_{kh}-1}{\rho_{kh}}}]^{\frac{\rho_{kh}}{\rho_{kh}-1}}$$

$$E_d = [(1-\eta_d^l)^{\frac{1}{\rho_{lkh}}} (E_d^l)^{\frac{\rho_{lkh}-1}{\rho_{lkh}}} + (\eta_d^l)^{\frac{1}{\rho_{lkh}}} (E_d^{eh})^{\frac{\rho_{lkh}-1}{\rho_{lkh}}}]^{\frac{\rho_{lkh}}{\rho_{lkh}-1}},$$

where E_d^{eh} is *equipped high-skill labor*, the output from the inner nest. ρ_{kh} ($\rho_{kh} < 1$) is the elasticity of substitution between high-skill labor and capital, and ρ_{lkh} ($\rho_{lkh} > 1$) is the elasticity of substitution between equipped high-skill labor and low-skill labor. η_d^h and η_d^l determine the region-specific factor shares in equipped composite labor and allow me to match skill premia by city.

Letting W_d^h / W_d^l be the wage rate for high-/low-skill labor, W_d^{eh} be the unit price for equipped high-skill labor, and W_d be the unit price for equipped composite labor, the optimization decision and the zero-profit conditions of equipped composite labor production imply the following:

$$\begin{aligned} W_d^{eh} &= [(1-\eta_d^h)(P_d^K)^{1-\rho_{kh}} + (\eta_d^h)(W_d^h)^{1-\rho_{kh}}]^{\frac{1}{1-\rho_{kh}}} \\ W_d &= [(1-\eta_d^l)(W_d^l)^{1-\rho_{lkh}} + (\eta_d^l)(W_d^{eh})^{1-\rho_{lkh}}]^{\frac{1}{1-\rho_{lkh}}} \end{aligned} \tag{11}$$

$$\begin{aligned} \frac{P_d^K K_d}{W_d^h E_d^h} &= (\frac{P_d^K}{W_d^h})^{1-\rho_{kh}} \frac{1-\eta_d^h}{\eta_d^h} \\ \frac{W_d^{eh} E_d^{eh}}{W_d^l E_d^l} &= (\frac{W_d^{eh}}{W_d^l})^{1-\rho_{lkh}} \frac{1-\eta_d^l}{\eta_d^l} \end{aligned} \tag{12}$$

Equation (12) expresses factor shares in equipped composite labor as a function of relative prices and technological parameters, η_d^h and η_d^l. Factor shares are endogenous. Nonetheless, to simplify notation, I use β_d^K, β_d^h, and β_d^l to denote the shares of capital, high-, and low-skill labor in equipped composite labor in region d: $\beta_d^K + \beta_d^h + \beta_d^l = 1$

3.3.5 Final Consumption Good Production

The final consumption good is non-tradable and produced with industry final goods using the following technology:

$$C_d = (C_d^A)^{s_A}(C_d^M)^{s_M}(C_d^S)^{s_S}, \quad s_A + s_M + s_S = 1,$$

in which C_d^s is the quantity of industry-s final good used in consumption.[11] The price of the final consumption good, P_d, is then given by:

$$P_d = (\frac{P_d^A}{s_A})^{s_A}(\frac{P_d^M}{s_M})^{s_M}(\frac{P_d^S}{s_S})^{s_S} \tag{13}$$

3.4 Goods and Labor Markets Clearing Conditions

It remains to describe market clearing conditions for labor and final goods. Let X_d^s denote total production of industry-s final good in region d. The demand for industry-s intermediate varieties produced in region o, denoted D_d^s, is given by:

$$D_d^s = \sum_{d'} X_{d'}^s \delta_{d'd}^s, \tag{14}$$

in which the summation is over $\mathbf{U} \cup$ RoW for urban industries and over $\mathbf{R} \cup$ RoW for agriculture. To make D_d^s amount intermediate varieties, the producers in region d use $\gamma_s^{s'} D_d^s$ worth of the industry-s' final good. The producers also employ $\gamma_s^L D_d^s$ worth of equipped composite labor, the payment to which will be distributed to capital and workers. The labor market clearing conditions, separate for rural and urban regions of a city, are

$$\begin{aligned} &\text{Rural (d}\in \mathbf{R}\text{): } E_d^h W_d^h = D_d^A \gamma_A^L \beta_d^h; \quad E_d^l W_d^l = D_d^A \gamma_A^L \beta_d^l, \\ &\text{Urban (d}\in \mathbf{U}\text{): } E_d^h W_d^h = \beta_d^h \sum_{s\in\{M,K,S\}} D_d^s \gamma_s^L; \quad E_d^l W_d^l = \beta_d^l \sum_{s\in\{M,K,S\}} D_d^s \gamma_s^L. \end{aligned} \tag{15}$$

The RoW labor market is not segregated so the market clearing condition is simply the sum of Equation (15) across sectors.

[11] The share of capital and equipment in final consumption is very small and hence omitted. Incorporating K in final consumption has a negligible impact on results.

The demand for industry final outputs in each region comprises demand from residents and intermediate variety producers. Recall that industry final outputs are freely tradable within a city and non-tradable across cities. I use d and d' to denote the urban and rural regions of the same city, respectively, and write the market clearing conditions for industry final outputs compactly as:

$$
\begin{aligned}
X^A_{d'} &= (C^A_d + C^A_{d'}) + D^A_{d'}\gamma^A_A + \sum_{s\in\{M,K,S\}} D^s_d \gamma^A_s, \quad d' \in \mathbf{R} \\
X^M_d &= (C^M_d + C^M_{d'}) + D^A_{d'}\gamma^M_A + \sum_{s\in\{M,K,S\}} D^s_d \gamma^M_s, \quad d \in \mathbf{U} \\
X^S_d &= (C^S_d + C^S_{d'}) + D^A_{d'}\gamma^S_A + \sum_{s\in\{M,K,S\}} D^s_d \gamma^S_s, \quad d \in \mathbf{U} \\
X^K_d &= D^A_{d'}\gamma^L_A \beta^K_{d'} + \sum_{s\in\{M,K,S\}} D^s_d \gamma^L_s \beta^K_d. \quad d \in \mathbf{U}
\end{aligned}
\tag{16}
$$

On the left of Equation (16) is the production of industry final outputs in a city; on the right, $D^A_{d'}\gamma^{s'}_A$ and $\sum_{s\in\{M,K,S\}} D^s_d \gamma^{s'}_s$ are the demands for final good in industry s' from intermediate variety producers in the agricultural industry and the three urban industries, respectively; $C^s_d + C^s_{d'}$ is the sum of consumption demand from rural and urban regions. It is calculated as $s_s[R_d + R_{d'} - (S_d + S_{d'})]$, where R_d is region d's aggregate income, $S_d + S_{d'}$ is the city's international trade surplus taken as exogenous from the data, scaled to the model economy,[12] and s_s is the share of industry s in the final consumption bundle. For the RoW, because there is no distinction between rural and urban regions, $C^s_d + C^{s'}_d$ is replaced with C^s_{RoW}.

The parameters in the economy are: spatial frictions, including migration costs $\{d^e_{od}\}$ and trade costs $\{\tau_{o,d}\}$; preference parameters $\{\sigma_A, \sigma_M, \sigma_K\}$, and $\{s_A, s_M, s_S\}$; production technology, including $\{\gamma^{s'}_s\}$, $\{\eta^e_d\}$, $\{\rho_{kh}, \rho_{lkh}\}$, and θ; local productivity and amenities, $\{T^s_d\}$ and $\{B^e_d\}$.

Definition 1 *Given above exogenous parameters and labor endowment $\{l^o_d\}$, a competitive equilibrium of the economy is a set of prices and allocations, such that optimization conditions for consumers and producers are all satisfied and all markets clear—Equations (3), (6), (7), (9), (10), (11), (12), (13), (15), (16).*

[12] I provide details on the construction of city-level surpluses in the appendix. Adjusting for trade surpluses ensures that the calibration of regional productivity takes into account the international trade imbalances, about 5% of the GDP of China in 2005. In counterfactual experiments, I do not allow for trade imbalances.

4 Parameterization

I parameterize the model using the data from the Chinese economy around the year 2005. This section describes the parameterization process, starting with data sources.

4.1 Data Description

Quantifying the model primarily requires the following information: as in Alvarez and Lucas (2007), to calibrate regional productivity, we need wages and employment for high- and low-skill workers in all regions; to calibrate region-specific parameters in the equipped composite labor production function, we need shares of factors in equipped composite labor; to estimate domestic migration costs we need migration flows; to further identify the component of migration costs created by the Hukou system we need variation from local Hukou reforms; to estimate trade costs we need information on domestic trade flows; finally we need measures of geographic and cultural distances between regions. I briefly discusses data sources here; Online Appendix B provides additional details.

I use the 2005 mini population census to estimate the wage rates for Chinese regions. I estimate the average wage for unskilled workers and the skill premium in each region as the regional fixed effects and the region-specific skill dummies, in an individual Mincer wage regression that controls for a rich set of individual demographic and occupation variables. This regression approach nets out the differences in demographics and detailed industry structures across regions, which are not explicitly modeled. Figure 2c plots a map of average urban wage estimated this way.

I also use the 2005 mini census to measure the *number of workers* employed in each city-industry. Once we have the estimates for migration costs $\{d^e_{od}\}$ and regional amenity-adjusted real wages $\{v^e_d\}$, we can use Equation (6) to convert the number of workers into the employment of *effective labor units*. Combined with effective wage estimated above, this information gives me the wage bill for high- and low-skill workers at the city-industry level.[13]

Using the data described above, we can readily compute the relative shares of wage payments to high- and low-skill workers. Determining η^h_d and η^l_d, the region-specific parameters in the equipped composite labor production functions, further requires the relative shares between capital and equipment (K) and labor. For the urban sector, I use

[13] I run into a small-sample problem and end up with zeros for employment in capital and equipment industry in some cities. To overcome this problem, I tabulate employments, differentiating only between agricultural and urban industries. I supplement this information with relative share of urban industries in each city, constructed from the manufacturing sub-sample of the 2004 economic census, to obtain the employment information at the city-industry level.

the 2004 Annual Survey of Industrial Production to construct wage bill and capital expenditures for each city; for the rural sector, due to the lack of regional data, I assume all cities have the same capital/labor share, and determine this share using the national input-output table.

To construct a database of inter-regional and inter-sectoral migration, I use the 2000 population census. It serves my purpose best because both birthplace and current residence information are reported, which allows me to measure life-time migration as defined in the model. For each worker, I identify her skill level, current city, birth province, type of Hukou, and whether she is currently working in a rural or urban industry. I then determine her migration status based on this information.[14]

I construct proxies for geographic distance and cultural distance between Chinese cities. The geographic distance between two cities is calculated as the greater-circle distance between the coordinates of their city centers, proxied by the locations of their local governments, extracted from the Google Maps. The cultural distance is constructed as $1 - \text{corr}(V_o, V_d)$, where V_o is a vector, the elements of which are the shares of various ethnic groups in the total ethnic minority population in o in the 1990 census. This measure is small if two cities had similar compositions of ethnic minorities in the 1990s.[15]

Finally, I use the 2002 inter-regional input-output table of China to construct trade flows between Chinese provinces. I collect city-level international import and export information from the 2005 provincial statistical yearbooks of foreign trade. This information will be used to estimate domestic trade costs.

4.2 Parameters Calibrated Independently

I calibrate the following parameters independently. The dispersion parameter ϵ_e governs the variance of the idiosyncratic component of workers' productivity draws. The parametric assumption in Equation (2) implies that the wage distribution of workers sharing the same migration origin and destination has the following coefficient of variation:

$$\frac{\text{Variance}}{\text{Mean}^2} = \frac{\Gamma(1 - \frac{2}{\epsilon_e(1-\rho)})}{(\Gamma(1 - \frac{1}{\epsilon_e(1-\rho)}))^2} - 1. \tag{17}$$

Guided by this relationship, I use the wage distribution of stayers to recover $\epsilon_e(1-\rho)$.

[14] I restrict the sample to workers who have finished schooling with age between 20 and 60. In Online Appendix B.6, I discuss this sample selection in detail, as well as the drawbacks of alternative ways of constructing migration flows, e.g., by defining migrants as people working in a city without local Hukou.

[15] Migrations were less common prior to 1990; therefore the correlation constructed this way captures the historical cultural distance between regions and is unlikely to be driven by current migration. Online Appendix B.2 provides background on ethnicity in China and the summary statistics of cultural distance.

Specifically, I regress the log wage of stayers on regional fixed effects, individual characteristics, and industry fixed effects, for high- and low-skill worker samples separately. I then take the exponents of the residuals and compute their coefficients of variation. I choose $\epsilon_e(1-\rho)$ so that Equation (17) gives the same value. This procedure determines $\epsilon^{\mathrm{h}}(1-\rho) = 2.73$ and $\epsilon^{\mathrm{l}}(1-\rho) = 2.5$. By deriving statistics for only stayers' wage distribution and matching them to their data counterparts, this procedure takes into account self-selection on productivity in migration.

The parameter ρ controls the correlation of individuals' productivity draws across regions. The correlation in wages of *migrants* before and after migration is informative about ρ. My strategy for calibrating ρ is therefore first to compute the explanatory power of individual fixed effects in an individual-*panel* wage regression using a sample of migrants only.[16] I then choose ρ so that in the simulated data, individual fixed effects have the same level of explanatory power. This procedure determines ρ to be 0.36.

Productivity dispersion in intermediate varieties, θ, is not separately identifiable from trade costs using my data. I assign a value of 4, the preferred estimate of Simonovska and Waugh (2014), to the productivity dispersion for A, M, and K industries.[17] The elasticities of substitution between high-skill labor and capital, and between low-skill labor and equipped high-skill labor, are set to the estimates in Krusell et al. (2000)—0.67 and 1.67, respectively. These values imply that capital and high-skill labor are complements, and both substitutes to low-skill labor. I also perform robustness exercises for different values of trade elasticity and capital-skill complementarity.

Finally, the composition of final consumption bundle, $\{s_A, s_M, s_S\}$, and shares of different inputs in intermediate variety production, $\{\gamma_s^{s'}\}$, are calibrated to the 2002 national input-output table.[18] The upper panel of Table (2) summarizes the sources and values of these parameters. The lower panel provides information on other parameters determined in the model, which I discuss in the rest of this section.

[16] In the panel regression I control for regional and time fixed effects as well as time-varying individual characteristics such as age and occupation. The residual of this regression can then be interpreted as the data counter-part of productivity draws for migrants from different regions. I then further add individual fixed effects to this regression. The *additional* explanatory power of these individual fixed effects tells us to what extent wages are correlated *overtime* for *movers*, which maps one-to-one into ρ.

[17] Simonovska and Waugh (2014) focuses on aggregate trade flows. Papers focusing on agricultural trade alone, for example, Donaldson (2018), report similar estimates for the elasticity of trade.

[18] Input shares for the RoW are taken as the median values from countries in Parro (2013). The values of these parameters are reported in Appendix B.4.

Table 2: Model Parameterization

A: Parameters Calibrated Independently			
Parameter	Description	Target/Source	Value
ρ	Worker productivity draw correlation	Correlation in wages for migrants	0.36
ϵ^h,ϵ^l	Worker productivity draw dispersion	Equation (17)	$\epsilon^h = \frac{2.73}{1-\rho}$, $\epsilon^l = \frac{2.5}{1-\rho}$
θ	Elasticity of trade	Simonovska and Waugh (2014)	4
ρ_{kh}, ρ_{lkh}	Elasticities in equipped composite labor	Krusell et al. (2000)	$\rho_{kh} = 0.67, \rho_{lkh} = 1.67$
s_A, s_M, s_S	Sectoral shares in final consumption	Aggregate consumption share	$s_A = 0.22, s_M = 0.24$ $s_S = 0.53$
$\gamma_s^{s'}$	Input-output linkages	National input-output tables	Appendix B.4
B: Parameters Estimates/Calibrated in Equilibrium			
Parameter	Description	Target/Source	Value
$\{d_{o,d}\}$	Migration costs	Migration flows	Table(3)
$\{\tau_{o,d}\}$	Domestic trade costs	Domestic trade & city import/export	Table(4)
$\{t_a, t_m, t_k\}$	Trade costs between ports and RoW	Sectoral international trade	Table(4)
$\{\eta_d^h\}$, $\{\eta_d^e\}$	Equipped labor production function	Corresponding factor shares	-

4.3 Migration Cost Estimation

4.3.1 Specification

The first step in parameterizing the rest of model is to estimate migration costs. I normalize $d_{oo} = 1$ and specify the cost of a migration move from o to d as

$$\ln(d^e_{o,d}) = \sum_{i=1}^{4} \beta^e_i I_i + \beta^e_5 * I_1 * \text{dist}_{o,d} + \beta^e_6 * I_2 * \text{dist}_{o,d} + \beta^e_7 * I_3 * \text{dist}_{o,d} + \beta_{e*\text{Cdist}_{o,d}+\mu_{o,d}}, \quad (18)$$

where I_1-I_4 are mutually exclusive dummy variables: I_1 indicates if o and d belong to different cities within the same province; I_2 indicates if o and d belong to different provinces within the same large region (of which there are seven in China, each containing five provinces on average); I_3 indicates if o and d belong to different large regions; and I_4 is an indicator for rural-urban migration. These indicators capture different institutional barriers to the free mobility of labor. $\text{dist}_{o,d}$ is the great-circle distance between o and d. I allow distance to have a nonlinear impact by interacting it with I_1-I_3. Finally, the literature has identified important network effects in international immigration (Munshi, 2003). As a proxy for these network effects, I include cultural distance $Cdist_{o,d}$.

4.3.2 Estimation Procedure

To estimate Equation (18) I use a nested nonlinear least square procedure. In the outer loop, I choose $\{\beta\}$ to minimize the deviations of migration flows in the model from their data counterpart. In the inner loop, I choose $\{v^e_d\}$, the amenity-adjusted real wages, so

that the number of workers in each city match that in the data exactly.[19] Matching the employment distribution allows the model to match inter-regional inequality in the calibrated equilibrium. The identification of $\{v_d^e\}$ relies on the distribution of employment across cities. Intuitively, controlling for migration costs, if a region is revealed to attract more migrants, then it must have higher amenities. Because the model is at the city level, and the migration flow data are at province-to-city level, in the outer loop I aggregate the model-predicted flows before comparing them to the data.

Formally, letting $p_i \in \mathbf{P}$ be individual provinces in China, $L_{p_i d}^{\text{data},e}$ be the number of workers migrating from province p_i to region d in the data, and $L_d^{data,e}$ be the number of workers in region d in the data, the loss function in the outer loop is:

$$\min_{\{\beta^e\}} \sum_{p_i \in \mathbf{P}, d \in G} \left(\log(\sum_{o \in p_i} l_o^e \pi_{od}^e) - \log(L_{p_i d}^{\text{data},e})\right)^2.$$

In the inner loop, I simply choose v_d^e so that given migration costs, $\sum_{o \in \mathbf{G}} l_o^e \pi_{od}^e = L_d^{data,e}$ holds for all regions.

4.3.3 The Effect of Hukou Reforms

After recovering the migration cost parameters $\{\beta^e\}$ from cross-sectional variation using the above procedure, I further use *over-time* variation from Hukou reforms to estimate the impacts of eliminating Hukou on migration costs.

Specifically, I assume that $\hat{\beta}_i^e$ $i \in \{1, 2, 3\}$, the estimated coefficients for inter-regional dummies in Equation (18), captures the cost of migrating to another city in 2000. The reforms documented in the companion note lead to variation in Hukou policies across cities afterwards. I assume $\beta_{i,d}^{e'}$ is cost of migrating to city d after a period of reforms:

$$\beta_{i,d}^{e'} = \hat{\beta}_i^e - \beta_{\text{Hukou}} \Delta\text{HukouScore}_d, i \in \{1, 2, 3\} \tag{19}$$

In the above definition, $\Delta\text{HukouScore}_d$ is the change in Hukou openness score between 2000 and 2005 in city d. Since the impact of reforms on migration might not be immediate, I define HukouScore for 2000 and 2005 by averaging over the five preceding years, and

[19] As shown in Appendix A.4, $\{v_d^e\}$ can be identified subject to normalization. Recall that $v_d^e = \frac{B_d^e W_d^e}{P_d}$. The estimation process determines v_d^e; W_d^e is from the data; once we estimate the trade cost, we know P_d. We can therefore back out amenities, B_d^e, as residuals. These amenities will be kept constant in all counterfactual experiments. An additional complication arises: I first use the 2000 data to estimate the migration costs; the wage and employment data, on the other hand, represents the 2005 economy. To ensure that the recovered $\{v_d^e\}$ are consistent with the 2005 employment distribution, after estimating $\{\beta\}$ for 2000 and update it to reflect the effects of Hukou reforms between 2000 and 2005, I use workers' birthplace and employment distribution in 2005, and solve the inner loop again to obtain the $\{v_d^e\}$ consistent with the 2005 economy.

calculate $\Delta\text{HukouScore}_d$ as the difference between the two average numbers.

I follow an indirect inference approach to estimate β_{Hukou}.[20] In the companion note, I estimate a reduced-form relationship between the percentage *increase* in the number of migrants into a city and $\Delta\text{HukouScore}_d$ of that city, controlling for potential confounding factors such as per-capita income, whether a city is a provincial capital or province-level municipality, a region's distance to the coast, as well as a proxy for the provision of local public goods. I find that on average, a one-point increase in HukouScore_d leads to a $19\% - 21\%$ increase in the number of migrants in city d.

I then use the model to simulate the effects of the exact *same reforms* as in the reduced-form regression.[21] Specifically, given the data $\Delta\text{HukouScore}_d$, for any value of β_{Hukou}, Equation (19) gives values for $\beta^{e'}_{i,d}$, with which I simulate the model and estimate the model counterpart of the relationship between $\Delta\text{HukouScore}_d$ and the changes in number of migrants. I choose β_{Hukou} so that the regression based on model-generated data also delivers the same coefficient. To preview the result, I find that $\hat{\beta}_{Hukou} = 0.13$, meaning that each point increase in the Hukou openness score reduces inter-city migration cost by 13%. The effective migration costs used in the following calibration, which targets the 2005 Chinese economy, will take into account changes in migration cost between 2000 and 2005 due to Hukou reforms estimated here.

4.4 Calibrating the Rest of the World

With the estimated migration costs, Equation (6) predicts the supply of effective labor in each place. Together with the regional wage estimated before, I compute the regional and, in turn, the national labor value added in China. I then use the share of Chinese value added in the world, calculated from Penn World Table 6.1, to determine the size of the RoW (labor value added=GDP in my model):

$$\text{GDP}_{\text{RoW}} = \text{GDP}_{\text{China}} * \frac{\text{Data GDP}_{\text{RoW}}}{\text{Data GDP}_{\text{China}}}$$

The total number of effective labor units in the RoW is then calculate as $\frac{\text{GDP}_{\text{RoW}}}{\text{Wage}_{\text{RoW}}}$. This is the endowment of the RoW and will be treated as fixed throughout all counterfactual experiments.

[20] Since in constructing the dataset I do not differentiate Hukou reforms by workers' skill type, I assume that β_{Hukou} is the same for both types of workers in empirical and quantitative analysis.

[21] Using the same variation in both reduced-form and structural analyses is important because Hukou reforms likely have heterogeneous effects across cities depending on their access to labor pools nearby.

4.5 Trade Cost Estimation

4.5.1 Specification

I jointly estimate trade costs and recover regional productivity for each sector. Following the gravity literature in international trade, I specify the trade cost between any two regions *within* China ($o, d \in \mathbf{G}$) as a log linear function of the geographic, institutional, and cultural distances between them:

$$\log(\tau_{o,d}) = \sum_{i=1}^{4} \gamma_i I_i' + \gamma_5 * I_1' * \text{dist}_{o,d} + \gamma_6 * I_2' * \text{dist}_{o,d} + \gamma_7 * I_3' * \text{dist}_{o,d} + \gamma_* \text{Cdist}_{o,d} + \epsilon_{o,d} \quad (20)$$

Dummy variables I_1'-I_3' in this specification are the same as I_1-I_3 in the migration cost specification. I_4' is an indicator for two provinces sharing a border. $\text{Cdist}_{o,d}$ and $\text{dist}_{o,d}$ are also defined in the same way as in the migration cost specification.

To capture the idea that it is more costly for inland cities to trade with the RoW, I assume that all trade between Chinese cities and the RoW need to go through one of the border/coastal cities. I specify the trade cost between a Chinese city and the RoW as the sum of two components: the trade cost between that city and *its nearest port/coastal city*, and a sector-specific parameter for international trade cost (t_a, t_m, t_k, respectively) that captures tariff and non-tariff barriers to international trade.[22] The international trade costs will be calibrated to match sectoral trade shares in China.

4.5.2 Estimation Procedure

I estimate trade cost parameters using a nested nonlinear least square procedure similar to how I estimate migration costs. The procedure consists of three steps. In the outer loop, I choose t_a, t_m, t_k to match sectoral level openness. In the middle loop, I choose $\{\gamma\}$ in Equation (20) to minimize the deviations of model-predicted bilateral trade flows from the data counterpart. In the inner loop, I choose sectoral-level productivity $\{T_d^s\}$ so that the production of goods in sector s by region d is consistent with the data.

The objective in the outer loop is simply sectoral trade openness, defined as the ratio between sectoral trade between China and the RoW and the sectoral production in China. The objective in the middle loop consists of two components. First, inter-provincial trade flows. Because the model is at the city level while the data are at the provincial level, I aggregate the trade flows in the model before taking them to the data. The second component in the objective function is city-level trade between Chinese cities and the RoW. By comparing international trade participation among cities within the same province, this

[22] $\log(\tau_{\text{interior,RoW}}) = \log(\tau_{\text{interior,port}}) + t_s$. $\log(\tau_{\text{d,port}})$ is the trade cost between d and the nearest port.

objective is informative about intra-provincial trade costs (γ_1 and γ_5). Formally, letting $X^{\text{model}}_{o,d}$ be the sum of trade flows from o to d in the model across all sectors and $X^{\text{data}}_{o,d}$ be its data counterpart, the problem in the middle loop is:

$$\begin{aligned} \max_{\{\gamma\}} \sum_{p_i,p_j\in\mathbf{P}} \big(\log(X^{\text{data}}_{p_i,p_j}) - \log(\sum_{o\in p_i, d\in p_j} X^{\text{model}}_{o,d})\big)^2 + \\ \sum_{o\in\mathbf{G},\, d=\text{RoW} \mid d\in\mathbf{G},\, o=\text{RoW}} \big(\log(X^{\text{model}}_{o,d}) - \log(X^{\text{data}}_{o,d})\big)^2, \end{aligned} \tag{21}$$

in which $p_i \in \mathbf{P}$ indicates a province and $o \in p_i$ indicates regions within province p_i.

The problem in the inner loop is to choose T^s_d so that Equation (16) holds for all regions and sectors. The left of Equation (16) can be directly calculated from the data. The right side depends on trade deficits (exogenously from data), and the total demand for goods produced in a region, given by Equation (14). The only unknown in solving Equation (16) is therefore trade shares, δ^s_{od}, which depends on wages (data) and T^s_d. We can therefore identify T^s_d (subject to normalization) by solving Equation (16). Intuitively, given trade costs and observed wages, if we see a region specializing in sector s, it must be relatively productive in that sector.

4.6 Estimation Results

4.6.1 Migration Costs

Table (3) reports migration cost estimates. The model fits the data well, as indicated by the high R-squared. The signs of coefficients are as expected: migration costs increase in all distance measures. The magnitudes are significant—the cost of migrating to other cities within the same province is around 130 log points for both types of workers. As a migration move covers more distance, it incurs a larger cost: for skilled workers, the *additional* cost of crossing a provincial border is about 30 log points, and the *additional* cost of crossing a regional border is another 23 log points; these costs are slightly higher for unskilled workers.

The continuous components of geographic distance have nonlinear effects on migration costs: when origin and destination regions are within the same province, additional geographic distance has a negligible and statistical insignificant effect; when they are in different provinces within the same large region, the marginal cost is sizable and statistically significant; when they are in different large regions, the marginal cost becomes somewhat smaller, but is still significant. This pattern holds for both types of workers, but the coefficients are in general much larger for unskilled workers. The estimation also reveals substantial costs, about 60 log points, associated with rural-urban migration. This

Table 3: Estimates of Migration Costs

	Skilled	Unskilled
I(Different Cities, Same Province)	1.29	1.32
	(0.05)	(0.05)
I(Different Provinces, Same Region)	1.60	1.66
	(0.03)	(0.03)
I(Different Regions)	1.83	1.96
	(0.03)	(0.03)
I(Rural to Urban)	0.62	0.64
	(0.02)	(0.02)
I(Different Cities, Same Province)*Distance	0.17	0.21
	(0.26)	(0.21)
I(Different Provinces, Same Region)*Distance	0.40	0.79
	(0.06)	(0.05)
I(Different Regions)* Distance	0.23	0.57
	(0.02)	(0.03)
Cultural Distance	0.15	0.16
	(0.03)	(0.04)
Observations	42160	42160
R^2	0.86	0.82

Notes: *Distance* is measured as the great circle distance between cities (in 1000 km); *Cultural Distance* is measured as one minus the correlation in lagged ethnic minority shares between cities. Robust standard errors are in parentheses.

coefficient, however, is only about a-third of the calibrated "labor wedge" for China in Swiecki (2017), calculated based on the rural-urban income gap. The difference underscores the importance of accounting for the geographic dimension: a large proportion of the measured rural-urban wedge could be a joint product of regional inequality and spatial frictions.

Finally, for both types of workers, the coefficients for the cultural distance are positive and significant. The inter-quartile range of cultural distance is around 0.5, so increasing cultural distance from 25th to 75th percentile leads to an increase of around 7-8 log points in migration costs.

4.6.2 Comparison to the U.S. and the Role of the Hukou System

The estimated migration costs could be attributed to both the Hukou system and workers' preference to stay close to family and friends. It is thus informative to compare my estimates to those based on the U.S. data, where policy distortions are likely less severe. Results from Diamond (2016), which estimates a discrete choice model of migration, suggest that living in a city outside the state of birth is equivalent to a 55 log point decrease in the real wage, and living in a city outside the census division of birth is equivalent to a 86

log point decrease in the real wage.[23] Since a U.S. state and a census division are similar to a Chinese province and a geographic region in my estimates, respectively, the corresponding numbers in my specification are around 160 and 183 log points—the migration costs are larger in China by around 100 log points.

To what extent can the differences be explain by the Hukou system in China? The indirect inference exercise in Section 4.3.3 finds that a one-point increase in the Hukou openness score reduces the cost of moving into a city by 13 log points. The average score in 2000 is around 1. If the Hukou system is completely eliminated, then all cities will have a score of 6. This would decrease migration costs from the estimated value by around 65 log point—about two-thirds of the gap between the U.S. and China. The remaining differences might be attributed to cultural differences, under-developed transportation infrastructures in China during the early 2000s, or other distortions.

This exercise also highlights the importance of using actual Hukou reforms for identification: if we were to take the U.S. estimate for the counterfactual of a complete overhaul of the Hukou system, we would overstate the effect by around 30%.

4.6.3 Trade Costs

Panel A in Table 4 presents domestic trade cost estimates. Trading with a different city in the same province incurs an iceberg trade cost of around 57 log points. Crossing a provincial border and a regional border further increases the cost by another 64 and 94 log points, respectively; sharing a common provincial border, on the other hand, could reduce costs by 6 log points. If these dummies variables indeed capture the institutional barriers to domestic trade, my estimates indicate that these barriers are large.

Geographic distance also significantly increases trade costs: for trading partners from different provinces within the same large region, each additional 1000 kilometers increases trade costs by 21 log points; for partners from two different regions, the impacts of distance are smaller—each additional 1000 kilometers increases trade costs by 4 log points. Between cities with in the same province, however, geographic distance does not appear to matter, and most of the cost is captured by the inter-city dummy. Finally, cultural distance also affects trade costs. Increasing the cultural distance from 25th to the 75th percentile increases trade costs by 10 log points.

[23] Since most worker in my sample period are not college graduates, I compare my results to estimates for the non-college worker group in Diamond (2016). According to the "fully flexible" model reported in her Tables 4 (Column 4) and 5 (Column 5), the coefficient associated with living in the same state of birth is 3.44, and that associated living in the same region of birth is 1.219 (the leave-out category being outside the census division of birth), whereas the wage coefficient is 4.026. So living outside the state of birth, but within the same census division is equivalent to 55 ($\frac{3.433-1.219}{4.026} * 100$) log points of the real wage, and living outside the census division of birth is equivalent to 86 ($\frac{3.433}{4.026} * 100$) log points in the real wage.

Table 4: Domestic and International Trade Costs

A. Domestic Trade Cost Estimates	Coefficient	Standard Error
I(Different Cities, Same Province)	0.57	(0.13)
I(Different Provinces, Same Region)	1.21	(0.10)
I(Different Regions)	1.51	(0.07)
I(Sharing Provincial Border)	-0.06	(0.06)
I(Same Province)*Distance	0.01	(0.12)
I(Different Provinces, Same Region)*Distance	0.21	(0.10)
I(Different Regions)* Distance	0.04	(0.03)
Cultural Distance	0.20	(0.08)
Observations	1580	
R^2	0.50	

B. International Trade Cost Calibration: Targets and Parameter Values	Trade/Production	International Trade Costs
Agricultural Industry	0.12	0.99
Manufacturing Industry	0.36	0.80
Capital and Equipment Industry	0.46	0.72

Notes: Panel A reports the estimates of domestic trade costs. See Table 3 for definitions of variables. Robust standard errors are in parentheses. Panel B reports the sectoral trade shares in the data and the calibrated trade costs between *port cities* and the RoW. Sectoral-level trade data are aggregated from the 2005 UN Comtrade database; sectoral production data are from the 2005 statistics yearbook.

Overall, the estimates suggest that trade costs between cities within China increase with both institutional and geographic distances. The former, captured by dummy variables in the regression, plays a more important role. The size of the inter-provincial dummy is smaller than in studies examining market fragmentation in China using earlier data (Poncet, 2005). On the other hand, relative to estimates for the U.S. (Crafts and Klein, 2014), the estimates here is about twice as large, reflecting larger barriers to trade flows at provincial border in China.[24]

Panel B of Table 4 presents sectoral trade shares. Capital and manufacturing industries are traded more heavily relative to the agricultural industry. This is also reflected in lower calibrated trade costs in these industries than in the agricultural industry.

[24]Since some of the variation used in estimation is from inter-provincial trade flows, one valid concern is I might mis-attribute the cost of trading within a province to provincial borders. If so, a further concern is whether counterfactual experiments would be affected. While not reported here, I estimate the model using only trade between Chinese cities and the rest of the world (the second term in Equation 21), and still find large provincial border effects. In addition, in Online Appendix C.6.2, I discuss related issues arising in the literature focusing on U.S. and perform a robustness exercise, in which I decrease the provincial border dummy in the economy and recalibrate the model. Results are robust to this alternative domestic trade cost structure.

Table 5: Non-targeted Moments

	Data	Model
Skill Premium Change and Distance to Port:		
Coastal minus non-coastal province	$\approx 5\%$	4.3%
-Distance measure	$\approx -1.9\%$	-1.6%
City Growth and Distance to Port:		
City-level regression	-3.4%	-4.8%
-within province variation only	-7.9%	-11.1%

Notes: The upper panel compares the model predictions on the impacts of trade liberalization on skill premia to the empirical estimates from Han et al. (2012). The lower panel compares the model predictions on population growth to the empirical counterparts, reported in Appendix C.2

4.7 Model Validation

Given the estimated domestic trade costs, international trade liberalization in China will have differential impacts on coastal and interior cities. I use these heterogeneous effects to validate the model.

Specifically, I solve for an autarky equilibrium by shutting down the trade between China and the RoW. I then calculate for each city the percentage *change* in the skill premium and population from the autarky equilibrium to the calibrated economy, and interpret these changes as the model-predicted impacts of trade liberalization. Using these simulated data, I estimate the heterogeneous impacts of trade on population and skill premia of cities with differential accesses to the international market. I then compare them to the empirical estimates identified of the same variation.

Table 5 summarizes the results; underlying regression tables and additional information are provided in Online Appendix C.2. The upper panel focuses on skill premium changes across cities. Han et al. (2012) estimates the effect of international trade on skill premia between 1988 and 2008, using two episodes of liberalization (Deng Xiaoping's Southern Tour in 1992 and the WTO accession in 2001). Their findings suggest that on average, trade liberalization increases skill premia by around 5% more in coastal provinces than in interior provinces. They also use distance to the coast as a measure for a city's exposure to trade and find a coefficient of -1.9%. Using the same set of provinces as in their study and run the same regression using model-simulated data, I find that skill premia increase more in coastal provinces than in interior province by 4.3%. Using distance to coast as a measure in the simulated regression, I find a coefficient of -1.6%, which is also close to its empirical counterpart.

The lower panel of Table 5 reports the comparison between the model predictions

and the empirical counterparts on trade-induced population reallocation. In the model, trade liberalization increases the relative income of coastal regions, which attract more migrants. The elasticity of population growth with respect to a city's distance to port is -4.8%. Using population growth during 2000 and 2010, a period of trade growth after China's WTO accession, I estimate the empirical relationship between city growth and distance to port, and finds a coefficient of -3.4%—coastal cities also experienced faster population growth during this period. I also verify whether the pattern holds within a province by controlling for provincial fixed effects. In this case, the coefficients in both simulated and real regressions increase.

Overall, these validation exercises confirm that the model's predictions on the heterogeneous impacts of trade are in line with the data.

5 Counterfactual Experiments

5.1 Impacts of Trade on Welfare and Inequality

I analyze the aggregate and distributional impacts of trade through the lens of this model. The first experiment is the same as in the validation exercise: I shut down international trade between China and the RoW, keeping all other parameters the same.

5.1.1 The Gains from Trade for Different Regions

I calculate the welfare gains from trade for each worker group as the percentage change in their real income from autarky to the calibrated equilibrium. Panel A of Table 6 reports the results. The population-weighted welfare gains across all worker groups are around 7.5%, so China as a whole benefits significantly from international trade. However, the welfare gains do not accrue to everyone equally. First, skilled and unskilled workers are affected differently by trade. The average gains from trade are about 13% for skilled workers, and only 6% for unskilled workers. Due to spatial frictions, skill premia differ across regions. Measuring the aggregate skill premium using the average percentage wage difference between skilled and unskilled workers, I find that international trade increases the skill premium by 5.5%. Second, within each skill group, the impacts of trade also differ dramatically: for both skilled and unskilled workers, the standard deviations of the distributions of the welfare gains are larger than the respective means. The most-benefited group receives a welfare improvement of 30%, while some unskilled workers could experience welfare losses.

These patterns suggest that international trade might have important impacts on inequality, between workers with different skills, and among similar workers from different

Table 6: Aggregate and Distributional Impacts of Trade

A. Gains from Trade for Different Worker Groups				
	Mean	std	5th percentile	95th percentile
Skilled	13.3	12.6	1.8	37.2
Unskilled	6.0	10.8	-1.6	28.4
National Average	7.5			
Skill Premium	5.5			
B. Impacts of Trade on Inter- and Intra-Regional Inequality				
	Between Region	Within Region	Theil Index	
Open Economy	0.223	0.030	0.253	
Autarky	0.211	0.026	0.237	
Increase (%)	5.7%	15.4%	6.8%	
Relative Contribution	75%	25%	100%	

Notes: Panel A of this table reports the summary statistics of the city-level welfare gains for skilled and unskilled workers. All numbers are in percentage points. Panel B reports the decomposition of inequality, measured by the Theil index, into within- and between-region components in both autarky and the open economy. The last row reports the relative contributions of the two components to the increase in the aggregate inequality after trade liberalization.

regions. I use the Theil index to measure the overall inequality in real wages in China, decompose it into between-region and within-region components, and examine the impacts of international trade on each component.

Panel B of Table 6 reports the results. The first row is the decomposition for the benchmark economy. The between-region component constitutes about 90% of the overall inequality in China, while the within-region inequality between skilled and unskilled workers contributes only 10%. The second row of the table is the decomposition for the autarky economy. Again, the between-region component explains most of the overall inequality. Given the large spatial inequality in China, this is hardly surprising.

The third row of Panel B indicates moving from the autarky economy to the open economy increases the overall inequality in the country by around 7%; both between- and within-region inequality increase. Although the within-region component accounts for only about 10% of inequality, its contribution to the increase is 25%. The between-region component accounts for the remaining 75% of the increase in aggregate inequality.

5.1.2 Trade and Domestic Inequality: the Role of Internal Geography

The decomposition in Table 6 suggests that both within- and between-region components matter for inequality, but the latter is more important. Since a major difference between regions is their accesses to foreign markets, I examine to what extent geography can explain the impacts of trade on different regions.

Figure 3 plots the relationship between access to foreign markets, proxied by a city's distance to the nearest port, and city-level average welfare gains from trade for skilled and unskilled workers. The size of dots indicates city size. For both types of workers, regions

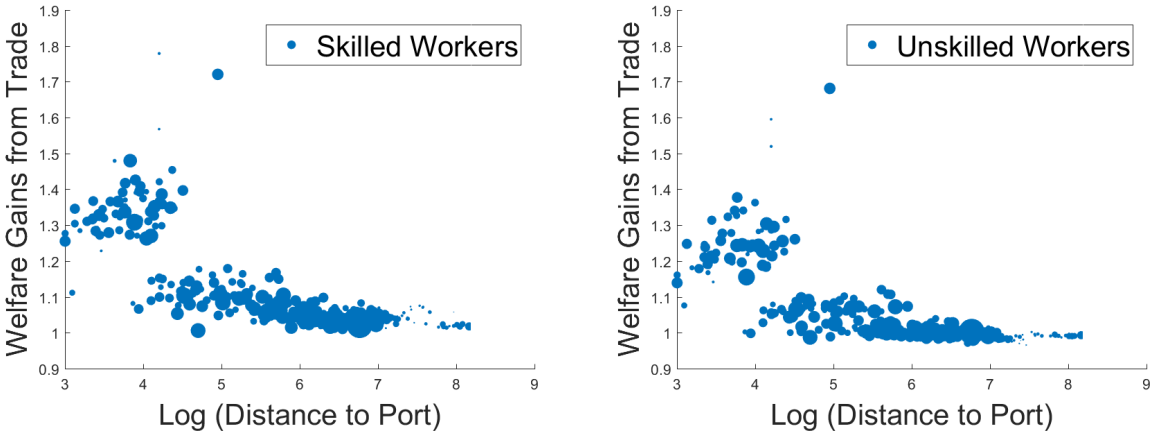


Figure 3: Geographic Distribution of Gains from Trade

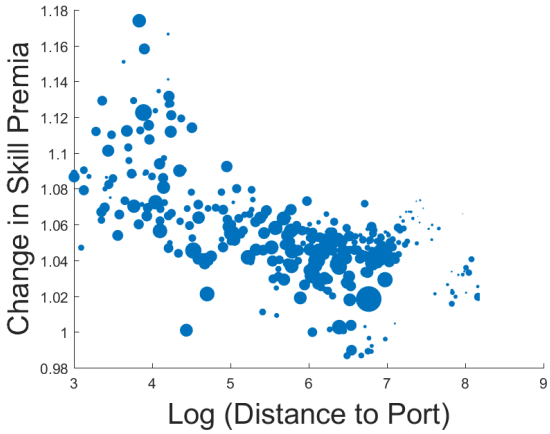


Figure 4: Trade and the Skill Premium

form two groups in terms of their gains from trade: a coastal group that reaps most of the benefits and an interior group that benefits very little. The segregation of gains from trade is reminiscent of the higher openness in the coast than in interior, documented by Panel (a) of Figure 2. By limiting free mobility of goods within the country, intra-national trade costs prevent interior regions from benefiting from trade.

To illustrate the impacts of international trade on within-region inequality and how the impacts differ across regions, Figure 4 plots *changes* in skill premia against distance to the nearest port. Trade liberalization increases skill premia in almost all cities, but the increases are far from even. The increases could be as high as 15% for cities along the coast, and as low as only 2% for cities in the interior. As discussed in Table 5, the negative correlation between skill premium increases and distance to port is also quantitatively consistent with empirical findings (Han et al., 2012).

Figures 3 and 4 illustrate clearly that within-country geography is relevant for work-

ers' gains from trade. The prediction that workers from coastal regions benefit more is intuitive: international trade is, on average, welfare-improving. Since coastal regions trade more, workers there also benefit more. On the other hand, the forces behind the differential impacts of trade on skilled and unskilled workers within the same region (i.e., the changes in skill premia) and how the impacts vary systematically across locations are less obvious. I explore the underlying mechanisms in the next subsection.

5.1.3 Explaining the Gradient of Changes in Skill Premia

The impacts of international trade on skill premia rest on its impacts on the relative demand and the relative supply of skilled and unskilled workers. I discuss forces affecting these two factors separately.

First, the factor content theory of trade predicts that cities specializing in sectors using skilled workers more intensively will see increases in skill premia. In my model, this mechanism interacts with domestic trade costs and operates through within-country specialization. Specifically, because urban tradable industries (K and M industries) employ intermediate goods more heavily than the agricultural industry, they are more "transportation intensive". When a country opens up to trade, coastal regions—due to their proximity to foreign suppliers—have stronger comparative advantages in these industries, and become more specialized in capital and manufacturing production. The interior regions, on the other hand, increasingly specialize in the agricultural industry. This shift in specialization patterns increases the relative demand for skilled workers in the coastal regions and decreases it in the hinterland, resulting in a negative relationship between increases in regional skill premia and distance to port. Because this channel works through the factor content of *intra-national* trade, I call it the "Domestic Stolper-Samuelson Effect".[25]

A second channel that affects relative demand for skilled workers is the capital-skill complementarity in production. China is a net importer of capital goods, which are complements to skilled workers. As a result, after the international trade liberalization, skill premia increase across the board. Because coastal regions experience larger drops in the prices of capital goods, skill premia increase more on the coast.

Now consider changes in relative supply of skilled workers after trade liberalization. Since coastal regions gain more from trade, they will experience a net inflow of migrants. As the estimated migration costs are lower for skilled workers, there will be more skilled

[25]The standard Stolper-Samuelson effect, which operates through international specialization, is less important in my setting: because the trade between China and the RoW is largely within the broad sectors defined in this paper, the change in relative demand for workers induced by the factor content of *international* trade is likely small. For example, the trade surplus is only 1% of production in the agricultural industry, and the trade deficit in capital and equipment industry accounts for less than 2% of production.

workers moving from interior into coastal regions, pushing down skill premia on the coast and driving them up in the interior. The differential mobilities between skilled and unskilled workers constitute a third channel that tends to offset the channels described above and flatten the gradient of changes in skill premia.

To illustrate the quantitative importance of these channels, I conduct a sequence of counterfactual exercises and plot the changes in skill premia in these experiments in Figure (5). For ease of comparison, I plot only the fitted value from a weighted least squares regression of changes in skill premia on regions' distances to port. "Benchmark" refers to the previous experiment. "Case 1" increases skilled workers' migration costs to the level of unskilled workers; "Case 2" further shuts down capital skill complementarity by setting both the elasticity between capital and skilled worker and the elasticity between equipped and unskilled worker to 1.1, the estimates of Dix-Carneiro (2014) using a symmetric CES specification. In both cases, I compute the open economy and autarky equilibria and calculate changes in skill premia from autarky to openness.

As shown in Figure 5, when migration costs are the same for both types of workers, the gradient of the changes in skill premia with respect to the distance to port becomes steeper. Coastal regions now experience around 13% increases in skill premia, 2 percentage points higher than in the benchmark experiment, and interior regions now experience barely any increase in the skill premium. When I further shut down capital-skill complementarity, the fitted line shifts downward and becomes flatter. Globalization now increases within-region inequality more evenly across regions.

Apart from capital-skill complementarity, the two other channels work through reallocation of workers across regions or sectors. The implications of these channels are that the coastal regions will: 1) trade more, 2) become more specialized in urban sectors, and 3) attract skilled workers disproportionately. Online Appendix C.3 shows that indeed these channels are at play. It further shows that, if we keep all regional heterogeneity but assume domestic trade is cost-free, the three above predictions are no longer true, and that the distributional impacts of trade along the geographic dimension also vanish—in short, the above results arise only because of domestic spatial frictions, rather than regional heterogeneity in production technologies.

Taken together, these experiments suggest that the various channels related to internal geography are all quantitatively important for both the geographic and the skill dimensions of the distributional impacts. In particular, to my knowledge, the "Domestic Stolper-Samuelson Effect" is novel to the literature. Operating through changes in domestic specialization, this channel has important implications for measuring regional trade exposures: since interior regions trade little with the RoW, most conventional measures of trade exposure will overlook these regions' exposures. However, because of international

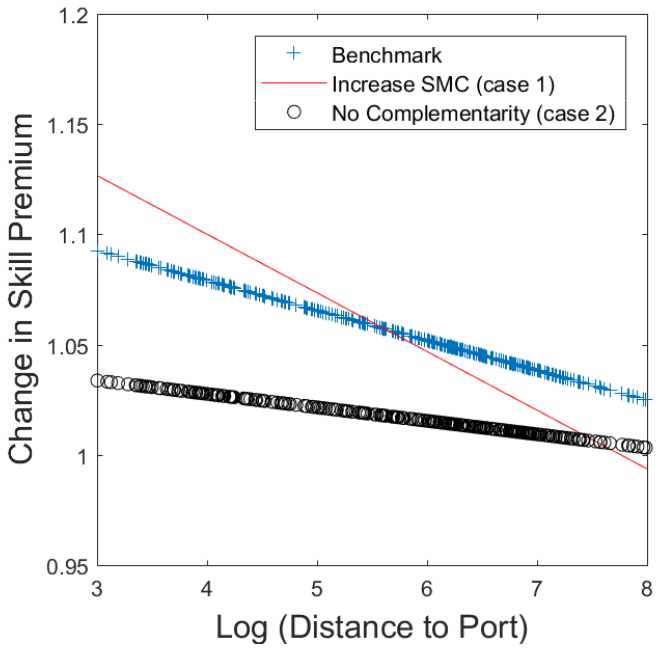


Figure 5: Changes in the Skill Premium

Notes: Each line is the fitted value from the weighted (by city size) least squares regression of changes in skill premia on the log of cities' distances to coast. Case 1: equal migration costs for both types of workers. Case 2: equal migration costs *and* no capital-skill complementarity.

trade liberalization, the economic environment of these regions change dramatically. In measuring the regional impacts of trade, it is therefore important to take into account not only a region's international trade participation but also its changing trade patterns with domestic partners.

5.1.4 Comparison to a Frictionless Model

It is instructive to compare my model to an otherwise similar model with free domestic trade and migration. To this end, I calibrate a frictionless model to match the spatial distribution of production and international trade. I then use this model to perform the same gains from trade experiment.

Similar to the benchmark model, this frictionless model predicts the national average welfare gains from trade to be 7.5%. However, the two models have very different implications on the effects of trade on inequality. In the frictionless economy, trade increases the aggregate Theil index by 4.7%, much lower than in the benchmark economy (6.8%). Intuitively, when everyone in the economy has the same access to the international market, the gains from trade are shared across regions, so the aggregate inequality increases by less.

Focusing on the inequality between skilled and unskilled workers, the predictions are

also significantly different—the aggregate skill premium increases by 12% in response to trade liberalization in the frictionless model, as opposed to 5.5% in the benchmark model. This contrast is driven by differential geographic distributions of trade shocks in these two models. In the benchmark model, the benefits of trade, which are higher for skilled workers, accrue primarily to coastal regions. This increases the share of skilled workers migrating to the coast. Selection on productivity in migration then implies that workers end up in the coast will have lower average productivity, while workers end up in the hinterland will have higher average productivity. Because skill premia increase by more in the coast, this compositional change dampens the effect of trade on the aggregate skill premium.

Overall, this comparison demonstrates that taking into account domestic spatial frictions matters for understanding not only the effects of trade for specific geographic regions or worker groups, but also the overall inequality and the aggregate skill premium.

5.2 Domestic Reforms and International Trade

5.2.1 Effects of Domestic Trade and Hukou Reforms

Over the past decade, along with many other developing countries, China has improved domestic transportation infrastructure and started structural reforms that aim to reduce domestic migration and trade costs. This give rise to a host of questions. How would these reforms affect the welfare and inequality within China? What are their implications for the distributional impacts of trade?

To answer these questions, I introduce three types of reforms into the benchmark economy. In the first reform, the provincial "border effect" in China's domestic trade—the trade costs associated with crossing a provincial border per se—is set to 0.65, the estimate for the U.S. in the literature. In the second reform, I eliminate distortions from the Hukou system for workers in the economy by setting the Hukou scores for all cities to 6. After this reform, the cost of moving to another city is given by Equation (19), with $\Delta\text{HukouScore}_d$ being the difference between 6 and city d's Hukou score in 2005. The third reform combines the first two. In each of these three cases, I solve the model for an open-economy equilibrium, changing only coefficients corresponding to a specific reform. Compared to the benchmark economy, these equilibria capture the effects of hypothetical reforms.

Panel (A) of Table 7 reports the effects of these reforms on the average welfare and various measures of domestic inequality. The second column is the experiment with lower intra-national trade costs. Freer domestic trade clearly benefits everyone—the average welfare gains from the reform are around 24%. Because of the improved access to capital and equipment, skilled workers, especially those in the hinterland, tend to benefit

Table 7: Domestic Reforms and Interaction with International Trade

A. Effects of Domestic Reforms	Benchmark	(2) Trade	(3) Hukou	(2)+(3)
Average Welfare	-	24.2	10.4	36.6
Skill Premium	-	12.1	-0.3	11.8
Inequality Increase	-	2.4	-18.5	-14.7
Contribution-Between	-	-18%	116%	141%
Contribution-Within	-	118%	-16%	-41%

B: Gains from Trade for Post-Reform Economies	Benchmark	(2) Trade	(3) Hukou	(2)+(3)
$\text{Trade/GDP}_{\text{China}}$	64.1	51.0	56.5	45
$\text{Trade/GDP}_{\text{RoW}}$	8.0	7.7	8.9	8.6
Average Welfare	7.5	5.9	6.7	5.4
Skill Premium	5.5	4.2	4.1	3.29
Inequality Increase	6.7	4.9	3.2	3.0
Contribution-Between	75%	70%	42%	43%
Contribution-Within	25%	30%	58%	57%

Notes: Panel A reports the effects of domestic reforms on the aggregate welfare and inequality. Panel B reports international trade statistics in the post-reform economies, and differences in welfare and inequality between post-reform economies and the corresponding autarky equilibrium.

more than the unskilled workers. As a result, the aggregate skill premium increases by 12%. The overall Theil index increases modestly by 2.4%, which is entirely driven by the within-region component.

The third column reports the effects of an overhaul of the Hukou system. The average welfare gains from such a reform are 10.4%. Since the reform apply to both skilled and unskilled workers, its impacts are largely skill-neutral—the aggregate skill premium barely changes. Because workers can now move more easily across regions, the overall domestic Theil index decreases significantly by 18.5%. This decrease is entirely due to the smaller between-region inequality. Finally, the last column reports the effects of the comprehensive reform. The welfare gains from this reform are significant, at 36.6%. The skill premium increase by 11.8%, largely because of the goods market reform. Not surprisingly, the overall inequality decreases significantly by 14.7%. As in the previous two experiments, the between-region component of the aggregate inequality decreases and accounts entirely for the decrease in the overall inequality.

5.2.2 International trade in the Post-Reform Economies

How do domestic reforms affect the distributional impacts of trade for China? I send the post-reform economies to autarky and calculate the changes. Importantly, these results

should be interpreted differently from those in Section 5.1.4 because they correspond to different thought experiments. In this subsection, we take the benchmark economy as given and ask how large the cost of moving to autarky would be if certain reforms are first implemented. Results in Section 5.1.4, on the other hand, are comparison between models: how the "inferred" gains from trade will change, if we force a frictionless model to fit the data.

Panel B of Table 7 reports international trade statistics in the post-reform economies (first two rows) and the effects of moving the post-reform economies to autarky (the rest). The first column is the calibrated benchmark economy. As we can see from the second column, after the domestic trade reform, international trade as a share of GDP decreases by one-fifth, from 64% to 51%. To disentangle changes in trade intensity from changes in trade volume, I normalize trade by the RoW GDP, which is not much affected by China's reform, and report this ratio as trade volume in the second row. This measure decreases by only 4%, from 8.0% to 7.7%, so the trade share is lower mainly because trade intensity is lower.

Two channels explain the trade share decreases. First, domestic reforms improve China's productivity, increasing its share in the world economy. As a large economy, China faces an upward-sloping supply curve from the RoW. As the Chinese economy expands, its terms-of-trade deteriorates. As a result, China trade less intensively with the RoW and benefits less from trade (see, Dornbusch et al., 1977, for a theoretical analysis of this channel). Second, the modest decrease in trade volume comes from the trade diversion effect. When domestic trade costs decrease, on the one hand, the trade between interior regions and the RoW increases; on the other hand, the foreign trade of the coastal region is diverted to with interior regions. Whether the trade volume increases or decreases depends on the strength of these two forces. In Online Appendix C.4, I show that at the calibrated equilibrium, as domestic trade costs are lowered, trade volume first decreases, and then increases. This result is robust to alternative values for the trade elasticity, θ.

Because of the lower international trade intensity, it is not surprising that a move to autarky is less costly—the gains from trade decrease to 5.9%. These gains, however, are distributed more evenly: the increases in the aggregate skill premium and the Theil index are both smaller than in the benchmark economy. The between-region component also becomes less important for the aggregate Theil index increase—with lower domestic trade costs, interior regions now benefit relatively more.

When the Hukou system is abolished, China's aggregate productivity improves, thus slightly reducing its trade share through the terms-of-trade effect, as shown in the third column. Because the "trade diversion" channel in the case of the domestic trade reform

is no longer at play, we see an increase in trade volume. In this case, the average gains from trade decrease (by 11%), but the trade-induced inequality decreases much more: the increase in the skill premium and the overall Theil index are 25% and 50% smaller, respectively. Since workers can now more freely move across regions, the spatial component is much less important for the overall inequality. Finally, when trade and Hukou reforms are combined, the increase in the skill premium and the Theil index are further reduced. On the flip side, the aggregate gains from trade also decrease—as China become more productive through internal reforms, international trade becomes relatively less crucial for China. Nevertheless, the trade volume still increases.

While domestic reforms might reduce a large country's intensity of international trade for *a given level* of international trade costs, it increases the response of the economy to *changes* in trade costs. Intuitively, with lower domestic migration costs, it will be easier for workers to reallocate to coastal regions; with lower domestic trade costs, tariff cuts will also have a larger impact on interior regions. The question is whether such effects are quantitatively significant.[26] Figure 6 plots percentage changes in trade share in both benchmark and post-reform economies against decreasing international trade costs. For small reductions in trade cost, domestic reforms matter very little. However, the post-reform economies are much more responsive to large reductions in trade costs: when the trade cost is 50% lower, the economy with both trade and Hukou reforms predict that the trade share will increase by 480%, about a-third larger than the prediction of the benchmark economy. Economies with only trade or Hukou reforms fall in-between the two cases.

5.3 Sensitivity Analysis

I perform a set of analysis to test the robustness of the results to external parameters. For each parameter, I recalibrate the model to match the open economy equilibrium, and then calculate the distributional effects of trade. These robustness exercises include: different degrees of capital-skill complementarity (ρ_{kh}, ρ_{lkh}); different correlations between individuals' productivity draws across regions (ρ); different trade elasticities (θ). In an extension, I also incorporate remittances, so that trade shocks might affect interior regions through remittances of migrants. I find similar results in these experiments. Details could be found in Appendix C.6.1.

In the quantitative analysis, I also make some assumptions, which, while standard in the literature, are admittedly strong. For example, I assume that workers learn their

[26] In this model, domestic trade and migration costs will not have a first-order effect on international trade elasticities, but for large shocks they could still matter quantitatively. See Online Appendix C.5 for a detailed discussion.

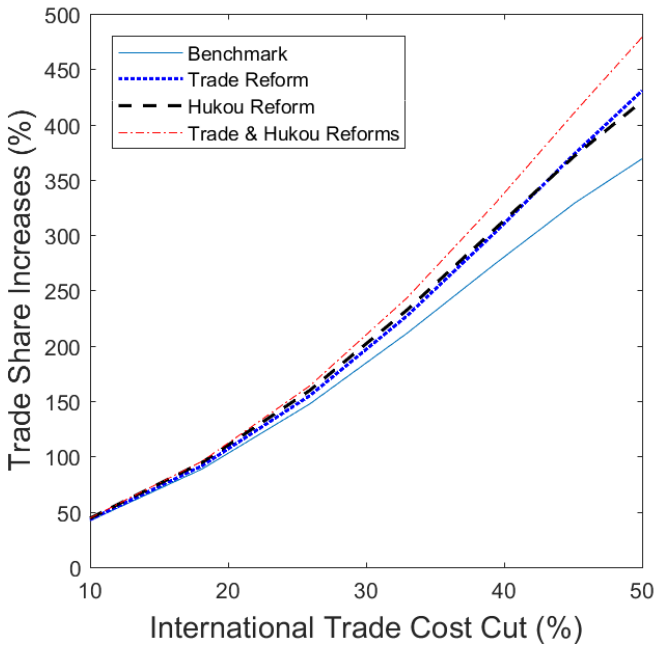


Figure 6: Future Trade Liberalization

Notes: Percentage changes in trade share as international trade costs decrease for: 1) Benchmark (solid line); 2) After trade reform (dotted line); 3) After Hukou reform (dashed line); and 4) After trade and Hukou reforms (dash-dot line)

idiosyncratic productivity shocks before migration, which rules out uncertainty in return to migration; I also calibrate a static model to China, which is a dynamic economy. In Appendix C.7, I discuss how the violation of these assumptions will change the results.

6 Conclusion

This paper studies the impacts of international trade liberalization on the inequality of China, a country characterized by high domestic trade and migration costs. I show that in this setting, international trade affects the welfare and skill premia differently across regions, and that overlooking domestic frictions will lead to significantly different predictions on inequality, measured by either the overall Theil index or the aggregate skill premium. Given that both trade and migration frictions have been documented to be important in many other developing countries (see, Footnote 1, for relevant literature), the main channels in this paper and the qualitative predictions likely generalize to these countries, too.

Hypothetical trade and Hukou reforms within China effectively increase the size of the Chinese economy, which might slightly reduce the gains from trade if the tariffs are kept at the current level. On the other hand, these reforms allow the gains from trade to

be distributed more evenly while increase the response of the economy to future liberalizations.

This paper abstracts from some interesting and important aspects of the real world that could affect the impacts of international trade liberalization. For example, regional agglomeration effects might amplify both the distributional and the aggregate impacts. Agglomeration and dynamic effects are potentially important features to incorporate into future research, especially when analyzing an emerging economy like China.

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.

Alvarez, Fernando and Robert E Lucas, "General Equilibrium Analysis of the Eaton–Kortum Model of International Trade," *Journal of Monetary Economics*, 2007, *54* (6), 1726–1768.

Atkin, David and Dave Donaldson, "Who's Getting Globalized? The Size and Nature of Intranational Trade Costs," *NBER Working Paper 16778*, 2015.

Au, Chun-Chung and J Vernon Henderson, "How Migration Restrictions Limit Agglomeration and Productivity in China," *Journal of Development Economics*, 2006, *80* (2), 350–388.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–68.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from indonesia," *Manuscript*, 2018.

Burstein, Ariel, Javier Cravino, and Jonathan Vogel, "Importing Skill-Biased Technology," *American Economic Journal: Macroeconomics*, 2013, *5* (2), 32–71.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2015.

_ , Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *NBER Working Paper*, 2017, *77999.*

Coşar, A Kerem and Banu Demir, "Domestic Road Infrastructure and International Trade: Evidence from Turkey," *Journal of Development Economics*, 2016, *118*, 232–244.

Cosar, A Kerem and Pablo D Fajgelbaum, "Internal Geography, International Trade, and Regional Specialization," *American Economic Journal: Microeconomics*, 2016, *8* (1), 24–56.

Crafts, Nicholas and Alexander Klein, "Geography and Intra-national Home Bias: U.S. Domestic Trade in 1949 and 2007," *Journal of Economic Geography*, 2014.

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

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885.

_ and Brian K Kovak, "Trade liberalization and regional dynamics," *American Economic Review*, 2017, *107* (10), 2908–46.

Donaldson, Dave, "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 2018, *108* (4-5), 899–934.

Dornbusch, Rudiger, Stanley Fischer, and Paul Anthony Samuelson, "Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods," *The American Economic Review*, 1977, *67* (5), 823–839.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi, "Slicing the pie: Quantifying the aggregate and distributional effects of trade," *NBER Working Paper*, 2017, *23737.*

Goldberg, Pinelopi Koujianou and Nina Pavcnik, "Distributional Effects of Globalization in Developing Countries," *Journal of Economic Literature*, 2007, *45* (1), 39–82.

Gollin, Douglas, David Lagakos, and Michael E Waugh, "The agricultural productivity gap," *The Quarterly Journal of Economics*, 2013, *129* (2), 939–993.

Han, Jun, Runjuan Liu, and Junsen Zhang, "Globalization and Wage Inequality: Evidence from Urban China," *Journal of International Economics*, 2012, *87* (2), 288–297.

Kambourov, Gueorgui, "Labour market regulations and the sectoral reallocation of workers: The case of trade reforms," *The Review of Economic Studies*, 2009, *76* (4), 1321–1358.

Kanbur, Ravi and Anthony J Venables, *Spatial Inequality and Development*, Oxford University Press, 2005.

_ and Xiaobo Zhang, "Fifty Years of Regional Inequality in China: a Journey Through Central Planning, Reform, and Openness," *Review of development Economics*, 2005, *9* (1), 87–106.

Kinnan, Cynthia, Shing-Yi Wang, and Yongxiang Wang, "Access to migration for rural households," *American Economic Journal: Applied Economics*, Forthcoming.

Krugman, Paul and Raul Livas Elizondo, "Trade Policy and the Third World Metropolis," *Journal of Development Economics*, 1996, *49* (1), 137–150.

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.

Limão, Nuno and Anthony J Venables, "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade," *The World Bank Economic Review*, 2001, *15* (3), 451–479.

Monte, Ferdinando, "Local Transmission of Trade Shocks," *Unpublished Manuscript*, 2016.

Munshi, Kaivan, "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 2003, *118* (2), 549–599.

Parro, Fernando, "Capital-skill Complementarity and the Skill Premium in a Quantitative Model of Trade," *American Economic Journal: Macroeconomics*, 2013, *5* (2), 72–117.

Poncet, Sandra, "A Fragmented China: Measure and Determinants of Chinese Domestic Market Disintegration," *Review of international Economics*, 2005, *13* (3), 409–430.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez, "Trade, domestic frictions, and scale effects," *American Economic Review*, 2016, *106* (10), 3159–84.

Redding, Stephen J, "Goods trade, factor mobility and welfare," *Journal of International Economics*, 2016, *101*, 148–167.

Simonovska, Ina and Michael E Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Sun, W, C Bai, and P Xie, "The Impact of Hukou Reforms on Rural Labor Migration in China," *Economic Research Journal*, 2011, pp. 28–41.

Swiecki, Tomasz, "Intersectoral Distortions and the Welfare Gains from Trade," *Journal of International Economics*, 2017, *104*, 138–156.

Tombe, Trevor and Xiaodong Zhu, "Trade, migration and productivity: A quantitative analysis of china," *Manuscript, University of Toronto*, 2017.

Topalova, Petia, "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India," *American Economic Journal: Applied Economics*, 2010, pp. 1–41.

Venables, Anthony J and Nuno Limão, "Geographical Disadvantage: A Heckscher–Ohlin–von Thünen Model of International Specialisation," *Journal of International Economics*, 2002, *58* (2), 239–263.