Online Appendix for:
Internal Geography, Labor Mobility, and the Distributional Impacts of Trade

Jingting Fan*
Pennsylvania State University

Contents

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

A Algebra

A.1 Deriving Equation (3)

$$\begin{aligned}
\pi^e_{o,d} &= \Pr(\frac{v^e_d z_d}{d_{o,d}} \geq \frac{v^d_g z_g}{d_{o,g}}, \forall g \in G) \\
&= \Pr(z_g \leq \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d, \forall g \in G,) \\
&= \int_0^\infty \Pr(z_g \leq \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d, \forall g \in G | z_d) f(z_d) dz_d \\
&= \int_0^\infty F_d(\frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_1}{d_{o,1}}} z_d, \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_2}{d_{o,2}}} z_d, ... \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d, ...) dz_d,
\end{aligned}$$

Where $F_d(\frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_1}{d_{o,1}}} z_d, \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_2}{d_{o,2}}} z_d, ... \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d, ...) := \frac{dF}{dz_d}|_{z_g = \frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d, \forall g \in G}$ is the probability that the draw from region d is z_d *and* this draw dominates all other draws.

Use the functional form of F, it follows that

$$\begin{aligned}
\pi^e_{o,d} &= \int_0^\infty \exp(-(\sum_{g \in G} (\frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d)^{-\epsilon_e})^{1-\rho}) * (1-\rho)\epsilon_e (\sum_{g \in G} (\frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d)^{-\epsilon_e})^{\rho} z_d^{-\epsilon_e - 1} dz_d \\
&= \frac{\int_0^\infty d\exp(-(\sum_{g \in G} (\frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}} z_d)^{-\epsilon_e})^{1-\rho})}{\sum_{g \in G} (\frac{\frac{v^e_d}{d_{o,d}}}{\frac{v^e_g}{d_{o,g}}})^{-\epsilon_e}} \\
&= \frac{(\frac{v^e_d}{d_{o,d}})^{\epsilon_e}}{\sum_{g \in G} (\frac{v^e_g}{d_{o,g}})^{\epsilon_e}}
\end{aligned}$$

A.2 Deriving Equation (5)

In the first step, I derive the expected value of the destination-specific component in workers' indirect utility (Equation 1), u^e_o, for workers moving from o to d, denoted $E(u^e_o | l^e_{o,d})$, where $u^e_o := \max_{d \in \mathbf{G}} \{\frac{W^e_d B_d z_d}{P_d d^e_{o,d}}\} = \max_{d \in \mathbf{G}} \{\frac{v^e_d z_d}{d^e_{o,d}}\}$.

I first derive the distribution of u_o^e, $u_o^e = \max_{d\in G}\{\frac{v_d^e z_d}{d_{o,d}}\}$,

$$
\begin{aligned}
F_{u_o^e}(u) :&= \text{Prob}(u_o^e \leq u) \\
&= \text{Prob}(\frac{v_d^e z_d}{d_{o,d}} \leq u, \quad \forall d \in G) \\
&= \text{Prob}(z_d \leq \frac{u d_{o,d}}{v_d^e}, \quad \forall d \in G) \\
&= F(\frac{u d_{o,1}}{v_1^e}, \frac{u d_{o,2}}{v_2^e}, ..., \frac{u d_{o,d}}{v_d^e}, ...) \\
&= \exp(-[\sum_{d\in G}(\frac{u d_{o,d}}{v_d^e})^{-\epsilon_e}]^{1-\rho}) \\
&= \exp(-[\sum_{d\in G}(\frac{d_{o,d}}{v_d^e})^{-\epsilon_e}]^{1-\rho} u^{-(1-\rho)\epsilon_e}) \\
&\equiv \exp(-\Phi_o^{e\,1-\rho} u^{-(1-\rho)\epsilon_e})
\end{aligned}
$$

It can be shown that, $\forall d \in G$, the cumulative distribution function of u for workers moving from o, to d, is

$$F_{u_{o,d}^e}(u) = F_{u_o^e}(u) = \exp(-\Phi_o^{e\,1-\rho} u^{-(1-\rho)\epsilon_e}),$$

which is a Frechet distribution with position parameter $\Phi_o^{e\,1-\rho}$ and dispersion parameter $(1-\rho)\epsilon_e$.[1]

[1]This is obtained by showing $F_{u_{o,d}^e}(u) := \text{Prob}(u_{o,d}^e \leq u | u_{o,d}^e \text{is the highest}) = \frac{\text{Prob}(u_{o,d}^e \leq u, u_{o,d}^e \text{is highest})}{\pi_{o,d}^e} = \frac{\int_0^u F_d(z_d) dz_d}{\pi_{o,d}^e} = F_{u_o^e}(u)$.

$$
\begin{aligned}
E(u_o^e|l_{o,d}) &= \int u dF_{u_o^e}(u) \\
&= \int u d(\exp(-[\sum_{d\in G}(\frac{d_{o,d}}{v_d^e})^{-\epsilon_e}]^{1-\rho}u^{-(1-\rho)\epsilon_e})) \\
&= \int u\epsilon_e(1-\rho)\exp(-[\sum_{d\in G}(\frac{d_{o,d}}{v_d^e})^{-\epsilon_e}]^{1-\rho}u^{-(1-\rho)\epsilon_e})[\sum_{d\in G}(\frac{d_{o,d}}{v_d^e})^{-\epsilon_e}]^{1-\rho}u^{-(1-\rho)\epsilon_e}du \\
&= \int u\epsilon_e(1-\rho)\exp(-\Phi_o^{e\,1-\rho}u^{-(1-\rho)\epsilon_e})\Phi_o^{e\,1-\rho}u^{-(1-\rho)\epsilon_e}du \\
&= -\int \epsilon_e(1-\rho)\exp(-y)ydu \quad (\text{change of variable}: y = \Phi_o^{e\,1-\rho}u^{-(1-\rho)\epsilon_e}) \\
&= -\int \epsilon_e(1-\rho)\exp(-y)yd(\frac{y}{\Phi_o^{e\,1-\rho}})^{\frac{1}{-(1-\rho)\epsilon_e}} \\
&= \int \exp(-y)y^{-\frac{1}{(1-\rho)\epsilon_e}}\Phi_o^{e\,\frac{1}{\epsilon_e}}dy \\
&= \Phi_o^{e\,\frac{1}{\epsilon_e}}\Gamma(1-\frac{1}{\epsilon_e(1-\rho)}) \quad (\text{Definition of Gamma function})
\end{aligned}
$$

In the second step, we use $E(u_o^e|l_{o,d}^e)$ to derive $E(z_d^e|le)$, the expected productivity of the workers who move from o to d,

$$
\begin{aligned}
E(z_d^e|l_{o,d}^e) &= E(\frac{u_o^e d_{o,d}^e}{v_d^e}|l_{o,d}^e)(\text{for } l_{o,d}^e, u_o^e = \max_{g\in\mathbf{G}}\{\frac{v_g^e z_g}{d_{o,g}^e}\} = \frac{v_d^e z_d}{d_{o,d}}) \\
&= \frac{d_{o,d}^e}{v_d^e}E(u_o^e|l_{o,d}) \\
&= \Phi_o^{e\,\frac{1}{\epsilon_e}}\Gamma(1-\frac{1}{\epsilon_e(1-\rho)})\frac{d_{o,d}^e}{v_d^e}
\end{aligned}
$$

A.3 Deriving Equation (17)

For workers staying in their hometown, $u_{o,o}^e = \frac{v_o^e z_o}{d_{o,o}} = v_o^e z_o$, hence the distribution of productivity draws for workers choosing to stay in o is:

$$
\begin{aligned}
F_{z_{o,o}^e}(z) :&= \Pr(z_{o,o}^e < z) \\
&= \Pr(\frac{u_{o,o}^e}{v_o^e} < z) \quad (\text{using } d_{o,o} = 1) \\
&= F_{u_{o,o}^e}(zv_o^e) \\
&= \exp(-[v_o^{-(1-\rho)\epsilon_e}\Phi_o^{e\,1-\rho}]z^{-(1-\rho)\epsilon_e}),
\end{aligned}
$$

which is also a Frechet distribution. For different regions, the productivity distribution of stayers there have different means, but their dispersions will be the same. Therefore, I regress stayers' log wages on regional fixed effects to net out the different average regional productivity draws and interpret the exponent of the residuals as random draws from a Frechet distribution with dispersion parameter $\epsilon_e(1-\rho)$. The coefficient of variations for this distribution is given by Equation (17).

A.4 Identification of Amenities-Adjusted Real Wages $\{v_d^e\}$

Proposition 1 is used in Section 3.C of the main text, where I estimate migration costs.

Proposition 1 *Given migration costs $\{d_{o,d}\}$, there exists a unique set of $\{v_d\}$ (up to normalization), such that the model-predicted number of workers employed in each region equals that in the data, i.e., $L_d^e = \sum_{o\in \mathbf{G}} \pi_{o,d}^e l_o^e$ is satisfied, where L_d^e is the number of workers working in d (data), l_o^e is the number of workers born in o (data), and $\pi_{o,d}^e$ is the model-predicted probability of workers born in o to move to d.*

Proof The proof follows Michaels, Redding and Rauch (2011) and Lemma 1, Lemma 2 in Ahlfeldt et al. (2015), so I only sketch the key steps here.

Consider Equation (4) in the text

$$L_d^e = \sum_{o\in \mathbf{G}} \pi_{o,d}^e l_o^e,$$

Where L_d^e and l_o^e are data, and $\pi_{o,d}^e = \frac{(\frac{v_d^e}{d_{o,d}})^{\epsilon^e}}{\sum_{g\in\mathbf{G}}(\frac{v_g^e}{d_{o,g}})^{\epsilon^e}}$. Given $\{d_{o,d}\}$, l_o^e, and L_d^e, the only unknowns in this equation is $\{v_d^e\}$. Let $\mathbf{v}^e$ be the vector $(v_1^e, v_2^e, ..., v_d^e, ...)$. Define $\text{WD}(\mathbf{v}^e)$ (worker deficits) as

$$\text{WD}(\mathbf{v}^e) = L_d^e - \sum_{o\in \mathbf{G}} \pi_{o,d}^e l_o^e.$$

WD is simply the gap between the number of workers working in region d in the data, and the number predicted by the model. $\text{WD}(\mathbf{v}^e)$ is a function of $\mathbf{v}^e$. To prove Proposition 1 we show the following:

1. $\text{WD}(\mathbf{v}^e)$ is continuous;

2. $\text{WD}(\mathbf{v}^e)$ is homogeneous of degree zero;

3. $\sum_{d\in\mathbf{G}} \text{WD}_d(\mathbf{v}^e) = 0, \forall \mathbf{v}^e \in \mathbf{R}_+^G$

4. $\text{WD}(\mathbf{v}^e)$ exhibits gross substitute property.

It is easy to verify that requirement (1) and (2) are satisfied. Requirement (3) can be shown to be satisfied by noting that $\sum_{d \in G} \pi^e_{o,d} = 1$; requirement (4) can be shown to be satisfied by computing the derivatives directly.

Requirements (1)–(2) guarantee the existence of a solution. The proof is a constructive one: by homogeneous of degree zero, we can normalize $\mathbf{v}^e$ to the simplex $\{\mathbf{v}^e \in R_+ : \sum \mathbf{v}^e = 1\}$. Define $\mathbf{WD}^+ = \max\{0, \mathbf{WD}\}$, and $\mathbf{f}(\mathbf{v}) = \frac{\mathbf{v}+\mathbf{WD}^+}{\sum_{d \in \mathbf{G}} v_d + \sum_{d \in textbf G} \mathbf{WD}(v)_d}$, then $\mathbf{f}$ is a continuous function mapping the unit simplex onto itself. The existence of a solution to $\mathbf{v} = \mathbf{f(v)}$ then follows from the Brouwer's fixed point theorem.

Requirement (3)-(4) then guarantee the uniqueness of the solution, see Ahlfeldt et al. (2015) for additional discussion. The implication of proposition 1 is that, given migration costs, we can solve Equation (4) for the unique set of amenity-adjusted real wages for all locations.

B Data

B.1 Overview of the Data Sources

The primary individual- and firm-level data I use are the following: random sub-samples of the 2005 Population Mini Census and the 2000 Population Census, manufacturing sub-sample of the 2004 Economic Census, and the universe of the 2004 Annual Survey of Industrial Production. In addition to these micro data sources, I also use the 2002 inter-regional and inter-sectoral input-output table, as well as data from the UN Comtrade Database (sectoral trade), national accounts (sectoral production), and provincial statistical yearbooks (city-level import, export, and surplus).

The 2005 Mini Population Census covers 1% of Chinese population. It records individual demographic and employment information. To my knowledge, this is the only dataset that provides individual-level income information for the entire country, so I use it to estimate the average income in each region. I also choose 2005 as the benchmark year, as the calibration procedure requires wage information.

The 2000 Population Census covers the entire Chinese population. Respondents in this sub-sample fill a longer form than others, which asks for information on migration, education, occupation, industry, and housing conditions, but unfortunately, not for income or earnings.

The 2004 Economic Census covers the universe of registered firms. The sample I have access to is its manufacturing sub-sample, with firm-level revenue and employment information.

The 2004 Annual Survey of Industrial Production covers all state-owned enterprises, as well as private enterprises with annual sales over 5 million RMB yuan. Different from the 2004 Economic Census, this dataset contains detailed firm-level financial information, rather than only employment and revenue information.

The rest of this section covers necessary details in sample selection and variable construction.

B.2 Cultural Distance

To proxy for the cultural distance between cities, I construct a cultural similarity index based on the compositions of ethnic minority groups. I extract the prefecture-level information on the compositions of ethnic minorities from the 1990 census. Migration was not as pervasive in 1990 as it was in 2000, and therefore the ethnic compositions largely reflect the cultural root of a city. Using the 1990 census data helps us avoid the endogeneity problem that would arise, if we used the 2000 census to construct cultural distance.

There are 56 ethnic groups in China, with Han ethnic being the dominating one. I exclude Han from the calculation because its share of population is so high that including it eliminates most of the variation in the similarity index. For each city, I am left with a 55 by 1 vector, each element of which is the share of one ethnic group in the total local ethnic minority population. I then compute the correlations between the vectors of all city pairs, and use these as the values of my cultural similarity index; the cultural distance is then defined as one minus this similarity index.

Figure (B.1) is the density distribution of the index. The mean, median and standard deviation of the similarity index are 0.2569, 0.0608, and 0.3645, respectively.

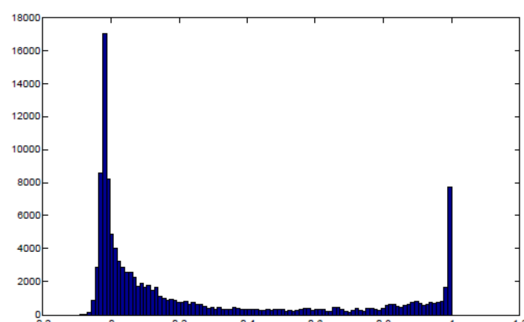


Figure B.1: Density Distribution of the Similarity Index
Source: Author's calculation based on the 1990 census

B.3 City-Level International Trade Surplus

To incorporate international trade imbalances into the calibration, I construct a dataset of city-level international trade surplus.

Each city's trade surplus in 2005 is extracted directly from the provincial statistical yearbooks. I make two adjustments to the data. First, Beijing trades a lot with the ROW, but the majority of the trade is done by big companies (especially the SOEs) with headquarters in Beijing. It is plausible that the trade is actually carried in the subsidiaries of these companies, spread out over the country. Fortunately, Beijing statistical yearbook reports "local trade" and "total trade" separately, the later including trade done by SOEs. I assign "local trade" to Beijing, and the remaining component of "total trade" to all Chinese cities, based on their relative size. The implicit assumption is that the operation of those SOEs headquartered in Beijing are distributed across all cities, proportionally to their size.

Second, sometimes the data are not well-behaved. For example, for Shaoshan, a city in Guangdong Province, one of the coastal provinces, the trade surplus is 13 times of its GDP. My conjecture is that there are many trade intermediaries. I make the following adjustments: I aggregate city-level trade surplus to the province level, and then allocate the trade surplus of a province to its cities according to their GDP. The underlying assumptions are that those trade intermediaries mostly work with other companies in the same province and that trade surplus is proportional to size of economy within a province.

I convert the city-level trade surpluses from the data to the scale of the model and divide allocate them to sectors.

B.4 Input-Output Linkages for China and the ROW

Table B.1: Input Shares in China and the ROW

$\gamma_s^{s'}$	Output Industry: China			
Input	A	M	K	S
L	0.57	0.30	0.59	0.48
A	0.19	0.07	0.00	0.03
M	0.15	0.44	0.26	0.21
S	0.09	0.20	0.16	0.28

$\gamma_s^{s'}$	Output Industry: ROW			
Input	A	M	K	S
L	0.58	0.42	0.56	0.63
A	0.19	0.00	0.00	0.00
M	0.16	0.41	0.26	0.11
S	0.07	0.17	0.18	0.26

In the model, the input-output parameters for China are constructed from the 2002 national input-output table, which records, at the 2-digit industry level, the usages of inputs in the economy. I aggregate the data to four industries—agricultural, capital and equipment, other manufacturing, and service, and four inputs—industry final outputs in the agricultural, other manufacturing, and service industries, as well as equipped composite labor.

The input shares of the ROW are assume to be the same as the median country in Parro (2013). Since industry classification in my paper is finer, for values not directly available in Parro (2013), I use the corresponding value from China, scaled appropriately. The underlying assumption behind this imputation that input-output linkages are similar across different countries are strongly supported by Jones (2013). All results in the paper are robust to changes in the input shares. Table (B.1) report the shares of inputs in each industry. The upper panel is for China and the lower panel is for the RoW. A, M, K, S represent agricultural, manufacturing, capital and equipment, and service industries, respectively. L represents equipped composite labor.

B.5 Skills, Wages, and Skill Premia

Since only a small proportion of workers had college degrees in China in the early 2000's, I classify a worker to be high-skill, if he or she has received more than nine years' formal education, equivalent to finishing junior high school.[2] Below I explain how I estimate average wage for each region by skill cell.

There are two types of workers, two types of local labor markets (rural and urban), and N cities in the economy, so in total there are 4N wages (mean wages for skilled and unskilled workers in all regions in the economy) to estimate. The data I use for this purpose is the 2005 mini census.

I estimate the following specification:

$$\begin{aligned} log(\text{Wage}_{e,i}) =& \beta_0 + \beta_1 \text{age} + \beta_2 \text{age}^2 + \beta_3 \text{sex} + \beta_4 I_{\text{Skilled}} * I_{\text{Agriculture}} \\ & + \text{F}_i + \text{S}_i * \text{F}_i I_{\text{Skilled}} + \text{A}_i * \text{F}_i I_{\text{Agriculture}} + \epsilon_{e,i}, \end{aligned}$$

where F_i is the regional fixed effect, $\text{F}_i * I_{\text{Skilled}}$ is the interaction between regional fixed effect and high-skill dummy, and $\text{F}_i * I_{\text{Agriculture}}$ is the interaction between regional fixed effect and a dummy for agricultural sector.

In this specification, I restrict the skill premium in the agricultural sector (relative to the skill premium in urban sector of the same city) to be the same across cities (that is, β_4

[2] The higher education reform started in 1999 in China, which expanded the scale of the higher education sector dramatically. Before the reform, the college admission rate in China was below 5%; in 1999, the college admission increased by 40%. The following years saw additional increase. But until 2005, college graduates constitute only a small proportion of the Chinese labor market.

is not city-specific). This choice is constrained by the power of the regression, as in the sample, the rural sector in many cities only hires a small number of high-skill workers. The omitted group in the regression is the unskilled worker in the urban sector in Beijing, whose average wage is β_0. Table B.2 illustrates how to recover regional wages from the regression.

The output of the regressions are presented in Table (B.3). The signs and magnitudes of coefficients are reasonable. The R^2 of the regression is 0.58, indicating that the regression has a strong explanatory power. Figure (B.2) presents the distribution of the p-values for the fixed effects in the wage regression. The distribution is heavily concentrated around zero (the spike in the figures corresponds to p-value<0.0005), suggesting that the fixed effects are very precisely estimated. Figure (B.3) shows the distribution of average wage for different worker groups across regions. Two patterns emerge: first, there is considerable heterogeneity across regions; second, overall, wages are higher for high-skill workers and urban workers. Figures 1c in the text cast the estimates for average wages of workers on the map of China.

Table B.2: Average wage for different groups

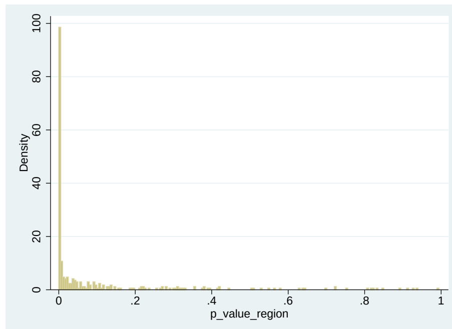
Education	Sector	Region	Wage
Unskilled	Urban	i	$\beta_0 + F_i$
Unskilled	Rural	i	$\beta_0 + F_i + A_i$
Skilled	Urban	i	$\beta_0 + F_i + S_i$
Skilled	Rural	i	$\beta_0 + \beta_4 + F_i + S_i + A_i$

Table B.3: Wage Regressions

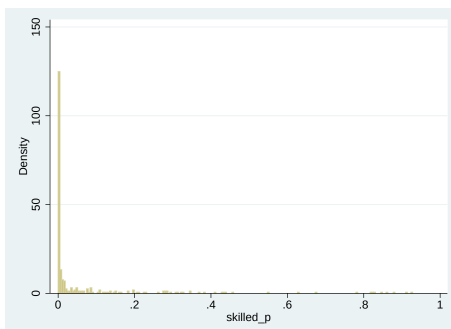
	(1) log_wage
Age	0.0327***
	(22.32)
Age_square	-0.000413***
	(-22.70)
Sex	-0.206***
	(-42.36)
Skilled_agri	-0.296***
	(-16.57)

t statistics in parentheses

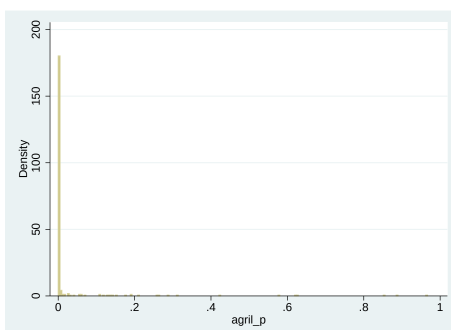
* p<0.10, ** p<0.05, *** p<0.01



(a) Regional Fixed Effects

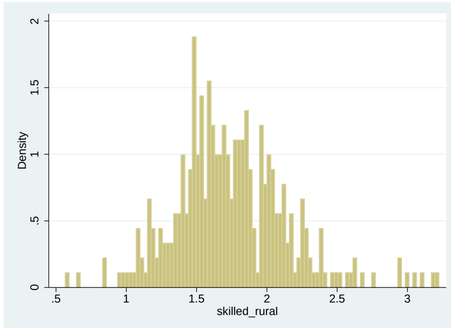


(b) Skilled*Regional Fixed Effects

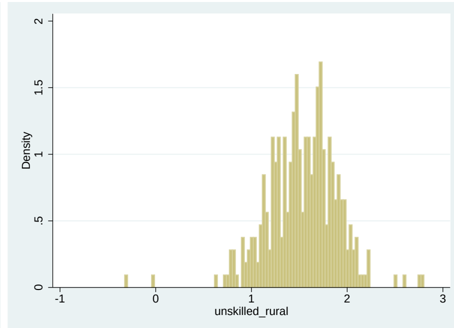


(c) Agriculture*Regional Fixed Effects

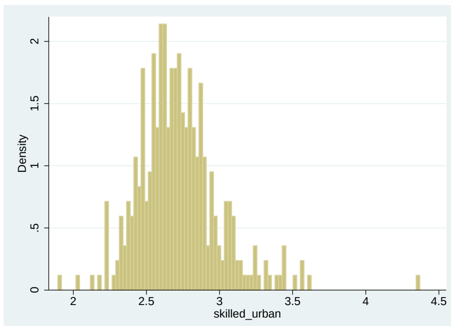
Figure B.2: Distribution of the P-value for Fixed Effects



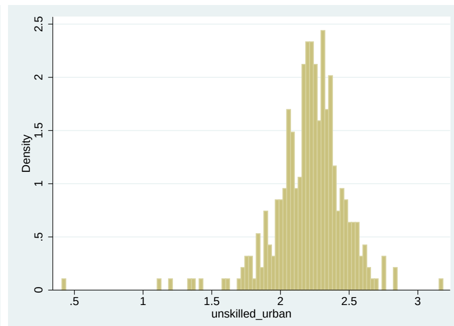
(a) Rural high-skill workers



(b) Rural low-skill workers



(c) Urban high-skill workers



(d) Urban low-skill workers

Figure B.3: Average Wages for Different Worker Groups

B.6 Migration

I use the 2000 population census to construct migration flows, following the procedures below: first of all, I restrict the sample to those who already finished their schooling, aged between 20 and 60 (60 is the official retirement age for urban male non-physical-labor workers in China). I drop those who are currently not working, unless the reason for not working is either "on vacation" or "on sick leave". I classify a worker as a migrant, if he or she is not residing in her or his birthplace. I identify the source sector (rural or urban) of a worker with the type of Hukou (rural or urban) the worker currently holds, and the destination sector of a worker by the locality at the survey.[3] From these procedures, for all workers in the economy, I identify their education level, source province, source sector, destination city, and destination sector. I use this to estimate inter-regional and inter-sectoral migration costs.

In the following, I discuss why measuring migration this way is most suitable for my analysis. Since I use the 2005 mini census to estimate regional wage and calibrate the model to the 2005 economy, ideally I would like to use this dataset to estimate migration

[3]To the extent that some rural Hukou holders have switched an urban Hukou in 2000, this classification underestimate rural-urban migration. However, before early 2000, switching a rural Hukou for an urban one was highly restricted. See also for discussion in the supplementary note on Hukou reforms.

costs, too. However, because the model neglects dynamic choices of individuals, the migration decision in the model is best interpreted as a life-time choice. So the model-consistent definition of migration is one that is based on birthplace. Unfortunately, the 2005 data does not report birthplace information, so I use the 2000 census to estimate the long-run migration costs.

Of course, the migration costs estimated this way apply only to 2000. Between 2000 and 2005 were partial Hukou reforms in some cities, which might reduce the cost of moving into these cities. I use the reduced-form estimate of the effect of Hukou reforms on migration to pin down the corresponding change in migration costs. In calibration, the migration costs I will use take into account the changes between 2000 and 2005 due to Hukou reforms.

An alternative migration measure is based on workers' Hukou and current residence from the 2005 mini census. We can define a worker as a migrant in Beijing, if he works in Beijing, but holds Hukou from elsewhere. One potential issue with this problem is that despite the strict control, many people were still able to obtain a local Hukou. This measure therefore will under-estimate the true extent of migration in China. To the extent that it was easier for more skilled worker to obtain local Hukou, this measure understates migration for skilled workers. Perhaps much more problematically, according to this measure, cities with more open Hukou policies might have fewer migrants simply because a larger fraction of migrants have been granted local Hukou.

For these reasons, I think that defining migration based on birthplace and current residence is most suitable for the purpose of this paper.

B.7 Worker Employment and Birthplace Distributions in 2005

After estimating the parameters governing migration costs, I solve the labor market clearing conditions (Equation 4 in the text) for one more time, to obtain the regional fixed effects—$\{v_d^e\}$—that are consistent with employment distribution in 2005. For this purpose, I need the distribution of workers' birthplace and employment in 2005 by skill.

I construct workers' birthplace distribution from the 2000 census. I restrict the sample to workers aged 15–55 in 2000. The birthplace distribution of this sample will be the birthplace distribution for workers aged 20–60 in 2005. To determine the skill level of workers for this sample, if a worker has finished schooling in 2000, I classify his or her skill level based on the education attainment directly; for workers that are above 15, but have not yet finished schooling, I assume they are skilled—by this age, a typical Chinese kid has received 8-9 years of education, so the possibility of (wrongly) classifying a student receiving less than 9 years education as skilled is minimized.

I construct the employment distribution from the 2005 mini census. For a few cities, because the sample size is small and the share of skilled workers are small, there are no skilled workers sampled. This creates a problem because the equilibrium does not allow for zero employment in any places—employment could be arbitrarily small, but not zero. To resolve this issue, If one region have 30 unskilled workers, and 0 skilled workers in my sample, I replace this zero with 30 * $\left(\frac{\text{Total Number of Skilled Workers}}{\text{Total Number of Unskilled Workers}}\right)$, where skill ratio $\left(\frac{\text{Total Number of Skilled Workers}}{\text{Total Number of Unskilled Workers}}\right)$ is computed based on the official publication calculated from the universe of mini census records for *that* city.

B.8 Factor Shares in Equipped Composite Labor

We need the shares of payments to capital, high-skill workers, and low-skill workers in each region to calibrate the region-specific equipped composite labor production functions. I compute the ratios between payment to high-skill workers and payment to low-skill workers directly from the estimated wages and the distribution of effective labor units, both of which are known once we have estimated migration costs. I further need the ratio between the payment to capital and the payment to labor in each region.

For the urban regions, I use the 2004 Survey of Industrial Production. I aggregate firm-level data to obtain the city-level ratio between wage bill and expenditures on capital and equipment. The firm-level wage bill is the “total salary payments" entry in the dataset; the firm-level expenditures on capital and equipment is the “total capital depreciations" entry in the dataset. In addition to depreciations to capital and equipment, the total depreciations entry also includes depreciations to properties and buildings. Therefore I adjust for this by subtracting the share of buildings among aggregate tangible fixed capital stock in China in 2004, calculated from the national statistical yearbook. The mean ratio across cities, constructed this way, is similar to the corresponding ratio from the national input-output table for the urban sector.

For the rural regions, since I am not aware of any data sources that contain information on capital share at the regional level, I assume the capital shares are the same for all rural regions and determine it using the national input-output table.

B.9 Inter-Regional Correlation of Worker Productivity ρ

Parameter ρ determines the correlation of workers' productivity draws across regions. The moment I use to pin down ρ is essentially the *residual* correlation in wages across regions for workers who have migrated. Using residual correlation is important because many factors outside the model, such as occupation, age, experience, and location, all determine a workers' wages and will likely drive the correlation in potential wages for a worker from different regions.

Specifically, I use an individual panel dataset from the 2004 and 2006 waves of China Nutrition and Health Survey (CHNS) and estimate a Mincer regression with regional fixed effects, time fixed effects, along with gender, education, age, and age squared as control variables. I restrict the regression to a sample of switchers—those who have either moved across cities or between the rural and urban regions of a city. For each individual, the two residuals (one for 2004 and one for 2006) therefore can be best thought of as two realizations of productivity draws for a worker in two regions.[4]

I further add individual fixed effects to this regression. I compare the R^2 of these two regressions and see how much of the variation unexplained in the first Mincer regression is explained by the individual fixed effects. Intuitively, if everything can be explained by the individual fixed effects, then the two draws must be the same, and ρ must be high. It turns out that around 66% of the unexplained variation in the first regression could be explained by individual fixed effects (if there were no correlation at all between two draws, this number should be 50%).

Given this estimate, for each given value of ρ, I simulate workers' productivity draws from different locations using a correlated Frechet distribution. Using this simulated draws, I estimate a regression specification with only individual fixed effects, and calculate the R^2. I chose the correlation parameter so that this R^2 is 66%. This procedure determines a value of 0.36 for ρ.

C Quantification

C.1 Numerical Algorithm

As discussed in Section 3.E in the text, I determine international trade costs, domestic trade costs, and regional productivity jointly. This section describes the numerical algorithm.

I start with an initial guess for international trade costs $\{t_A, t_M, t_K\}$, and the parameters governing domestic trade costs, $\{\gamma\}$, with which I compute the trade cost between any trade partners, $\{\tau_{o,d}\}$. I then guess a distribution for regional productivity, solve the trade model for prices and trade shares, and check if the demand for intermediate va-

[4] I thank the National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center, the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, and R01-HD38700) and the Fogarty International Center, NIH for financial support for the CHNS data collection and analysis files from 1989 to 2006, and for kindly making the data publicly available.

rieties produced by each region equals the supply (Equation 16 in the text).[5] If not, I update the guess for the distribution by increasing productivity in regions with excess supply and decreasing productivity in regions with excess demand.[6]

Once the distribution of regional productivity that clear all intermediate variety markets are found, I compute the bilateral trade flows, and evaluate the objective function, given by Equation (21) in the text.

I search over the space of $\{\gamma\}$ until the global minimum is reached, after which I calibrate international trade costs to match the sectoral openness, keeping both domestic trade costs and regional productivity fixed. I repeat the process until convergence.

C.1.1 Recovering Equipped Composite Labor Production Function η_d^h, η_d^l

As discussed in Footnote 5 of this Appendix, in solving the trade model, we need to compute the prices of tradable goods, taking regional wages and the distribution of technology $\{T_d^s\}$ as given. Computing the prices, however, requires η_d^h and η_d^l.

To proceed with the estimation algorithm not knowing η_d^h, η_d^l, I substitute the relative factor shares, $\frac{\text{Capital Share}}{\text{Skilled Share}}$ and $\frac{\text{Equipped Skilled Share}}{\text{Unskilled Share}}$ (data from Section B.8 of this Appendix), at the regional level, to the left hand side of Equation (12) in the text, and express η_d^h, η_d^l as

$$\eta_d^h = \frac{(\frac{P_d^K}{W_d^h})^{1-\rho_{kh}}}{\frac{\text{Capital Share}}{\text{Skilled Share}} + (\frac{P_d^K}{W_d^h})^{1-\rho_{kh}}}, \quad \eta_d^l = \frac{(\frac{W_d^{eh}}{W_d^l})^{1-\rho_{lkh}}}{\frac{\text{Equipped Skilled Share}}{\text{Unskilled Share}} + (\frac{W_d^{eh}}{W_d^l})^{1-\rho_{lkh}}} \tag{C.1}$$

I then substitute Equation (C.1) into (11), and solve the model without actually knowing η_d^h or η_d^l. Once the whole procedure is over and the model is solved, we can then back out η_d^h and η_d^l from (C.1). These are interpreted as the true parameter values, which I keep fixed for all counterfactual experiments.

[5] In the step where we solve the trade model, if we know η_d^h and η_d^s, Equations (7), (10), and (11) in the text can be viewed as a system of equations with prices being the only unknowns. Although η_d^h and η_d^s are unknown before the model is parameterized, in Section C.1.1 of this appendix I show that, conditional on information on the shares of different factors in the equipped composite labor, η_d^h and η_d^s are unnecessary in solving the model. Once the model is solved, however, we can use Equation (12) to back out η_d^h and η_d^s, to be used in policy experiments.

[6] The feasibility of this approach requires that, for any given level of trade costs, we can find a set of unique T_d^s that clear all intermediate variety markets in all locations. Redding (2016) proves this is true in a single-sector model. An earlier version of this paper extends the proof to a multi-sector model with input-output linkages within the same broad sector. In the general model here with flexible input-output linkages and capital-skill complementarity, the uniqueness cannot be established. But in implementation, I find the update rule always converge uniformly to one unique object.

C.2 Model Validation

C.2.1 The Gradient of Changes in Skill Premia

I provide additional information about the external validation exercises reported in Section 3.G of the text.

Table C.4: Trade Liberalization and Skill Premia

	(1)	(2)
	Skill Premium Change (Model)	
Distance to Port	-0.016***	
	(0.002)	
Coastal Province		0.043***
		(0.007)
Constant	0.157***	0.049***
	(0.011)	(0.005)
Observations	78	78
R^2	0.426	0.349

Notes: This table reports regressions of changes in skill premia on city-level distance to port and a dummy for coastal province using model-simulated data. Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Using two recent trade liberalization episodes in China (Deng Xiaoping's Southern Tour in 1992 and the WTO accession in 2001), Han, Liu and Zhang (2012) separately estimates the effects of trade liberalization on high school and college premia. They find that after reforms (Column 1, Table 2 of their paper), the college premium increases by 14.2% (0.067+0.075) more in coastal provinces than in interior provinces ; the high school premium increases by 4.27% (0.054-0.0113) more in coastal provinces than in interior provinces. In my model, both types are considered skilled, so assuming that the fraction of high school graduates with a college degree is 5%, the corresponding change in skill premium is $(14.2\% + 4.27\%) * 0.05 + 4.27\% * 0.95 = 5\%$. Using distance as an alternative measure for trade exposure, they find that after reform (in their Appendix Table 4), college premium increases by 4.4% (0.018+0.026) more in cities that are 100% geographically closer to the coast; high school premium increases by 1.7% (0.008+0.009) more in cities that are geographically closer to the coast. The weighted average across these two types is therefore $(4.4\% + 1.7\%) * 0.05 + 1.7\% * 0.95 = 1.9\%$.

Table C.4 reports the model counterpart of these results. I regress percentage changes in skill premia from autarky to the calibrated equilibrium on distance to port and a coastal province dummy, restricting the samples to the same set of provinces used in Han, Liu and Zhang (2012). Column 1 shows that the coefficient associated with distance is 1.6%.

Column two shows that the coefficient associated with the coastal province dummy is 4.3%. Both numbers are close to the data (1.9% and 5%, respectively).

C.2.2 The Gradient of Population Growth Rates

I also compare the model's prediction on city growth to the empirical counterpart. The model predicts that, after trade liberalization, workers will reallocate from the interior to the coast. This reallocation implies a gradient in city growth: coastal cities will experience faster population growth during a period of rapid trade integration.

Table C.5: Trade Liberalization and Population Reallocation

	(1)	(2)	(3)	(4)
	Population Growth Rate (2000-2010)			
	Data		Model	
Distance to Port	-0.034*	-0.079***	-0.048***	-0.111***
	(0.007)	(0.021)	(0.004)	(0.009)
Constant	0.246***	0.621***	0.259***	0.416***
	(0.041)	(0.102)	(0.023)	(0.054)
Prov FE		Y		Y
Observations	339	339	340	340
R^2	0.058	0.280	0.305	0.573

Notes: Columns 1-2 report results from regressions of city-level population growth rates between 2000-2010 on distances to port. Columns 3-4 report results from the same regression using simulated data. Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I verify whether this prediction holds up in the data, and if it does, whether the magnitude is similar to that implied by the model. Table C.5 reports regression results using real and simulated data. The first two columns are based on the real data. The dependent variable is log population growth between 2000 and 2010, a period of rapid trade integration in China the WTO accession. The key independent variable is a city's log distance to port. The first column does not include provincial fixed effects. The coefficient is -0.034 and statistically significant at 10%. The second column includes provincial fixed effects, so the identification comes from within province variation. The coefficient increases in magnitude to -0.079.

The third and fourth columns use the model-simulated data. The dependent variable is population growth rate from autarky to the calibrated equilibrium. The independent variable is the same as in Columns 1 and 2. Without provincial fixed effects, the coefficient of interest is -0.048. When provincial fixed effects are included, it becomes -0.111. Both coefficients are statistically significant. The magnitudes are also in line with the empirical counterparts.

C.3 Reallocation After Trade and the Role of Domestic Trade Costs

In this subsection, I first provide direct evidence on the reallocation mechanisms driving spatial distribution effects in Section 4.A.1 of the text. I then show that without domestic trade costs, none of the results relating to the geographic dimension would exist, so the spatial distributional impacts of trade in this model are indeed driven by internal geography of a country, rather than regional heterogeneity in production (including heterogeneous productivity T_d^s and factor shares η_d^e).

The main mechanisms in the model are as follows. Because of the domestic trade costs, trade liberalization improves the wage of coastal regions by more than that of interior regions. This leads workers to leave the interior and to move to the coast. Because skilled workers are estimated to have lower migration costs, more skilled workers than unskilled workers migrate, which increases the share of skilled workers in coastal regions. This, however, does not reduce skill premia in coastal regions for two reasons. First, because capital and equipment goods are complementary to skilled workers in production. In the calibrated economy, the RoW has a comparative advantage in capital and equipments, so after trade liberalization, the price of these goods decrease in China. The price decreases are especially large in coastal regions, which trade more intensively with the RoW because of the geographic proximity. Second, because manufacturing and capital industries are calibrated to be more intensive in the use of intermediate goods, regions with better access have a comparative advantage in these industries. Trade liberalization increases the access to foreign market by more in coastal regions. This helps coastal regions specialize in manufacturing and capital industries, which are more skilled intensive.

I provide direct evidence of the reallocation patterns predicted by these channels in Figure C.4. Panel (a) plot the changes in population, share of skilled workers, and share or urban industries for the benchmark economy. As the figures show, coastal cities expand after trade liberalization, while cities of the coast shrink. Because skilled workers are more mobile, the share of skilled workers increases in the coast. Finally, as predicted by the "Domestic Stolper-Samuelson Effect", the share of urban value added in local economies increases on the coast and decreases in the hinterland.

The model features rich regional heterogeneity in production. To rule out that these patterns are due to heterogeneous production functions across regions, Panel (b) of Figure C.4 plots the changes in a re-calibrated economy with the same regional heterogeneity but frictionless domestic trade. As is clear from Panel (b), when we shut down domestic trade costs, non of the reallocation patterns exist. Figure C.5 further shows that, in this alternative model with frictionless domestic trade, the distributional impacts of trade along the geographic dimension also disappear.

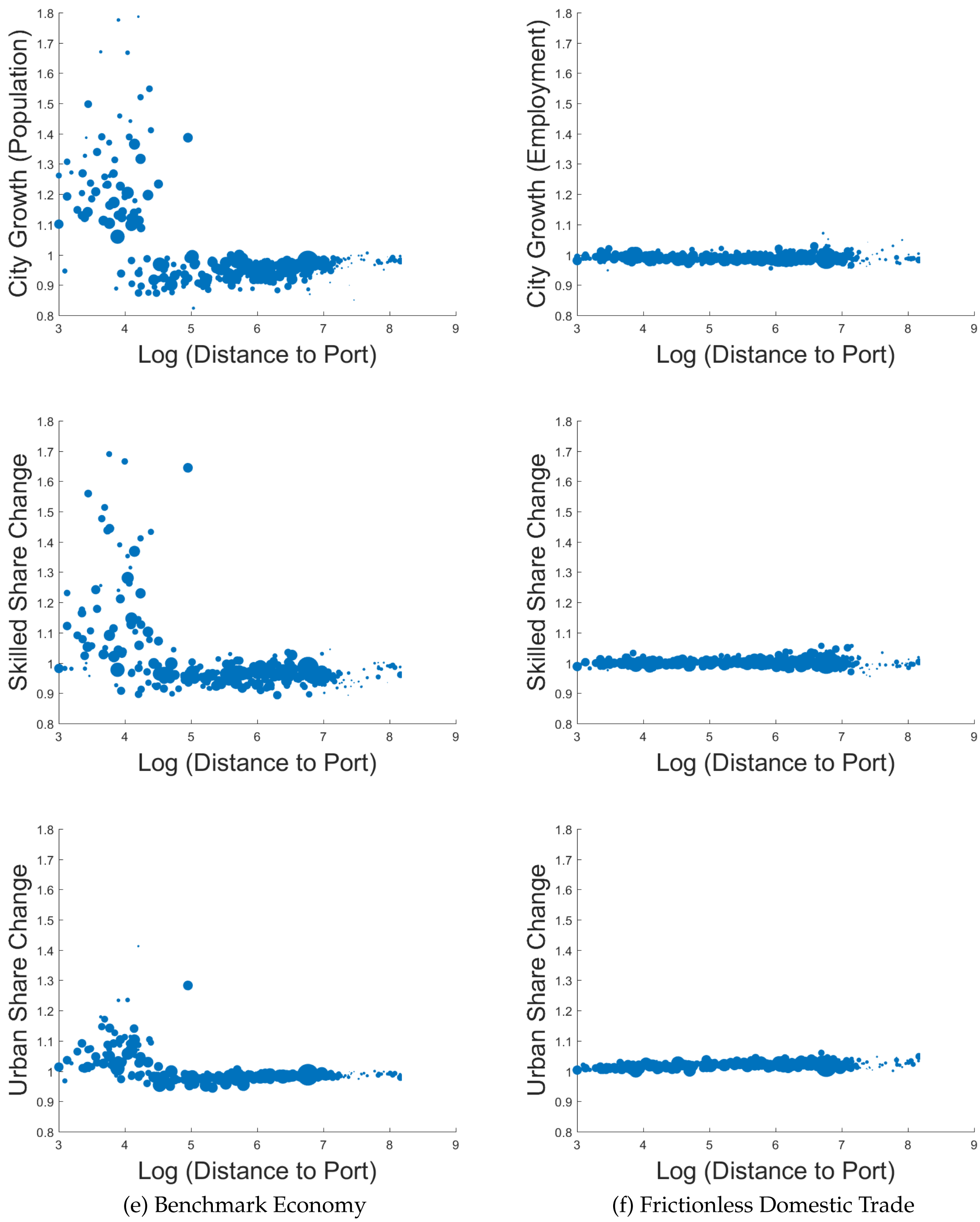
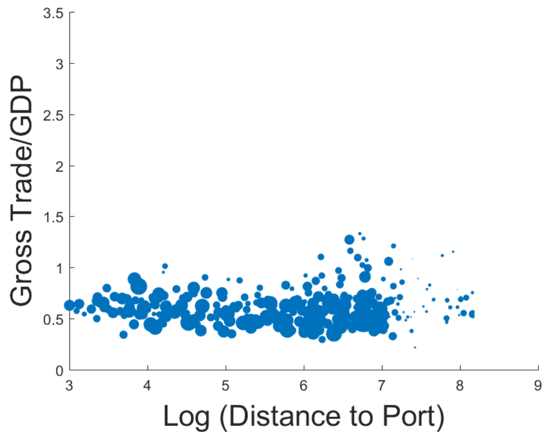
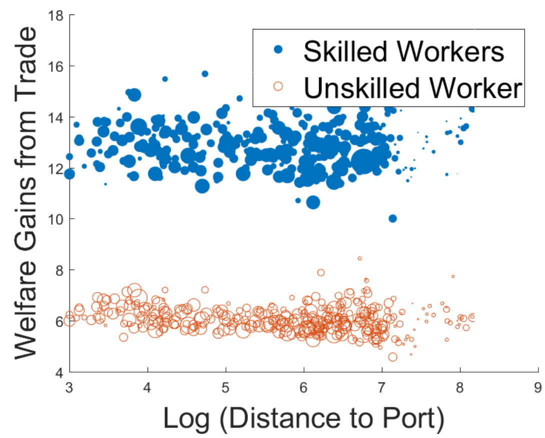


Figure C.4: Trade-induced Reallocations

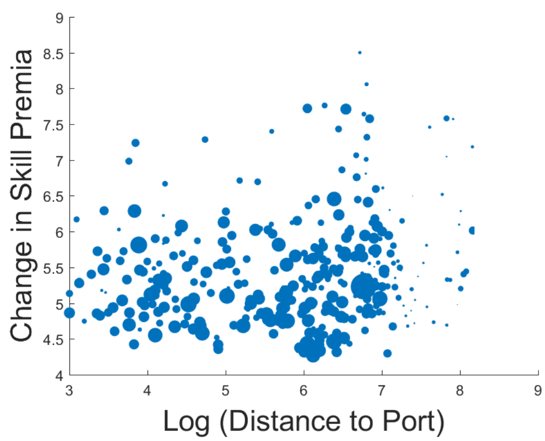
Notes: The vertical exercises are the ratio between open-economy value and the autarky value of each variable.



(a) Trade Intensity



(b) The Gains from Trade



(c) Skill Premia Change

Figure C.5: The Distribution of the Gains from Trade with Frictionless Domestic Trade

Notes: The welfare gains and skill premium changes are in percent.

C.4 Domestic Trade Reforms and International Trade Participation

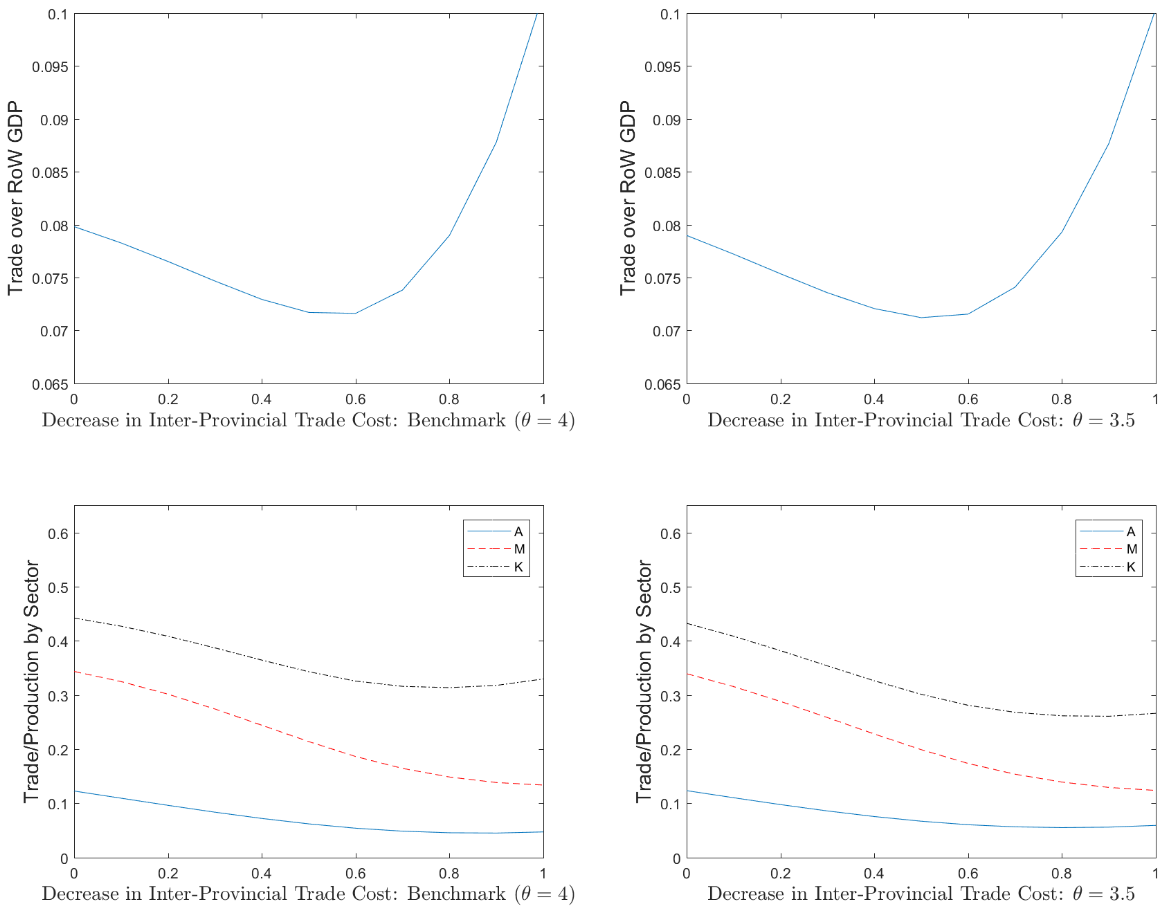
In Section 4.B of the text, I show that decreasing domestic trade costs in the benchmark economy might reduce the country's participation in international trade. This happens for two reasons. First, with lower domestic trade costs, coastal regions will trade more intensively with interior regions and less intensively with the RoW. This trade diversion effect might reduce the *volume* of international trade.[7] Second, in the model, China is not a small open economy. It faces an upward-sloping supply curve and a downward-sloping demand curve from the RoW. As domestic reforms effectively increase the size and productivity of China, imports from the RoW become relatively more expensive (the terms-of-trade deteriorate for China), so *trade intensity* might decrease. A natural question is in what parameter regions this could arise. I show that this result is impervious to different depths of domestic reforms and alternative values of trade elasticities.

Figure C.6 plots international trade between China and the RoW, as domestic trade costs in China are gradually lowered. The left panel is for the benchmark economy; the right panel is for an economy with a lower trade elasticity (θ=3.5), re-calibrated to match the same production distribution and international trade in 2005. The horizontal axis is the decrease in inter-provincial and inter-regional dummies from the estimated value of 1.21 and 1.51, respectively (Table 4). For example, a value of 0.5 in the horizontal axis indicates these two dummies decrease by 0.5.

In the two upper panels, the vertical axis is the trade volume between China and the RoW as a share of RoW GDP. Because reforms within China only have a very modest effect on the RoW ($\frac{\text{GDP}_{\text{China}}}{\text{GDP}_{\text{RoW}}} = 0.126$ in calibration; the real income barely changes in the RoW after reforms), the RoW GDP can be thought of as a numeraire and the ratio could be interpreted as trade *volume* between China and the RoW. The figure shows that initially, the trade diversion effect dominates and trade volume decreases as domestic trade costs are gradually lowered. However, as interior regions trade more and more intensively with the RoW, eventually this force dominates. After some point, trade volume starts increasing as domestic trade costs decrease further. The upper right figure shows that the patterns are similar if trade elasticity is 3.5.

The lower figures plot trade intensities for three tradable sectors, defined as sectoral international trade over sectoral production in China. The trade intensity decreases in all three sectors as we just start to reduce domestic trade costs. This is a combined result from two effects. First, as discussed earlier, trade volume between China and the RoW decreases. Second, domestic reforms increases the size of the Chinese economy, which

[7] Of course, interior regions will trade more intensively with both the RoW and coastal regions. Whether the net effect increases or decreases trade volume depends on the relative strength of these two forces.



(c) Benchmark Economy

(d) Lower Trade Elasticity

Figure C.6: Trade Volume, Trade Intensity, and Domestic Trade Costs deteriorates China's terms of trade, leading to further decline in the trade intensity. As domestic trade costs further decrease, eventually trade intensities stop declining and start gradually increasing. The point at which this minimum is reached vary across sectors, but because of the terms-of-trade effect, these points are all located to the right of the minimum *trade volume* point shown in the upper panel. The lower right panel shows that the same pattern holds when θ is lower.

C.5 Domestic Reforms and Further Liberalization

Section 4.B of the text shows that freer domestic factor and goods market increase the effectiveness of future trade liberalizations. I discuss the underlying channels.

First, I show that domestic trade and migration costs will not have a first-order impact on the economy's response to trade liberalization. With the gravity structure in my model, the total import of an interior region, indexed by *int*, from the RoW is given by (for clarity sectoral index s suppressed):

$$\begin{aligned}\log(X_{\text{int,RoW}}) &= \log(X_{int}\delta_{\text{int,RoW}}) \\ &= \log(X_{\text{int}}) - \log(\sum_{o'} T_{o'}(c_{o'}\tau_{\text{int,o}'})^{-\theta}) + \log(T_{\text{RoW}}) - \theta\log(c_{\text{RoW}}) - \theta\log(\tau_{\text{int,RoW}}),\end{aligned} \tag{C.2}$$

in which $\delta_{d,o}$ is the trade share, defined in Equation 9, and X_{int} is the total expenditure in region *int* on intermediate varieties.

Recall that $\log(\tau_{\text{int,RoW}}) = \log(\tau_{\text{int,Port}}) + \log(t)$, where $\log(\tau_{\text{int,Port}})$ is the trade cost between the interior region and the nearest port, and t is the international trade cost. Now consider the elasticity of trade with respect to international trade cost, $\frac{\partial\log(X_{\text{int,RoW}})}{\partial\log(t)}$. Given that there are many regions, a change in the trade cost between China and the RoW will only have a small impact on the multilateral resistance term $\log(\sum_{o'} T_{o'}(c_{o'}\tau_{\text{int,o}'})^{-\theta})$. The first order effect of international trade cost on the trade flow between an interior region and the RoW is therefore given by θ. The conclusion also holds true for port cities, which can access the international market directly. Therefore domestic trade and migration costs do not have a first order effect on the international trade elasticity in the model.

Nevertheless, Figure 5 in the text shows that post-reform economies respond more strongly to future tariff cuts. In the case of domestic Hukou reforms, trade elasticity is higher because of the reallocation channel. Specifically, even though a given decrease in t increases international trade in coastal and interior regions by the same *percentage*, because coastal regions originally traded more intensively with the RoW, the increase in export as a share of local production is higher in coastal regions, which leads to a larger percentage increase in local labor demand and wage.[8] This draws workers from interior to coastal regions.This reallocation increases the size of coastal regions, which trades more intensively with the RoW. By making this reallocation easier, Hukou reforms increases the economy's response to international trade liberalization.

The explanation for higher trade elasticity after the domestic trade reform is similar. With high domestic trade cost, the trade volume increase comes primarily from coastal regions. This drives up coastal wages and reduces export in these regions through the GE effect (higher production cost). When domestic trade costs are lower, trade volume increase will be more evenly spread across regions. The GE effect is therefore attenuated.

C.6 Sensitivity Analysis

C.6.1 Robustness to Different External Parameters

The parameterization uses several externally calibrated parameters, which might affect the predictions of the model. I show that the main conclusion of this paper are insensitive to alternative values of these parameters. In all cases, I recalibrate the model to match the 2005 data.

The elasticity of substitution between capital and skilled workers, ρ_{kh}, and the elasticity between equipped skilled workers and unskilled workers, ρ_{lkh}, are important parameters in the model. For robustness, I first reduce ρ_{lkh}, to 1.1, implying that the upper nest is close to the Cobb-Douglas production function. I then increase ρ_{kh} to 1.1, keeping ρ_{lkh} at the benchmark level. Finally, I treat capital, skilled workers and unskilled workers as symmetric input into composite labor production by setting both ρ_{lkh} and ρ_{kh}, to 1.1. Rows 2–4 in Table (C.6) report the findings. As we can see, the aggregate gains from trade remain similar, while the changes in aggregate inequality and the contribution from the within-region component become smaller, as expected.

In the benchmark analysis, I calibrate the correlation between an individual's productivity draws across regions, ρ, to 0.36, to match the correlation in residual wages of migrants before and after migration. I perform the policy experiment again, for $\rho = 0.3$ and $\rho = 0.4$. Rows 5 and 6 in Table (C.6) report the findings. The results do not change much.

I also change the elasticity of trade, θ, from 4 to 3.5 and 4.5, and conduct the same exercise. Rows 7 and 8 in Table (C.6) reports the results. When trade is more elastic, the gains from trade tend to be smaller and so are the inequality increase. But overall, the

[8] Similarly, while the price of imported goods decrease by the same percentage point everywhere, because coastal regions initially import from the RoW more intensively, they see a larger decrease in the local price index.

results are similar to the benchmark model.

Finally, I introduce remittances into the model, so that the impacts of trade could be spread across space through migrant's remittances. Following Akay et al. (2012), I assume migrants send 10% of their income home. The last row of the table reports the results. The predictions are very similar to those of the benchmark model.

Table C.6: Sensitivity Analysis

Parameters	Average Gains	Inequality Increase	Contribution (%)	
			Between	Within
Benchmark	7.5	6.7	74.7	25.3
$\rho_{kh} = 0.67, \rho_{lkh} = 1.1$	7.3	6.0	75.1	24.9
$\rho_{kh} = 1.1, \rho_{lkh} = 1.67$	7.2	5.1	80.1	19.9
$\rho_{kh} = 1.1, \rho_{lkh} = 1.1$	7.5	4.3	86.5	13.5
$\rho = 0.3$	7.1	6.3	73.0	27.0
$\rho = 0.4$	7.1	6.3	72.8	27.2
$\theta = 3.5$	8.1	7.7	76.5	23.5
$\theta = 4.5$	6.3	4.7	65.4	35.6
10% remittances	7.2	6.4	72.6	27.4

Notes: This table reports the effects of trade on welfare and inequality under alternative parameterizations. All numbers are in percentage points. Measures for average gains from trade and inequality are the same as in Table 6.

C.6.2 Discussion on the Estimated Inter-Provincial Trade Cost and Additional Exercise

In Table 4 of the text, I report my estimates of the domestic trade costs. It is useful to compare my estimates to those obtained using the U.S. Commodity Flow Survey data. In the literature, the comparable coefficient for state border, after scaled appropriately by the elasticity of trade, is on the range of 0.38 (Wolf, 2000) to 0.65 (Crafts and Klein, 2014, using 2007 data). My estimate of the state-border effect is therefore about twice as large as the comparable estimates for the U.S., reflecting larger barriers to trade flows at provincial borders in China. One lesson from the U.S. state border effect literature is that, the estimates might be driven up by the wholesale industry (Hillberry and Hummels, 2003), and might suffer from the aggregation bias—a lot of trade costs are actually due to geographic distance, but might be captured by the state-border dummy when state-level aggregate data is used. When these two factors are taken into account, the estimates shrink (Hillberry and Hummels, 2008).

Therefore, as discussed in Footnote 24 of the text, one natural concern is whether due to the poor quality or the high level of aggregation of the data in China, the estimates might also overestimate the effect of provincial borders; and if that is the case, whether

Table C.7: Counterfactual Experiment with an Alternative Domestic Trade Cost Structure

Panel A: Statistics by Skill		
	Mean	std
Skilled Workers	12.55	10.45
Unskilled Workers	5.40	8.30

Panel B: Aggregate Statistics	
National Average	6.93
Skill Premium Increase	5.6
Inequality Increase	7.48
Contribution-Between(%)	78
Contribution-Within (%)	22

the results from the counterfactual experiments are still valid. My use of city-level international imports and exports should partially alleviate this concern. Intuitively, this information allows me to identify domestic trade costs using variation in distance to port within a province. The estimated provincial border effects are also large when I use only this as the objective function.

To further address the concern, I use an additional experiment to show that even if there are biases in the estimation, they will not affect main conclusions of the counterfactual experiment. Specifically, I perform a robustness exercise, in which I reduce inter-provincial and inter-regional trade costs to 0.65, the level of the U.S. economy, while recalibrate the model to match regional production and sectoral trade. I compute the gains from trade and inequality changes in that economy. The results are reported in Table (C.7). Overall the findings are similar. The between-region component accounts for the most of the increase in inequality.

C.7 Discussion of Model Assumptions

I make several assumptions in the model. In this subsection, I discuss how the violations of these assumptions would affect the main results.

In terms of the timing of migration, I assume that workers learn their idiosyncratic productivity draws in all regions prior to their move. Admittedly, in reality, there is substantial uncertainty about the payoffs to migration, which can be inferred from the fact that many migrants return to their birthplace shortly after their migration (Kennan and Walker, 2011). In the empirical analysis, I classify workers as migrants if they are currently not in their birthplaces. Some of them might be temporary migrants who will shortly return to their hometowns. However, these migrants are unlikely to constitute an important part of the total migrants: even if 50% of migrants are temporary workers who return to their hometowns within two months, over a period of twenty years, the stock of migrants in each place will mostly be the permanent ones. My estimates of the migration costs, then, correspond to the long-run migration costs, which could be interpreted as reduced-form approximations of the real migration costs when there is uncertainty.

I use the Frechet distribution to model individuals' productivity draws. This distribution is a reasonable approximation of the wage distribution. In particular, it has a fat right tail. Most existing work in the migration literature makes similar parametric assumptions, using Logit or Pareto distributions. Instead of treating idiosyncratic migration decisions as outcomes of idiosyncratic individual preference shocks, an approach commonly adopted in migration literature, I assume they are driven by idiosyncratic productivity shocks as in Ahlfeldt et al. (2015). The advantage of this approach is that, while individual preference is unobservable, parameters governing productivity shocks can be inferred directly from the wage distribution. The Frechet distribution is particularly attractive because under this assumption, we have tractable expressions to aggregate supply of efficiency units in each region. That being said, all the channels discussed in this paper would apply under other distributional assumptions.

Since China's economy is growing quickly, and my model is static, one might worry that this discrepancy will make my results less useful. In analyzing the potential problems, it is important to be clear what dynamics one has in mind. First of all, the demographic structures are changing over time. My framework is general enough to incorporate multiple age groups, but I abstract from this mainly because of the limited sample size. Hence, my estimates could be interpreted as average migration costs across different age groups. If we want to simulate how the economy would evolve in the long run for a future policy change in twenty years, it might be problematic because the future demography is different. However, the main counterfactual experiments are backward looking; the counterfactuals aim to analyze the implications of China's *past* trade integration on welfare, when there are different magnitudes of internal frictions. Hence, the changing demography will not invalidate the results.

Another potential threat is that in 2005, the domestic labor markets are not yet in equilibrium; that is, there is potential migration that has not been realized. The existence of those workers will result in overestimating regional fixed effects for the regions experiencing migration outflows, and given the observed wages, this will in turn be reflected in overestimated amenities in these regions; similarly, I will under-estimate the amenities in popular migration destinations. In quantification, I find large dispersion in amenities, and if this argument is true, the real dispersion will be larger. In counterfactual experiments, however, since I keep the amenities fixed, the biases in the measured amenities will not affect the relative changes in the variables of interest, between trade and autarky equilibria.

References

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

Akay, Alpaslan, Corrado Giulietti, Juan D Robalino, and Klaus F Zimmermann. 2012. "Remittances and Well-being among Rural-to-urban Migrants in China." *Review of Economics of the Household*, 1–30.

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

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

Hillberry, Russell, and David Hummels. 2003. "Intranational Home Bias: Some Explanations." *Review of Economics and Statistics*, 85(4): 1089–1092.

Hillberry, Russell, and David Hummels. 2008. "Trade Responses to Geographic Frictions: A Decomposition Using Micro-data." *European Economic Review*, 52(3): 527–550.

Jones, Charles I. 2013. "Misallocation, Economic Growth, and Input-Output Economics." Vol. 2, 419, Cambridge University Press.

Kennan, John, and James R Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211–251.

Michaels, Guy, Stephen J Redding, and Ferdinand Rauch. 2011. "Technical Note: An Eaton and Kortum (2002) Model of Urbanization and Structural Transformation." *Mimeo*.

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

Redding, Stephen J. 2016. "Goods Trade, Factor Mobility and Welfare." *Journal of International Economics*, 101: 148–167.

Wolf, Holger C. 2000. "Intranational Home Bias in Trade." *Review of Economics and Statistics*, 82(4): 555–563.