Ageing, Domestic Integration, and External Trade Frictions: Welfare and Spatial Inequality in China

June 30, 2026

Abstract

JEL Classification: F11, F43, J11, R12

Keywords: Demographic ageing; domestic market integration; trade fragmentation; spatial inequality

1 Introduction

China is entering a new stage in which demographic change, external shocks, and domestic market integration jointly shape regional economic adjustment. The sharp decline in births and the rise in deaths have pushed China into negative natural population growth since 2022, while the share of the elderly population continues to rise. This demographic transition occurs at a time when China faces persistent foreign trade shocks, especially the U.S.–China trade war and broader geoeconomic fragmentation, and is simultaneously promoting a dual-circulation strategy that places greater weight on domestic demand and internal economic circulation. These forces make the study of population ageing vital. Ageing affects not only aggregate labor supply and consumption demand, but also the spatial allocation of workers, firms, and fiscal resources. Whether China can reduce migration costs, lower domestic trade costs, and strengthen domestic market integration will determine how regions adjust to ageing and external shocks. The issue is therefore closely connected to China's economic slowdown, the sustainability of domestic demand, and concerns over income and spatial inequality.

1. One-to-two paragraphs to motivate paper
   (a) Start w/ well-known fact or result of general importance of a research question
   (b) Then move to a puzzle or a part of the question that has not been addressed
   (c) By the end of the second paragraph, it should be very natural and obvious that your particular question should be pursued
2. One paragraph summary of paper should start before end of page 1
3. 2-3 pages on details of paper: each part of paper Emprical analysis
   (a) Theoretical foundation: 1-2 paragraphs on key model features and mechanisms
   (b) Empirical design: how Theoretical foundation help us buliding our emprical model
      i. Paragraph on data
      ii. Paragraph on methodology: specification, identification, etc.
      iii. 2-3 paragraphs on results and interpretation
   (c) Model
      i. 1 paragraph on calibration, results and interpretation

ii. 2-3 paragraphs on counterfactuals, results and interpretation

4. 2-4 paragraphs on literature review
   (a) Each paragraph captures a different strand of literature
      i. Each paragraph ends with a sentence or two on your contribution relative to the literature
   (b) There should be a sentence on “the most closely related paper”.

2 Stylized facts

Now we have different color for comments. Here are some examples:

- Peiyang
- Jinwei
- ~~Here if we want to delete something~~

2.1 Data

2.2 Empirical Strategy

2.3 Empirical results

Figure 1 plots China’s annual births and deaths from 2005 to 2025, measured in units of 10,000 persons. It shows a sharp reversal in China’s demographic dynamics. From 2005 to 2016, births remained relatively high, fluctuating around 16–18 million per year, while deaths increased gradually from about 8.5 million to around 9.8 million. The birth series reached its peak in 2016, after which it declined rapidly and continuously.

A key turning point appears in 2022, when deaths exceeded births for the first time in this period. Since then, the gap between the two series has widened. By 2025, births had fallen to 7.92 million, while deaths had risen to 11.31 million. This divergence indicates that China has entered a phase of negative natural population growth.

Overall, the figure highlights the urgency of studying population ageing in China. The decline in births reduces the future working-age population, while the steady rise in deaths reflects the growing weight of older cohorts. Together, these trends suggest that demographic

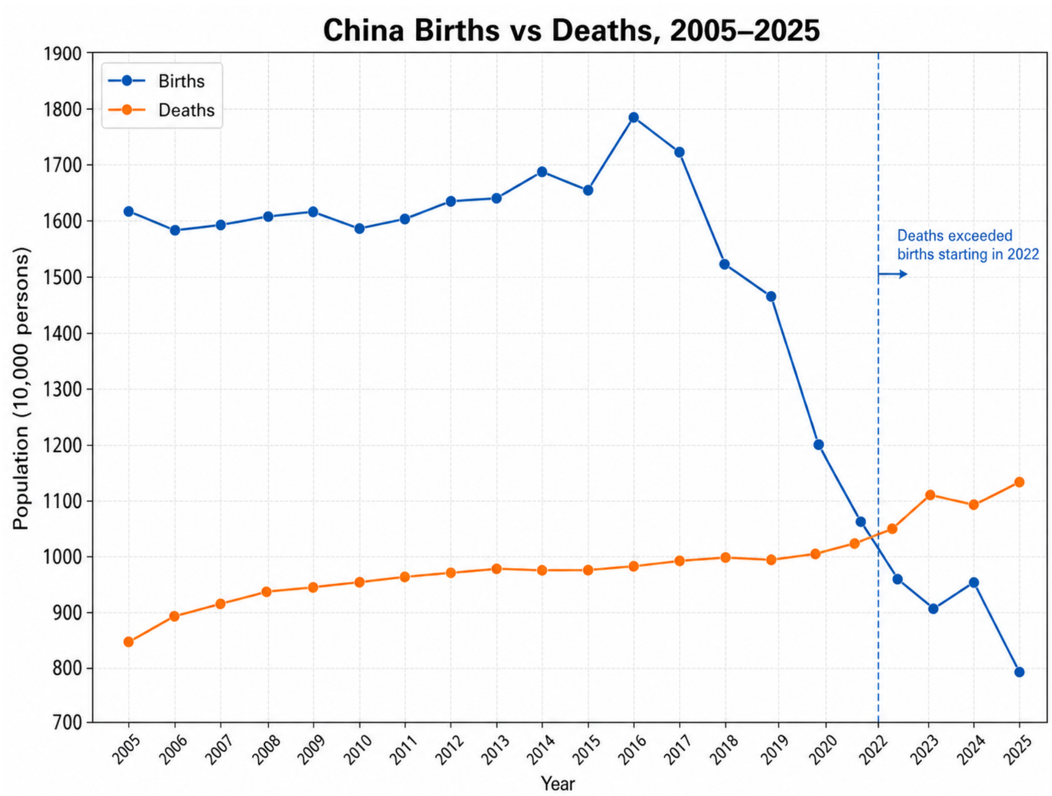

Figure 1: China Births and Deaths during 2005–2025

Notes: The figure plots China's total numbers of birth and death population data during 2005–2025. Source: National Bureau of Statistics of China

ageing is becoming a central force shaping China's labor supply, domestic demand, regional migration, fiscal pressure, and long-term economic growth.

Table 1 shows clear heterogeneity in age structure across China's regions in 2015. Guangdong has a relatively young demographic profile, with a mean age of 34.71 for its resident population and 35.00 for its registered, or hukou, population. By contrast, Northeast China is substantially older: the mean age is 41.51 for residents and 41.45 for the hukou population. The share of older individuals is also much higher in Northeast China. People aged 56 and above account for 25.52% of Northeast China's resident population, compared with only 16.06% in Guangdong. These patterns indicate that population ageing is much more severe in Northeast China, while Guangdong remains demographically younger.

The table also shows that migrant populations are generally much younger than both resident and hukou populations. In Guangdong, the mean age of in-migrants is 32.36, lower than that of both residents and the hukou population. In Northeast China, the difference is even more striking: in-migrants have a mean age of 32.44, almost nine years younger than the

Table 1: Age Structure of the Population in Guangdong and Northeast China (2015)

Region	Population type	Sample size	Mean age	Ages 0–15	Ages 16–25	Ages 26–55	Ages 56–65	Ages 66+
Guangdong	Resident	158,139	34.71	17.42%	18.17%	48.35%	8.86%	7.20%
Guangdong	In-migrant	21,262	32.36	10.19%	19.98%	65.02%	3.38%	1.42%
Guangdong	Out-migrant	2,254	30.77	11.00%	23.74%	61.45%	2.44%	1.38%
Guangdong	Hukou	139,131	35.00	18.42%	17.99%	46.01%	9.59%	7.99%
Northeast	Resident	158,923	41.51	11.71%	10.57%	52.19%	15.10%	10.42%
Northeast	In-migrant	1,371	32.44	11.60%	23.49%	56.16%	5.62%	3.14%
Northeast	Out-migrant	4,777	36.86	9.17%	15.51%	58.68%	11.60%	5.05%
Northeast	Hukou	162,329	41.45	11.64%	10.60%	52.35%	15.08%	10.33%

Notes: Northeast China includes Liaoning, Jilin, and Heilongjiang. The resident population is defined by place of residence, while the Hukou population is defined by place of household registration. The in-migrant population refers to individuals whose place of residence is in the region but whose household registration is outside the region. The out-migrant population refers to individuals whose household registration is in the region but whose place of residence is outside the region. Age-group shares are calculated using unweighted individual-level observations.

overall resident population. Migrants are also highly concentrated in the prime working-age group. For example, 65.02% of Guangdong's in-migrants and 56.16% of Northeast China's in-migrants are aged 26–55. This suggests that migration is strongly age-selective, with young and working-age individuals being more mobile.

The scale and direction of migration further distinguish Guangdong from Northeast China. Guangdong receives a much larger inflow population, with 21,262 in-migrants in the sample, while its outflow population is relatively small, at 2,254. In contrast, Northeast China has a much smaller inflow population, only 1,371, but a larger outflow population, 4,777. This contrast is consistent with Guangdong's role as a major destination for internal migration, while Northeast China faces stronger population outflows. Overall, the table suggests that regional demographic divergence in China is shaped not only by differences in local age structure, but also by selective migration flows that bring younger workers into Guangdong and reinforce ageing pressures in Northeast China.

To explore the relationship between labor income and age, we conduct a quadratic fit of age and individual labor income as follows:

$$\log(\text{Income}) = \beta_1 \times \text{Age} + \beta_2 \times \text{Age2} + \text{Controls} + \epsilon. \tag{1}$$

The estimation results are demonstrated in Table 2. We clearly find a hump-shape of income with age as depicted in Figure 2. The pattern is evident, no matter we examine the full population survey (Columns (1)–(3)) or only the working population surveyed (Columns (4)–(6)), with or without more individual characteristics such as gender, education, marital status, ethnicity, hukou type (urban or rural), occupation, sector of employment, and region controlled for. In general, the implied age with max income is around 38 (calculated as $\frac{\beta_1}{2\beta_2}$). These facts thus motivate a spatial quantitative general equilibrium structural model with labor mobility, age heterogeneity, interregional and international trade as will be elaborated in the following section.

Table 2: The Hump-Shape of Income and Age

	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.0551***	0.0385***	0.0297***	0.0634***	0.0388***	0.0285***
	(0.0003)	(0.0002)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
Age2	-0.0008***	-0.0005***	-0.0004***	-0.0009***	-0.0005***	-0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	1,382,263	1,382,262	1,381,554	1,345,600	1,345,600	1,344,912
Adjusted R^2	0.0596	0.6322	0.6419	0.0421	0.6305	0.6404
Implied max income age	33.4518	40.1970	37.8982	33.5509	40.0219	37.7915

Notes: This table reports regression results for how individual labor income varies with age and age squared, following Equation (1). Columns (1)–(3) use the full population sample, while Columns (4)–(6) use the working-age sample. Columns (1) and (4) include only age controls. Columns (2) and (5) additionally include gender, education, sector of employment, and region fixed effects. Columns (3) and (6) further include gender, education, marital status, ethnicity, hukou type (urban or rural), occupation, sector of employment, and region fixed effects. The implied income-maximizing age is calculated as $-\beta_1/(2\beta_2)$ in Equation (1). Standard errors are reported in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

3 Model

The theoretical framework builds on the quantitative trade literature of Eaton and Kortum (2002) and Caliendo and Parro (2015), as well as the quantitative spatial literature on trade, labor mobility, and the geography of economic activity, including Allen and Arkolakis (2014), Monte, Redding, and Rossi-Hansberg (2018), Redding and Rossi-Hansberg (2017), and Caliendo, Dvorkin, and Parro (2019). The economy consists of N regions. Let $\mathbb{N}_0 = \{1, \ldots, N-1\}$ denote the set of Chinese regions, and let region N denote the aggregated rest of the world. There are J sectors. Production features roundabout input–output linkages, so sectors use intermediate inputs from each other and goods are traded both across Chinese regions and internationally. Labor does not move across national borders. Within China, however, workers reallocate across regions subject to migration frictions and idiosyncratic uncertainty.

The model is designed to study how population ageing, domestic market integration, and external shocks jointly shape aggregate income and its spatial distribution in China. Relative to the existing literature, it introduces three features into a unified framework. First, labor

Figure 2: The Hump-Shape of Income and Age

Notes: The figure plots the hump-shape of (log) labor income with respect to age, following Equation(1). The red curve is the quadratic fitted line; while the green vertical line indicates the implied age with max income by the regression coefficients (calculated as $\frac{\beta_1}{2\beta_2}$).

is heterogeneous by age: different age groups differ in labor force participation, efficiency units of labor, and migration frictions, so population ageing affects both the quantity and the spatial mobility of effective labor. Second, domestic market integration is captured by declines in trade costs for both final goods and intermediate inputs across Chinese regions, together with declines in internal migration frictions, thereby reshaping the allocation of production and labor within China. Third, Chinese regions are linked to the external market through exports and imported intermediate inputs, so external shocks—such as weaker foreign demand, higher international trade costs, or disruptions to imported inputs—affect regions asymmetrically depending on their external exposure.

3.1 Households

Households are heterogeneous by age. Economic activity in the model begins at age 16 and ends with retirement after age 65, so the household side focuses on the working-age population. Let

$$\mathcal{G}^W \equiv \{16, 17, \ldots, 65\}$$

denote the set of working ages. For each Chinese region $n \in \mathbb{N}_0$ and age $g \in \mathcal{G}^W$, let $N_{n,g}$ denote the number of age-g households attached to home region n. The age distribution $\{N_{n,g}\}$ is taken as exogenous.

Among working-age households, let $\phi_{n,g} \in [0,1]$ denote the share of age-g households from home region n that do not participate in the labor market. The corresponding labor force participation rate is $1 - \phi_{n,g}$. Conditional on participation, each worker supplies l^g efficiency units of labor, where l^g captures the age profile of labor efficiency. We normalize $l^{16} = 1$.

Define the number of participating age-g workers attached to home region n by

$$\bar{L}_{n,g} = (1 - \phi_{n,g})N_{n,g}, \qquad g \in \mathcal{G}^W, \tag{2}$$

and the corresponding effective labor endowment by

$$\bar{L}^e_{n,g} = (1 - \phi_{n,g})\, l^g\, N_{n,g} = l^g \bar{L}_{n,g}. \tag{3}$$

Aggregating across working ages,

$$\bar{L}_n = \sum_{g \in \mathcal{G}^W} \bar{L}_{n,g}, \qquad \bar{L}^e_n = \sum_{g \in \mathcal{G}^W} \bar{L}^e_{n,g}. \tag{4}$$

Thus, the age structure of the labor force affects labor supply both through the size of age cohorts and through the composition of effective labor.

A participating age-g worker who lives and works in region $n \in \mathbb{N}_0$ consumes sectoral composites $\{c^j_{n,g}\}^J_{j=1}$ and aggregate consumption $c_{n,g}$, with period utility

$$U(c_{n,g}) = \log(c_{n,g}).$$

Following Caliendo, Dvorkin, and Parro (2019), we model trade imbalances as arising from income transfers across regions. Since W_n is the price of one efficiency unit of labor in region n, labor income of an age-g worker is $W_n l^g$. Let $c^j_{n,g}$ denote the worker's consumption of the sector-j composite and let $c_{n,g}$ denote aggregate consumption. The worker chooses $\{c^j_{n,g}\}^J_{j=1}$ subject to

$$\sum_{j=1}^{J} P^j_n c^j_{n,g} = P_n c_{n,g} = W_n l^g + \chi_{n,g}. \tag{5}$$

Let φ_n denote the remittance ratio in region n. Each worker remits a fraction φ_n of labor income to a global portfolio. In return, the pooled portfolio rebates a common transfer T per efficiency unit of labor. Formally, the net transfer received by an age-g worker in region n is

$$\chi_{n,g} = -\varphi_n W_n l^g + T l^g, \qquad T = \frac{\sum_{n=1}^{N} \varphi_n W_n L_n^e}{\sum_{n=1}^{N} L_n^e}. \tag{6}$$

The associated regional trade imbalance is

$$D_n = \sum_{g \in \mathcal{G}^W} \chi_{n,g} L_{n,g} = -\varphi_n W_n L_n^e + T L_n^e. \tag{7}$$

We model internal migration within China as workers choosing across Chinese regions subject to migration frictions that generate barriers to interregional labor flows. Each participating age-g worker is registered in a home region $m \in \mathbb{N}_0$ and chooses a destination region $n \in \mathbb{N}_0$; in the model, the work location coincides with the residence location.

The migration utility has three components. First, by moving to region n, the worker obtains the real payoff $\Lambda_{n,g}$, which determines consumption $c_{n,g}$ and thus utility from goods. Should we write down explicitly the formula for $\Lambda_{n,g}$? Second, the proportional term $\nu_{nm,g} > 0$, referred to as the migration friction, captures the utility loss from leaving the home region; geographic distance, hukou-related barriers, relocation costs, and other impediments to interregional mobility all affect this cost. For workers who remain in their home region, we normalize $\nu_{mm,g} = 1$. Third, the idiosyncratic term $z_{nm,g}(\epsilon)$ is a preference shifter for moving and generates heterogeneity among workers making the same origin–destination choice.

The indirect utility for an age-g worker from region m that chooses to live and work in region n is therefore

$$V_{nm,g}(\epsilon) = \frac{z_{nm,g}(\epsilon)\, \Lambda_{n,g}}{\nu_{nm,g}}. \tag{8}$$

For each worker ϵ, the idiosyncratic preference shifter $z_{nm,g}(\epsilon)$ is independent and identically distributed across all destination choices. For each age g, it follows a Fréchet distribution with cumulative distribution function

$$G^g(z) = \exp\left(-z^{-\kappa^g}\right), \qquad z > 0,$$

where $\kappa^g > 0$ is the age-specific migration elasticity. A larger κ^g implies less dispersion in idiosyncratic tastes and hence a higher elasticity of migration with respect to the deterministic component $\Lambda_{n,g}/\nu_{nm,g}$.

Let $\mu_{nm,g}$ denote the fraction of age-g workers registered in region m that choose to work in region n, so that

$$\sum_{n\in\mathbb{N}_0} \mu_{nm,g} = 1 \qquad \forall(m,g).$$

By the law of large numbers, the migration shares are

$$\mu_{nm,g} = \frac{\left(\dfrac{\Lambda_{n,g}}{\nu_{nm,g}}\right)^{\kappa^g}}{\sum_{i\in\mathbb{N}_0}\left(\dfrac{\Lambda_{i,g}}{\nu_{im,g}}\right)^{\kappa^g}}. \tag{9}$$

Given the initial distribution of workers across home regions and ages, destination-specific real payoffs, and age-specific migration frictions,

$$\{\bar{L}_{m,g}, \bar{L}^e_{m,g}, \Lambda_{n,g}, \nu_{nm,g}\},$$

these expressions imply a unique partial-equilibrium allocation of labor flows $\{\mu_{nm,g}\}$. The number of age-g workers located in destination region n is

$$L_{n,g} = \sum_{m\in\mathbb{N}_0} \mu_{nm,g}\bar{L}_{m,g}, \tag{10}$$

and the corresponding effective labor supply is

$$L^e_{n,g} = \sum_{m\in\mathbb{N}_0} \mu_{nm,g}\bar{L}^e_{m,g} = l^g L_{n,g}. \tag{11}$$

Aggregating across ages,

$$L_n = \sum_{g\in\mathcal{G}^W} L_{n,g}, \qquad L^e_n = \sum_{g\in\mathcal{G}^W} L^e_{n,g}. \tag{12}$$

For the aggregated rest of the world, labor is internationally immobile; L_N and L^e_N are taken as exogenous.

3.2 Production

In each region $n = 1, \ldots, N$ and sector $j = 1, \ldots, J$, competitive firms produce a continuum of tradable varieties indexed by $\omega \in [0,1]$. Production uses effective labor and sectoral composite intermediates under a constant-returns-to-scale technology:

$$y_n^j(\omega) = q_n^j(\omega) \left(\ell_n^j(\omega)\right)^{\gamma_n^j} \prod_{k=1}^{J} \left(m_n^{jk}(\omega)\right)^{\gamma_n^{jk}}, \tag{13}$$

where $\ell_n^j(\omega)$ is effective labor employed in producing variety ω, and $m_n^{jk}(\omega)$ is the quantity of the sector-k composite input used in sector j. Constant returns imply

$$\gamma_n^j + \sum_{k=1}^{J} \gamma_n^{jk} = 1.$$

Following Eaton and Kortum (2002), productivity is heterogeneous across varieties. The productivity draw $q_n^j(\omega)$ follows a Fréchet distribution,

$$F_n^j(q) = \exp\left(-\lambda_n^j q^{-\theta}\right),$$

where λ_n^j is the region–sector productivity parameter and $\theta > 0$ governs dispersion.

Cost minimization implies that the unit cost of producing a sector-j variety in region n is

$$c_n^j = \Upsilon_n^j \, W_n^{\gamma_n^j} \prod_{k=1}^{J} (P_n^k)^{\gamma_n^{jk}}, \tag{14}$$

where W_n is the price of one efficiency unit of labor in region n, P_n^k is the price of the sector-k composite in region n, and

$$\Upsilon_n^j = (\gamma_n^j)^{-\gamma_n^j} \prod_{k=1}^{J} \left(\gamma_n^{jk}\right)^{-\gamma_n^{jk}}.$$

3.3 Trade

Goods are traded both across Chinese regions and between China and the rest of the world. Trade is subject to iceberg costs. Delivering one unit of a sector-j variety from region i to region n requires shipping $\kappa_{ni}^j \geq 1$ units, with $\kappa_{nn}^j = 1$.

In each destination region n, sector-j varieties are aggregated into a sectoral composite according to a CES aggregator,

$$Q_n^j = \left[\int_0^1 r_n^j(\omega)^{\frac{\sigma^j - 1}{\sigma^j}} \, d\omega\right]^{\frac{\sigma^j}{\sigma^j - 1}}, \tag{15}$$

where $r_n^j(\omega)$ denotes the quantity of sector-j variety ω used in region n, and $\sigma^j > 1$ is the elasticity of substitution across varieties within sector j.

Standard Eaton–Kortum arguments imply that the price of the sector-j composite in region n is

$$P_n^j = A^j \left[\sum_{i=1}^{N} \lambda_i^j \left(\kappa_{ni}^j c_i^j \right)^{-\theta} \right]^{-1/\theta}, \tag{16}$$

where

$$A^j = \left[\Gamma\left(\frac{1+\theta-\sigma^j}{\theta} \right) \right]^{\frac{1}{1-\sigma^j}}.$$

Let π_{ni}^j denote the share of region n's expenditure on sector-j goods sourced from region i. Then

$$\pi_{ni}^j = \frac{\lambda_i^j \left(c_i^j \kappa_{ni}^j \right)^{-\theta}}{\sum_{m=1}^{N} \lambda_m^j \left(c_m^j \kappa_{nm}^j \right)^{-\theta}}. \tag{17}$$

These trade shares summarize sourcing patterns across all Chinese regions and the rest of the world.

3.4 Aggregation and Market Clearing

Workers consume sectoral composites and aggregate them into a final good. Let $c_{n,g}^j$ denote per-worker consumption of the sector-j composite by an age-g worker in region n, and let $c_{n,g}$ denote the associated final consumption level. We assume Cobb–Douglas preferences across sectors:

$$c_{n,g} = \prod_{j=1}^{J} \left(c_{n,g}^j \right)^{\alpha_n^j}, \qquad \sum_{j=1}^{J} \alpha_n^j = 1. \tag{18}$$

The corresponding final-good price index is

$$P_n = \prod_{j=1}^{J} \left(\frac{P_n^j}{\alpha_n^j} \right)^{\alpha_n^j}. \tag{19}$$

Conditional on total expenditure $P_n c_{n,g}$, Cobb–Douglas demand implies

$$P_n^j c_{n,g}^j = \alpha_n^j P_n c_{n,g}. \tag{20}$$

Aggregate final consumption in region n is obtained by summing individual consumption across participating workers:

$$C_n = \sum_{g \in \mathcal{G}^W} c_{n,g} L_{n,g}, \qquad C_n^j = \sum_{g \in \mathcal{G}^W} c_{n,g}^j L_{n,g}. \tag{21}$$

Since expenditure shares are common across workers within a region, aggregation gives

$$P_n^j C_n^j = \alpha_n^j P_n C_n. \tag{22}$$

Summing the worker-level budget constraints across ages yields

$$P_n C_n = \sum_{g \in \mathcal{G}^W} \left(W_n l^g + \chi_{n,g} \right) L_{n,g} = W_n L_n^e + D_n, \tag{23}$$

where the regional trade imbalance is defined by

$$D_n \equiv \sum_{g \in \mathcal{G}^W} \chi_{n,g} L_{n,g} = -\varphi_n W_n L_n^e + T L_n^e. \tag{24}$$

By construction,

$$\sum_{n=1}^{N} D_n = 0. \tag{25}$$

Let

$$X_n^j \equiv P_n^j Q_n^j \tag{26}$$

denote total expenditure in region n on the sector-j composite. Total expenditure on sector j consists of final demand and intermediate demand:

$$X_n^j = P_n^j C_n^j + \sum_{k=1}^{J} \gamma_n^{kj} Y_n^k = \alpha_n^j P_n C_n + \sum_{k=1}^{J} \gamma_n^{kj} Y_n^k, \tag{27}$$

where γ_n^{kj} is the share of sector-k revenue spent on sector-j intermediates, and Y_n^k is the value of sector-k sales produced in region n.

Goods-market clearing requires that sales of sector-j goods produced in region n equal expenditure on those goods from all destinations:

$$Y_n^j = \sum_{i=1}^{N} \pi_{in}^j X_i^j. \tag{28}$$

Labor-market clearing requires that total payments to labor equal total labor income in each region:

$$W_n L_n^e = \sum_{j=1}^{J} \gamma_n^j Y_n^j, \qquad n = 1, \ldots, N. \tag{29}$$

Together with the household block, these conditions determine wages, goods prices, trade shares, expenditure levels, trade imbalances, and the spatial allocation of effective labor across Chinese regions.

Table 3: Equilibrium conditions

ID	**Level equation**	**Scope**
H1	$\bar{L}_{n,g} = (1-\phi_{n,g})N_{n,g}, \qquad \bar{L}^e_{n,g} = l^g \bar{L}_{n,g}$	$\forall (n,g),\ n \in \mathbb{N}_0$
H2	$\bar{L}_n = \sum_{g\in\mathcal{G}^W} \bar{L}_{n,g}, \qquad \bar{L}^e_n = \sum_{g\in\mathcal{G}^W} \bar{L}^e_{n,g}$	$\forall n \in \mathbb{N}_0$
H3	$\chi_{n,g} = -\varphi_n W_n l^g + T l^g, \qquad T = \frac{\sum_{n=1}^{N} \varphi_n W_n L_n^e}{\sum_{n=1}^{N} L_n^e}$	$\forall (n,g)$
H4	$\Lambda_{n,g} = \frac{W_n l^g + \chi_{n,g}}{P_n}$	$\forall (n,g),\ n \in \mathbb{N}_0$
H5	$\mu_{nm,g} = \frac{(\Lambda_{n,g}/\nu_{nm,g})^{\kappa^g}}{\sum_{i\in\mathbb{N}_0} (\Lambda_{i,g}/\nu_{im,g})^{\kappa^g}}$	$\forall (n,m,g),\ n,m \in \mathbb{N}_0$
H6	$L_{n,g} = \sum_{m\in\mathbb{N}_0} \mu_{nm,g}\bar{L}_{m,g}, \qquad L^e_{n,g} = l^g L_{n,g}$	$\forall (n,g),\ n \in \mathbb{N}_0$
H7	$L_n = \sum_{g\in\mathcal{G}^W} L_{n,g}, \qquad L^e_n = \sum_{g\in\mathcal{G}^W} L^e_{n,g}$	$\forall n$
P1	$c_n^j = \Upsilon_n^j W_n^{\gamma_n^j} \prod_{k=1}^{J} (P_n^k)^{\gamma_n^{jk}}, \qquad \Upsilon_n^j = (\gamma_n^j)^{-\gamma_n^j} \prod_{k=1}^{J} (\gamma_n^{jk})^{-\gamma_n^{jk}}$	$\forall (n,j)$
P2	$P_n^j = A^j \left[\sum_{i=1}^{N} \lambda_i^j (\kappa_{ni}^j c_i^j)^{-\theta}\right]^{-1/\theta}, \qquad A^j = \left[\Gamma\left(\frac{1+\theta-\sigma^j}{\theta}\right)\right]^{\frac{1}{1-\sigma^j}}$	$\forall (n,j)$
P3	$\pi_{ni}^j = \frac{\lambda_i^j (c_i^j \kappa_{ni}^j)^{-\theta}}{\sum_{m=1}^{N} \lambda_m^j (c_m^j \kappa_{nm}^j)^{-\theta}}$	$\forall (n,i,j)$
A1	$P_n = \prod_{j=1}^{J} \left(\frac{P_n^j}{\alpha_n^j}\right)^{\alpha_n^j}, \qquad P_n^j c_{n,g}^j = \alpha_n^j P_n c_{n,g}$	$\forall (n,j,g)$
A2	$C_n = \sum_{g\in\mathcal{G}^W} c_{n,g} L_{n,g}, \qquad C_n^j = \sum_{g\in\mathcal{G}^W} c_{n,g}^j L_{n,g}$	$\forall (n,j)$
A3	$P_n^j C_n^j = \alpha_n^j P_n C_n$	$\forall (n,j)$
A4	$D_n = \sum_{g\in\mathcal{G}^W} \chi_{n,g} L_{n,g} = -\varphi_n W_n L_n^e + T L_n^e$	$\forall n$
A5	$P_n C_n = W_n L_n^e + D_n$	$\forall n$
A6	$X_n^j = P_n^j Q_n^j = P_n^j C_n^j + \sum_{k=1}^{J} \gamma_n^{kj} Y_n^k = \alpha_n^j P_n C_n + \sum_{k=1}^{J} \gamma_n^{kj} Y_n^k$	$\forall (n,j)$
A7	$Y_n^j = \sum_{i=1}^{N} \pi_{in}^j X_i^j$	$\forall (n,j)$
A8	$W_n L_n^e = \sum_{j=1}^{J} \gamma_n^j Y_n^j$	$\forall n$

Equilibrium definition. The model is characterized by exogenous demographic objects $\{N_{n,g}, \phi_{n,g}, l^g\}$, preference parameters $\{\alpha_n^j\}$, technology parameters $\{\lambda_n^j, \gamma_n^j, \gamma_n^{jk}, \theta, \sigma^j\}$, bilateral trade costs $\{\kappa_{ni}^j\}$, migration frictions $\{\nu_{nm,g}\}$, migration elasticities $\{\kappa^g\}$, and transfer wedges $\{\varphi_n\}$. Chinese regions are indexed by $n \in \mathbb{N}_0$, while region N denotes the aggregated rest of the world.

We define a competitive equilibrium below, with the associated conditions summarized in Table 3. A competitive equilibrium consists of allocations

$$\{\mu_{nm,g},\ L_{n,g},\ L_{n,g}^e,\ L_n,\ L_n^e,\ c_{n,g},\ c_{n,g}^j,\ C_n,\ C_n^j,\ Q_n^j,\ X_n^j,\ Y_n^j,\ D_n\},$$

and prices

$$\{W_n,\ T,\ \Lambda_{n,g},\ P_n,\ P_n^j,\ c_n^j,\ \pi_{ni}^j\},$$

such that: (i) participating workers choose destination regions to maximize utility given wages, prices, migration frictions, and transfers; (ii) firms minimize costs and price competitively under constant-returns technologies, Fréchet productivity draws, and iceberg trade costs; (iii) final and intermediate demands are consistent with household and firm optimization; and (iv) all goods and labor markets clear, with global trade imbalances summing to zero.

4 Calibration

This section describes the calibration strategy used to quantify the model. The benchmark economy contains 30 mainland Chinese province-level regions, including provinces, autonomous regions, and municipalities, plus an aggregated rest-of-world region. Thus, $N = 31$, with

$$\mathbb{N}_0 = \{1, \ldots, 30\}$$

denoting Chinese regions and region N denoting the rest of the world. There are three sectors,

$$\mathcal{J} = \{\text{agr.}, \text{ind.}, \text{ser.}\},$$

corresponding to agriculture, industry, and services. Working ages are indexed by

$$\mathcal{G}^W = \{16, 17, \ldots, 65\}.$$

Table B.1 lists the 30 mainland Chinese province-level administrative regions used in the benchmark calibration. These regions include provinces, autonomous regions, and municipalities directly under the central government. Region N is reserved for the aggregated rest of the world.

The calibration combines demographic, input–output, trade, and geographic data. We first use population survey data to discipline the demographic and labor-market objects. The 2005 National 1% Population Sample Survey is used to estimate the age profile of labor efficiency and to discipline age-specific migration elasticities, because the 2005 micro data report income and detailed labor-market information.[1] Since comparable migration-flow information is not available for 2017, we use age-specific migration shares from the 2015 National 1% Population Sample Survey to approximate 2017 migration patterns (Department of Population and Employment Statistics, National Bureau of Statistics of China, 2016). For the remaining demographic objects, including the population distribution across province-level regions and age groups and age-specific labor-force participation rates, we use the corresponding 2017 population data. We then use Penn World Table 10.0 to discipline aggregate population and labor input for China and the rest of the world (Feenstra, Inklaar, and Timmer, 2015; Groningen Growth and Development Centre, 2021). In particular, PWT provides internationally comparable population and employment measures, which we use to construct aggregate China and ROW population and worker totals. China's aggregate population is then allocated across provinces and age groups using the 2017 population data.

We next use input–output data to construct production, expenditure, and trade objects. The China Multi-Regional Input–Output Table for 2017 is used to construct internal trade flows, regional production structure, intermediate input shares, and final expenditure shares across Chinese provinces (Zhao et al., 2024). This table covers 30 province-level regions and 42 socioeconomic sectors. We aggregate the original sectors into agriculture, industry, and services to match the model's three-sector structure. We then use the OECD Inter-Country Input–Output Tables for 2017 to construct trade flows between China and the rest of the world, as well as the production structure and expenditure shares of the rest-of-world region (Yamano et al., 2021; OECD, 2021). The OECD ICIO tables provide a globally consistent accounting framework for production, intermediate input use, final demand, and international trade flows.

We combine the Chinese provincial MRIO table with the OECD ICIO information to obtain a unified dataset linking Chinese regions to the external market. All values are

[1] The 2005 micro data contain income, occupation, gender, sector of employment, hukou status, and other individual controls.

reported in millions of U.S. dollars. To make the two data sources comparable, we rescale the Chinese provincial MRIO table using China's aggregate value added in the OECD ICIO table for 2017 and the corresponding aggregate value added in the Chinese provincial MRIO table.[2]

Finally, we use geographic information to construct bilateral distances and border indicators for the 32-region economy. For the 31 Chinese province-level regions, we approximate each region's location by the geographic coordinates of its provincial capital. For the rest-of-world aggregate, we collect the geographic coordinates of foreign capitals for countries included in the ROW region. We then compute pairwise great-circle distances between all locations. For two Chinese regions $n, i \in \mathbb{N}_0$, $dist_{ni}$ is the great-circle distance between their provincial capitals. For a Chinese region n and the rest-of-world region N, we construct $dist_{nN} = dist_{Nn}$ as a population-weighted average of the distances from region n's provincial capital to foreign capitals.[3] This procedure converts many country-level distances between China and foreign countries into a single distance between each Chinese region and the aggregated ROW. Together, the demographic, input–output, trade, and geographic data allow us to map observed population, labor, production, expenditure, trade, and migration moments into the model.

The calibration proceeds in three steps. First, we set a small number of time-invariant parameters, including the trade elasticity, the within-sector elasticity of substitution, and the age-specific migration elasticities. Second, we construct demographic objects, effective labor supply, production shares, expenditure shares, and trade imbalances directly from the data. Third, we recover sectoral price levels, bilateral trade costs, fundamental productivity, migration elasticities, and migration frictions using model-implied equilibrium relationships and gravity equations. Table 5 summarizes the calibrated objects, notation, and data sources.

4.1 Exogenous parameters

Age-efficiency profile The age-efficiency profile l^g captures the number of efficiency units supplied by one participating age-g worker. We estimate this profile from the 2005 National

[2] Specifically, all Chinese MRIO values are multiplied by the ratio of China's aggregate value added in the OECD ICIO table to China's aggregate value added in the Chinese provincial MRIO table. This normalization preserves the internal provincial structure of the Chinese MRIO table while matching the aggregate scale of China in the OECD ICIO table. When a bilateral trade flow takes the value zero in the constructed trade-flow matrix, we replace it with 1×10^{-19} to avoid numerical problems in calibration.

[3] The ROW aggregate includes 221 foreign countries. The weights are based on year-2000 population from PWT. Thus, the distance between Chinese region n and ROW is $\sum_{r \in ROW} \omega_r^{ROW} dist_{nr}$, where $\omega_r^{ROW} = Pop_{r,2000} / \sum_{s \in ROW} Pop_{s,2000}$.

1% Population Sample Survey, which reports individual income and labor-market characteristics. Specifically, we regress log individual labor income on age dummies and a set of individual controls:

$$\log y_i = \sum_{g \in \mathcal{G}^W} \beta_g \mathbf{1}\{age_i = g\} + X_i'\Gamma + \varepsilon_i, \tag{30}$$

where i indexes individuals, y_i is individual labor income, and X_i includes controls such as gender, education, marital status, ethnicity, hukou type (urban or rural), occupation, sector of employment, and region fixed effects. The age fixed effects capture the life-cycle component of labor efficiency after controlling for observable individual characteristics. We normalize labor efficiency for the youngest working age group to one and construct

$$l^g = \exp(\beta_g - \beta_{16}). \tag{31}$$

Table 4 reports the labor efficiency for different age groups. It can be seen clearly that the middle age group has the highest labor efficiency, compared to the youngest or oldest group. The estimates are robust to different age group classifications and with detailed individual characteristics controlled for. In the following content, we choose the estimates of column (3) as the benchmark, as more individual control variables are incorporated and the age grouping seems more reasonable. Thus, the age-efficiency profiles l^g for young, middle, and old groups are 1.00, 1.06, and 0.92, respectively.[4]

Age-related migration elasticities To estimate age-related migration elasticities κ^a for each of the three age groups using census-based migration flows, we can follow the income-dispersion method in Tombe and Zhu (2019). Under the Fréchet migration structure, log real earnings follow a Gumbel distribution with standard deviation $\pi/(\kappa\sqrt{6})$, so the migration elasticity can be inferred from within-group dispersion in log earnings. In practice, for each age group a, we compute the within-group standard deviation of log nominal earnings in the 2005 Population Survey, where groups are defined by source region, destination region, sector of employment, and age group, and then recover

$$\kappa^a = \frac{\pi}{\sqrt{6}\, sd_a(\log y)}.$$

[4] Given the estimates of column (3) in Table 4, we have $\exp(0.0623) \approx 1.06$ for middle age group and $\exp(-0.0888) \approx 0.92$ for old age group, with young age group efficiency normalized to 1.

Table 4: Age Group Efficiency

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\{age_i = [26,55]\}$	0.0513***	0.1474***	0.0623***			
	(0.0021)	(0.0014)	(0.0018)			
$\mathbf{1}\{age_i = [56,65]\}$	-0.4915***	0.0002	-0.0888***			
	(0.0033)	(0.0022)	(0.0026)			
$\mathbf{1}\{age_i = [30,51]\}$				-0.0020	0.1114***	0.0367***
				(0.0018)	(0.0012)	(0.0014)
$\mathbf{1}\{age_i = [52,65]\}$				-0.4389***	0.0067***	-0.0729***
				(0.0025)	(0.0017)	(0.0019)
Observations	1,345,600	1,345,600	1,344,912	1,345,600	1,345,600	1,344,912
Adjusted R^2	0.0259	0.6290	0.6393	0.0291	0.6276	0.6388

Notes: The table reports age-related labor efficiency, relative to the youngest working age group, with individual labor income as dependent variable. Columns (1)–(3) use age groups 16–25, 26–55, and 56–65; while Columns (4)–(6) use age groups 16–29, 30–51, and 52–65, respectively. Columns (1) and (4) only consider age group indicators, Columns (2) and (5) include gender, education, sector of employment, and region fixed effects; Columns (3) and (6) includes gender, education, marital status, ethnicity, hukou type (urban or rural), occupation, sector of employment, and region fixed effects. Standard errors are presented in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated by ***, **, and *, respectively.

It turns out that age-specific migration elasticities are 3.21, 2.68, 2.21 for young, middle, and old labor groups, respectively. A larger elasticity implies a worker is more sensitive to the migration costs and thus has more willingness to mobilize given the same migration cost shocks. Clearly, the estimates indicate that younger workers are more likely to migrate relative to the older within China.

Other time-invariant elasticities We set the trade elasticity to $\theta = 4$ in the benchmark calibration, following Simonovska and Waugh (2014). The within-sector elasticity of substitution across varieties is set to $\sigma^j = 2$ for all three sectors, following the estimates and calibration practice in Broda and Weinstein (2006).

Population, labor force, and effective labor The model requires population by source region and age, labor-force participation by region and age, and effective labor supply. Let $N_{n,g}$ denote the number of age-g individuals attached to source region n. The source region is identified by hukou registration or hukou ID-number location. Let $\tau_{n,g}$ denote the labor-force participation rate for age group g in region n. Since the model uses $\phi_{n,g}$ to denote the non-participation rate, we set $\phi_{n,g} = 1 - \tau_{n,g}$.

We construct $N_{n,g}$ and $\tau_{n,g}$ from the population survey and macro population data. First, we use census sampling weights to compute the age distribution by hukou/source region.

Second, we compute labor-force participation rates by age and destination region, where the destination region is measured by current residence or work location. Third, we combine the age distribution from the 2015 Population Survey with aggregate population totals, so that region-level population matches the macro totals while still varying by age. Combining this age-varying population object with the participation rate gives the number of participating workers by source region and age:

$$\bar{L}_{n,g} = \tau_{n,g} N_{n,g}. \tag{32}$$

Here, $\bar{L}_{n,g}$ is the number of participating age-g workers attached to source region n before migration choices are applied.

After migration choices are applied, destination-region labor supply is obtained by reallocating source-region workers across destinations using the migration shares $\mu_{nm,g}$. Effective labor in destination region n is then

$$L_n^e = \sum_{g \in \mathcal{G}^W} l^g L_{n,g}, \qquad L_{n,g} = \sum_{m \in \mathbb{N}_0} \mu_{nm,g} \bar{L}_{m,g}. \tag{33}$$

Thus, the age structure affects regional labor supply through both the number of workers in each age group and the age-efficiency profile l^g. For the rest of the world, labor is internationally immobile, so L_N^e is taken as exogenous.

Production shares and input–output linkages We calibrate production shares from the input–output data. Let GO_n^j denote gross output and VA_n^j denote value added in sector j and region n. The effective labor share is

$$\gamma_n^j = \frac{VA_n^j}{GO_n^j}. \tag{34}$$

Let M_n^{jk} denote the value of sector-k intermediates used in sector-j production. The intermediate-input share is

$$\gamma_n^{jk} = \frac{M_n^{jk}}{GO_n^j}, \tag{35}$$

so that

$$\gamma_n^j + \sum_{k=1}^{J} \gamma_n^{jk} = 1. \tag{36}$$

This mapping is consistent with the constant-returns production function in the model.

Final expenditure shares and trade imbalances Let FD_n^j denote final expenditure on sector j in region n. The Cobb–Douglas final expenditure share is

$$\alpha_n^j = \frac{FD_n^j}{\sum_{k=1}^{J} FD_n^k}, \qquad \sum_{j=1}^{J} \alpha_n^j = 1. \tag{37}$$

Total expenditure on the sector-j composite is the sum of final demand and intermediate demand:

$$X_n^j = FD_n^j + \sum_{k=1}^{J} M_n^{kj}. \tag{38}$$

Let X_{ni}^j denote region n's expenditure on sector-j goods sourced from region i. The bilateral expenditure share is

$$\pi_{ni}^j = \frac{X_{ni}^j}{X_n^j}. \tag{39}$$

We measure the regional trade imbalance as total expenditure minus total sales:

$$D_n^{data} = \sum_{j=1}^{J} X_n^j - \sum_{j=1}^{J} Y_n^j, \tag{40}$$

where Y_n^j is total sales of sector j produced in region n. A positive value of D_n^{data} indicates that region n runs a trade deficit.

The transfer wedge φ_n is chosen to match the observed trade imbalance:

$$D_n^{data} = -\varphi_n W_n L_n^e + T L_n^e, \qquad T = \frac{\sum_{n=1}^{N} \varphi_n W_n L_n^e}{\sum_{n=1}^{N} L_n^e}. \tag{41}$$

Because only net transfers matter for allocations, we normalize $T = 0$. Under this normalization,

$$\varphi_n = -\frac{D_n^{data}}{W_n L_n^e}. \tag{42}$$

4.2 Exogenous shocks

This subsection describes how we recover the model-implied shocks and wedges. These objects include bilateral trade costs, sectoral price levels, fundamental productivity, and migration frictions. The recovery uses the structural gravity relationships implied by the Eaton–Kortum environment with input-output linkages, as in Eaton and Kortum (2002).

Gravity equation We use the structural gravity equation to discipline sectoral price levels, bilateral trade costs, and fundamental productivity. The model implies

$$\pi_{ni}^{j} = \frac{\lambda_i^j (c_i^j \kappa_{ni}^j)^{-\theta}}{\sum_{m=1}^{N} \lambda_m^j (c_m^j \kappa_{nm}^j)^{-\theta}}, \qquad \kappa_{nn}^j = 1. \tag{43}$$

This trade-share equation follows the Eaton–Kortum structure and its multi-sector input-output extension (Eaton and Kortum, 2002; Caliendo and Parro, 2015). Define

$$\mathcal{F}_n^j \equiv \log \left[\lambda_n^j (c_n^j)^{-\theta} \right], \tag{44}$$

which summarizes region-sector competitiveness before bilateral trade costs.

Taking the ratio of bilateral expenditure to home expenditure gives

$$\log \left(\frac{X_{ni}^j}{X_{nn}^j} \right) = \mathcal{F}_i^j - \mathcal{F}_n^j - \theta \log \kappa_{ni}^j, \qquad n \neq i. \tag{45}$$

Here, X_{ni}^j is region n's expenditure on sector-j goods sourced from region i, and X_{nn}^j is region n's expenditure on its own sector-j goods.

Following the gravity literature (Anderson and van Wincoop, 2003; Eaton and Kortum, 2002), we parameterize bilateral trade costs as

$$\log \kappa_{ni}^j = \rho_1^j \log dist_{ni} + ex_i^j + u_{ni}^j, \tag{46}$$

where $dist_{ni}$ is bilateral distance and ex_i^j captures origin-specific trade-cost components.

Combining equations (45) and (46), the estimating equation becomes

$$\log \left(\frac{X_{ni}^j}{X_{nn}^j} \right) = \underbrace{\mathcal{F}_i^j - \theta ex_i^j}_{\text{exporter fixed effect}} - \underbrace{\mathcal{F}_n^j}_{\text{importer fixed effect}} - \theta \rho_1^j \log dist_{ni} + \varepsilon_{ni}^j, \qquad n \neq i. \tag{47}$$

We estimate equation (47) separately by sector. The importer fixed effects recover $-\mathcal{F}_n^j$, up to a normalization, while the exporter fixed effects absorb $\mathcal{F}_i^j - \theta ex_i^j$. We normalize the fixed effect of a reference region, or equivalently normalize the national average price level, to pin down the level of sectoral prices and productivity. When bilateral trade flows are zero or extremely small, we replace them with a small positive number before taking logs.

Bilateral trade costs We directly invert the trade-share equation. This inversion follows the structural gravity logic in Eaton and Kortum (2002) and Caliendo and Parro (2015). From equation (45),

$$\log \kappa_{ni}^j = -\frac{1}{\theta}\left[\log\left(\frac{X_{ni}^j}{X_{nn}^j}\right) - \mathcal{F}_i^j + \mathcal{F}_n^j\right], \qquad n \neq i. \tag{48}$$

Equivalently, since $\pi_{ni}^j = X_{ni}^j / X_n^j$,

$$\kappa_{ni}^j = \left[\left(\frac{\pi_{ni}^j}{\pi_{nn}^j}\right)\exp(\mathcal{F}_n^j - \mathcal{F}_i^j)\right]^{-1/\theta}, \qquad \kappa_{nn}^j = 1. \tag{49}$$

This expression recovers the bilateral iceberg-equivalent trade cost that rationalizes the observed trade share π_{ni}^j.

Wages, unit costs, and sectoral price levels Labor-market clearing pins down the wage per efficiency unit of labor:

$$W_n L_n^e = \sum_{j=1}^{J} \gamma_n^j Y_n^j, \qquad W_n = \frac{\sum_{j=1}^{J} \gamma_n^j Y_n^j}{L_n^e}. \tag{50}$$

Given wages, input shares, and sectoral prices, the unit cost of producing sector j in region n is

$$c_n^j = \Upsilon_n^j W_n^{\gamma_n^j} \prod_{k=1}^{J} (P_n^k)^{\gamma_n^{jk}}, \qquad \Upsilon_n^j = (\gamma_n^j)^{-\gamma_n^j} \prod_{k=1}^{J} (\gamma_n^{jk})^{-\gamma_n^{jk}}. \tag{51}$$

This cost structure follows the multi-sector input-output production environment in Caliendo and Parro (2015).

The sectoral price index follows from the Eaton–Kortum price system:

$$P_n^j = A^j \left[\sum_{i=1}^{N} \lambda_i^j (\kappa_{ni}^j c_i^j)^{-\theta}\right]^{-1/\theta}, \qquad A^j = \left[\Gamma\left(\frac{1+\theta-\sigma^j}{\theta}\right)\right]^{\frac{1}{1-\sigma^j}}. \tag{52}$$

Using the home expenditure share and the definition of $\mathcal{F}_n^j$, we can rewrite the price index as

$$P_n^j = A^j \left[\frac{\exp(\mathcal{F}_n^j)}{\pi_{nn}^j}\right]^{-1/\theta}. \tag{53}$$

Thus, after estimating the gravity equation and observing the home share π_{nn}^j, we recover P_n^j. Since price levels are identified only up to a normalization, we normalize the aggregate price index of a reference region, or the national average price level, to one. In practice, we solve equations (51) and (53) jointly by iterating on $\{P_n^j\}$ until convergence.

Fundamental productivity After recovering $\mathcal{F}_n^j$, P_n^j, and c_n^j, we recover fundamental productivity. By definition,

$$\mathcal{F}_n^j = \log\left[\lambda_n^j (c_n^j)^{-\theta}\right], \tag{54}$$

so

$$\lambda_n^j = \exp(\mathcal{F}_n^j)(c_n^j)^{\theta}. \tag{55}$$

Equivalently, combining the price-index equation with the home expenditure share gives

$$\lambda_n^j = \pi_{nn}^j \left(\frac{A^j c_n^j}{P_n^j}\right)^{\theta}. \tag{56}$$

This expression follows the Ricardian selection logic in Eaton–Kortum models (Eaton and Kortum, 2002). For a given unit cost and price level, a lower home share implies a lower calibrated λ_n^j, because the region is less likely to be the lowest-cost supplier in its own market. We report productivity as $(\lambda_n^j)^{1/\theta}$ and normalize it relative to a reference region or the national average.

Migration shares and migration frictions We use census micro data to construct age-specific migration shares. Let $M_{nm,g}$ denote the number of age-g individuals whose source region is m and destination region is n. The source region is identified by hukou location, while the destination region is identified by current residence or work location. The migration share is

$$\mu_{nm,g} = \frac{M_{nm,g}}{\sum_{i\in\mathbb{N}_0} M_{im,g}}, \qquad \sum_{n\in\mathbb{N}_0} \mu_{nm,g} = 1. \tag{57}$$

The model implies

$$\mu_{nm,g} = \frac{\left(\Lambda_{n,g}/\nu_{nm,g}\right)^{\kappa^g}}{\sum_{i\in\mathbb{N}_0}\left(\Lambda_{i,g}/\nu_{im,g}\right)^{\kappa^g}}. \tag{58}$$

Using the stayer share as the reference and imposing $\nu_{mm,g} = 1$, we recover bilateral migration frictions as

$$\nu_{nm,g} = \left(\frac{\mu_{nm,g}}{\mu_{mm,g}}\right)^{-1/\kappa^g} \frac{\Lambda_{n,g}}{\Lambda_{m,g}}, \qquad n \neq m. \tag{59}$$

The real payoff is

$$\Lambda_{n,g} = \frac{W_n l^g + \chi_{n,g}}{P_n}. \tag{60}$$

Thus, conditional on real-payoff differences, a lower migration share from m to n implies a higher migration friction $\nu_{nm,g}$.

4.3 Calibration Results and Model Fit

This subsection summarizes the calibration results and then assesses the quality of the calibration by feeding the calibrated parameters and wedges back into the model. We compare the model-generated values with their data counterparts for several key equilibrium objects, including wages, sectoral expenditures, trade shares, and migration shares.

Table 5 reports the main calibrated parameters and exogenous objects used in the benchmark economy. For objects that vary across regions, sectors, ages, or bilateral pairs, the table reports the mean and the 25th, 50th, and 75th percentiles in parentheses. Sector-specific objects are reported separately for agriculture, manufacturing, and services, denoted by a, m, and s. For γ_n^{jk}, moments are reported by producing sector j, pooling over regions n and input sectors $k \in \{a, m, s\}$. Age-specific objects are reported by age group g, while scalar elasticities are reported as single values. For fundamental productivity, trade costs, and migration frictions, the reported means are weighted averages using value added, bilateral trade flows, and bilateral migration flows, respectively; the percentile moments are unweighted. For population shares, the first number is the national population share of each age group, while the percentile moments describe the cross-region distribution of age-group population shares. Fundamental productivity is normalized relative to Beijing, the first region in the dataset. Trade-cost and migration-friction moments exclude own-region observations, $n = i$ for trade costs and $n = m$ for migration frictions.

Table 5: Calibration Summary

Object	**Notation**	**Value / Moments**	**Source / Method**
Panel A. Parameters			
Trade elasticity	θ	4	Following Simonovska and Waugh (2014)
Variety elasticity	σ^j	2, $\forall j$	Following Broda and Weinstein (2006)
Migration elasticity	κ^g	{3.21, 2.68, 2.21}	2005 Population Survey
Age-efficiency profile	l^g	{1.00, 1.06, 0.92}	2005 Population Survey
Labor shares	γ_n^a	0.59 (0.57, 0.59, 0.65)	China MRIO and OECD ICIO tables
	γ_n^m	0.26 (0.22, 0.24, 0.29)	China MRIO and OECD ICIO tables

Continued on next page

Table 5: Calibration Summary (continued)

Object	**Notation**	**Value / Moments**	**Source / Method**
	γ_n^s	0.46 (0.44, 0.46, 0.47)	China MRIO and OECD ICIO tables
Input-output shares	$\gamma_n^{ak},\ \forall k$	0.19 (0.05, 0.16, 0.26)	China MRIO and OECD ICIO tables
	$\gamma_n^{mk},\ \forall k$	0.19 (0.06, 0.14, 0.25)	China MRIO and OECD ICIO tables
	$\gamma_n^{sk},\ \forall k$	0.18 (0.07, 0.14, 0.26)	China MRIO and OECD ICIO tables
Final expenditure shares	α_n^a	0.05 (0.03, 0.05, 0.06)	China MRIO and OECD ICIO tables
	α_n^m	0.24 (0.21, 0.25, 0.27)	China MRIO and OECD ICIO tables
	α_n^s	0.71 (0.68, 0.71, 0.76)	China MRIO and OECD ICIO tables
Panel B. Exogenous shocks			
Fundamental productivity	$(\lambda_n^a)^{1/\theta}$	0.40 (0.16, 0.25, 0.28)	Home shares, prices, and unit costs
	$(\lambda_n^m)^{1/\theta}$	0.68 (0.57, 0.79, 1.04)	Home shares, prices, and unit costs
	$(\lambda_n^s)^{1/\theta}$	0.23 (0.34, 0.39, 0.42)	Home shares, prices, and unit costs
Migration frictions	ν_{nm,g_1}	8.16 (5.81, 9.28, 16.75)	Migration shares and real payoffs
	ν_{nm,g_2}	11.73 (10.71, 17.71, 31.13)	Migration shares and real payoffs
	ν_{nm,g_3}	28.53 (30.69, 73.16, 100.00)	Migration shares and real payoffs
Migration frictions, all age groups	$\nu_{nm,g}, \forall g$	11.75 (9.40, 20.15, 65.93)	Migration shares and real payoffs
Internal trade costs	κ_{ni}^a	1.75 (4.02, 13.03, 40.04)	Price and trade-share inversion
	κ_{ni}^m	2.07 (2.01, 2.71, 3.81)	Price and trade-share inversion
	κ_{ni}^s	2.58 (2.60, 3.60, 5.86)	Price and trade-share inversion
Internal trade costs, all sectors	$\kappa_{ni}^j,\ \forall j$	2.23 (2.44, 3.75, 8.20)	Price and trade-share inversion
External trade costs	κ_{ni}^a	3.02 (2.46, 4.28, 11.60)	Price and trade-share inversion
	κ_{ni}^m	2.96 (1.76, 3.37, 6.29)	Price and trade-share inversion
	κ_{ni}^s	3.97 (4.89, 10.15, 22.56)	Price and trade-share inversion
External trade costs, all sectors	$\kappa_{ni}^j,\ \forall j$	3.13 (2.40, 5.63, 10.83)	Price and trade-share inversion
Population shares	s_{n,g_1}^N	0.24 (0.12, 0.16, 0.18)	Population Survey and macro totals
	s_{n,g_2}^N	0.63 (0.66, 0.67, 0.69)	Population Survey and macro totals
	s_{n,g_3}^N	0.14 (0.14, 0.17, 0.20)	Population Survey and macro totals
Participation rate	τ_{n,g_1}	0.46 (0.42, 0.47, 0.50)	Population Survey
	τ_{n,g_2}	0.76 (0.73, 0.76, 0.79)	Population Survey
	τ_{n,g_3}	0.43 (0.39, 0.46, 0.48)	Population Survey
Trade imbalance wedges	φ_n	−0.09 (−0.14, −0.00, 0.06)	Matched to regional trade deficits

Notes: This table reports the main parameters and calibrated objects in the benchmark economy. For varying objects, entries report the mean and the 25th, 50th, and 75th percentiles in parentheses, except for population shares. For population shares, the first number is the national population share of age group g, $s_g^N = \sum_n N_{n,g} / \sum_n \sum_{g'} N_{n,g'}$, while the values in parentheses report the 25th, 50th, and 75th percentiles of regional population shares, $s_{n,g}^N = N_{n,g} / \sum_{g'} N_{n,g'}$. Sectors a, m, and s denote agriculture, manufacturing, and services. For γ_n^{jk}, moments are reported by producing sector j, pooling over regions n and input sectors $k \in \{a, m, s\}$. Means for productivity, trade costs, and migration frictions are weighted by value added, trade flows, and migration flows, respectively; percentiles are unweighted. Fundamental productivity is normalized to Beijing. Trade-cost and migration-friction moments exclude own-region observations, $n = i$ and $n = m$, respectively. Trade elasticity, variety elasticity, migration elasticity, and the age-efficiency profile are taken from the 2005 calibration; all other calibrated moments use 2017 data, except that migration shares are taken from 2015 and used to approximate 2017 migration patterns.

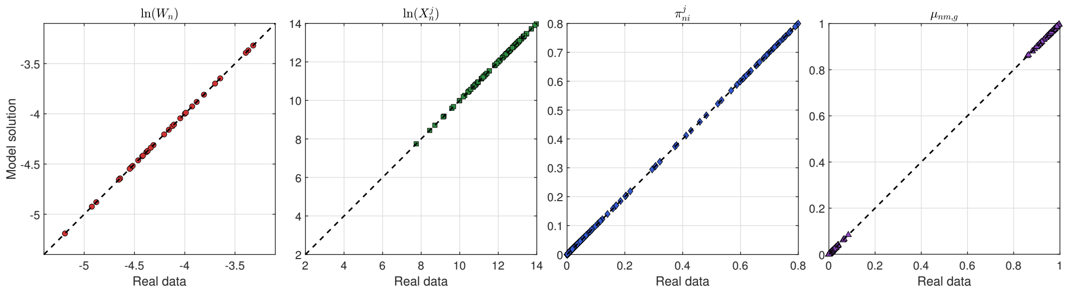

Figure 3: Model Fit

Notes: The scatter plots display data on the x-axis and model-generated values on the y-axis. The four panels report log wages, log sectoral expenditures, trade shares, and migration shares. The dashed line denotes the 45-degree reference line.

Panel A summarizes the time-invariant parameters. The age-specific migration elasticity κ^g and the age-efficiency profile l^g have been discussed above. The remaining entries describe the production and expenditure structure of the benchmark economy. Labor shares are highest in agriculture and lowest in manufacturing, while services lie in between. The input-output shares indicate substantial intermediate-input linkages across sectors. Final expenditure is concentrated in services, followed by manufacturing and then agriculture, which is consistent with the expenditure pattern in the data.

Panel B reports the recovered exogenous shocks and wedges. Fundamental productivity, reported as $(\lambda_n^j)^{1/\theta}$, shows stronger cross-region dispersion in agriculture than in manufacturing and services. Bilateral trade costs are also highest in agriculture, suggesting that agricultural goods face larger internal trade barriers, while manufacturing goods are relatively more tradable. Migration frictions rise with age, indicating that older individuals are less mobile across regions. The population entries reflect large differences in regional size after normalization to Beijing, especially because the ROW aggregate is included. Participation rates follow a standard life-cycle pattern: they are highest for prime-age workers and lower for young and older workers. The trade-imbalance wedge is centered around zero but varies across regions, capturing regional deficits and surpluses in the benchmark economy.

Figure 3 reports the model fit. The x-axis shows the data, while the y-axis shows the corresponding model-generated values. The four panels compare log wages, log sectoral expenditures, trade shares, and migration shares. The points lie close to the 45-degree reference line, indicating that the calibrated objects allow the model to reproduce the main patterns observed in the data.

Overall, the calibration performs well in matching the main patterns in the data. We next use the calibrated model to quantify the economic importance of the underlying forces.

5 Quantitative Analysis

This section quantifies the aggregate and spatial implications of three changes in China's economic environment: demographic ageing, domestic market integration, and external fragmentation. These changes operate through three distinct model margins: age-specific labor supply, internal market access, and external market access. The benchmark economy is the calibrated 2017 equilibrium.

We consider three counterfactual shocks. The first replaces the 2017 demographic structure with its 1997 counterpart. The second removes the improvement in domestic market integration between 1997 and 2017 by shifting internal trade costs and migration frictions back to their 1997 calibrated levels. The third shifts the external trade-cost block linking Chinese regions with the rest of the world back to its 1997 calibrated level. In each exercise, we change only the relevant block of fundamentals, keep the remaining fundamentals at their 2017 benchmark values, solve for the counterfactual equilibrium, and compare aggregate welfare and regional real-income dispersion with the benchmark.

Table 6 reports the empirical moments used to construct the counterfactual shocks. All hats denote calibrated 1997 values relative to calibrated 2017 values. The demographic moments imply that, relative to 2017, the population mass in 1997 was smaller for the young group and larger for the middle-age and older groups. The weighted means are 0.39, 1.40, and 1.75 for g_1, g_2, and g_3, respectively. Migration frictions were also higher in 1997: the flow-weighted mean is 1.95 when all age groups are pooled. Internal trade costs were higher as well, with weighted means of 1.69, 1.27, and 1.34 in agriculture, manufacturing, and services. External trade costs also fall between 1997 and 2017; shifting them back to 1997 implies weighted means of 1.93, 1.31, and 2.08 across the three sectors, with a pooled mean of 1.42.

5.1 Counterfactual Design

We construct the counterfactual shocks from two calibrated economies, 1997 and 2017, using the same model structure and calibration procedure. Throughout this section, a hat denotes the ratio of a calibrated 1997 object to its calibrated 2017 value. For a generic object z,

$$\widehat{z} \equiv \frac{z_{1997}}{z_{2017}}.$$

Applying $\widehat{z}$ to the 2017 benchmark replaces the relevant 2017 object with its 1997 calibrated counterpart along that margin.

Demographic ageing The demographic counterfactual removes the change in population structure between 1997 and 2017. We replace the 2017 distribution of population across regions and age groups with the corresponding 1997 distribution, while keeping the aggregate population scale fixed at its 2017 level. Age-specific participation rates are also set to their 1997 values. All other fundamentals remain at their 2017 benchmark values.

Let $N_{n,g,t}$ denote the population of age group g in region n at time t. The population shock is

$$\widehat{N}_{n,g} = \frac{N_{n,g,1997}}{N_{n,g,2017}}. \tag{61}$$

The counterfactual population is

$$N^A_{n,g} = \widehat{N}_{n,g} N^{2017}_{n,g}, \tag{62}$$

with an aggregate rescaling so that total population remains at the 2017 level. This isolates the effect of demographic change from productivity, trade-cost, and migration-friction changes.

Domestic market integration The domestic-integration counterfactual removes the improvement in internal market access between 1997 and 2017. Domestic integration operates through two bilateral friction blocks: internal trade costs and internal migration frictions.

For internal trade costs, the calibration delivers sector-specific bilateral trade costs $\kappa^j_{ni,t}$. For two different Chinese regions n and i, the internal trade-cost shock is

$$\widehat{\kappa}^j_{ni} = \frac{\kappa^j_{ni,1997}}{\kappa^j_{ni,2017}}, \qquad n, i \in \mathbb{N}_0, \quad n \neq i. \tag{63}$$

The counterfactual internal trade cost is therefore

$$\kappa^{D,j}_{ni} = \widehat{\kappa}^j_{ni} \kappa^j_{ni,2017} = \kappa^j_{ni,1997}, \qquad n \neq i, \tag{64}$$

while own-region trade costs remain normalized to one.

For migration frictions, the calibration delivers age-specific bilateral frictions $\nu_{nm,g,t}$. For age group g, the migration-friction shock is

$$\widehat{\nu}_{nm,g} = \frac{\nu_{nm,g,1997}}{\nu_{nm,g,2017}}, \qquad n, m \in \mathbb{N}_0, \quad n \neq m. \tag{65}$$

The counterfactual migration friction is

$$\nu^{D}_{nm,g} = \widehat{\nu}_{nm,g}\nu_{nm,g,2017} = \nu_{nm,g,1997}, \qquad n \neq m. \tag{66}$$

Thus, the domestic-integration counterfactual shifts the 2017 economy back to the 1997 internal trade and migration environment.

External fragmentation The external-fragmentation counterfactual shifts the trade-cost block between Chinese regions and the rest of the world back to its 1997 calibrated level. Region N denotes the rest of the world. For each sector j, the external block includes trade costs from Chinese regions to the rest of the world, κ^{j}_{Nn}, and from the rest of the world to Chinese regions, κ^{j}_{nN}, for $n \in \mathbb{N}_0$.

The external trade-cost shocks are

$$\widehat{\kappa}^{j}_{nN} = \frac{\kappa^{j}_{nN,1997}}{\kappa^{j}_{nN,2017}}, \qquad \widehat{\kappa}^{j}_{Nn} = \frac{\kappa^{j}_{Nn,1997}}{\kappa^{j}_{Nn,2017}}, \qquad n \in \mathbb{N}_0. \tag{67}$$

In this counterfactual, only these China–ROW and ROW–China trade-cost elements are changed. Demographics, internal trade costs, and migration frictions remain at their 2017 benchmark values unless the external-fragmentation shock is combined with other counterfactual shocks.

Table 6: Moments Used to Construct Counterfactual Shocks

Object	**Notation**	**Mean Ratio**	**p25, p50, p75 Ratios**
Panel A. Demographic ageing: 1997 relative to 2017			
Change in population shares	$\widehat{s}^{N}_{g_1}$	1.21	(1.75, 1.44, 1.50)
	$\widehat{s}^{N}_{g_2}$	0.97	(0.97, 1.00, 0.99)
	$\widehat{s}^{N}_{g_3}$	0.79	(0.64, 0.59, 0.55)
Panel B. Domestic integration: migration frictions, 1997 relative to 2017			
Change in migration frictions	$\widehat{\nu}_{nm,g_1}$	2.15	(1.56, 2.02, 5.97)
	$\widehat{\nu}_{nm,g_2}$	1.86	(1.40, 1.59, 3.21)
	$\widehat{\nu}_{nm,g_3}$	2.17	(2.86, 1.37, 1.00)
	$\widehat{\nu}_{nm,g}, \forall g$	1.95	(1.62, 2.99, 1.52)

Continued on next page

Table 6: Moments Used to Construct Counterfactual Shocks (continued)

Object	**Notation**	**Mean Ratio**	**p25, p50, p75 Ratios**
Panel C. Domestic integration: internal trade costs, 1997 relative to 2017			
Change in internal trade costs	$\widehat{\kappa}^{a}_{ni}$	1.69	(0.77, 0.28, 0.12)
	$\widehat{\kappa}^{m}_{ni}$	1.27	(1.29, 1.12, 0.93)
	$\widehat{\kappa}^{s}_{ni}$	1.34	(1.34, 1.16, 0.85)
	$\widehat{\kappa}^{j}_{ni}, \forall j$	1.26	(1.21, 0.95, 0.55)
Panel D. External fragmentation: external trade costs, 1997 relative to 2017			
Change in external trade costs	$\widehat{\kappa}^{a}_{nN}, \widehat{\kappa}^{a}_{Nn}$	1.93	(1.26, 1.43, 1.13)
	$\widehat{\kappa}^{m}_{nN}, \widehat{\kappa}^{m}_{Nn}$	1.31	(1.24, 0.96, 1.68)
	$\widehat{\kappa}^{s}_{nN}, \widehat{\kappa}^{s}_{Nn}$	2.08	(1.02, 0.84, 0.55)
	$\widehat{\kappa}^{j}_{nN}, \widehat{\kappa}^{j}_{Nn}, \forall j$	1.42	(1.20, 1.18, 1.10)

Notes: All entries are reported as 1997 values relative to 2017 values. Panel A reports changes in population shares rather than changes in population levels. The first number in each Panel A row is the ratio of the national age-group population share, $s^N_g = \sum_n N_{n,g} / \sum_n \sum_{g'} N_{n,g'}$, between 1997 and 2017. The values in parentheses in Panel A are the ratios of the 25th, 50th, and 75th percentiles of regional age-group shares, $s^N_{n,g} = N_{n,g} / \sum_{g'} N_{n,g'}$, between 1997 and 2017. For migration frictions, internal trade costs, and external trade costs, the first number is the ratio of the corresponding weighted means between 1997 and 2017; weights are migration flows, internal trade flows, and external trade flows, respectively. The values in parentheses report the ratios of the corresponding 25th, 50th, and 75th percentiles between 1997 and 2017. Domestic trade costs and migration frictions exclude own-region observations; external trade costs include China–ROW and ROW–China observations only. Migration frictions are capped at 100 before computing the reported moments.

Counterfactual cases Using these shocks, we consider three single-shock exercises:

$$A, \qquad D, \qquad F,$$

where A denotes the demographic-ageing counterfactual, D denotes the domestic-integration counterfactual, and F denotes the external-fragmentation counterfactual. These exercises measure the direct effect of each margin separately.

We then consider three pairwise interaction exercises:

$$A + D, \qquad A + F, \qquad D + F.$$

These exercises quantify whether the effect of one margin depends on the level of another margin. For example, the effect of domestic integration may differ under the 2002 demographic structure because the age composition changes both labor supply and migration responses. Similarly, the regional consequences of external fragmentation may depend on the degree of domestic market integration.

Finally, we implement all three shocks jointly:

$$A + D + F.$$

This final counterfactual combines changes in population structure, internal market access, and external market access in a single environment. Comparing this economy with the 2017 benchmark provides the overall quantitative effect of the three forces on aggregate welfare and spatial inequality.

5.2 Aggregate Effects and Spatial Inequality

Table 7: Counterfactual Results: Domestic and External Frictions

	Case 0	**Case 1A**	**Case 2A**	**Case 1B**	**Case 2B**
Set to 1997	None	Dom. MC	Dom. TC	Dom. MCTC	Ext. TC
RW_{pw}	1.0000	0.9756	0.9218	0.9057	0.9687
RW_{pew}	1.0000	0.9756	0.9218	0.9057	0.9687
$Gini^{Le}_{RW,pew}$	1.0000	0.9871	0.9974	1.0042	0.9536
$Gini^{Le}_{W,pew}$	1.0000	1.0035	1.0603	1.0785	0.9654
$Gini^{L}_{RW,pw}$	1.0000	0.9858	0.9970	1.0023	0.9535
$Gini^{L}_{W,pw}$	1.0000	1.0020	1.0601	1.0764	0.9654
	Case X	**Case 1AX**	**Case 2AX**	**Case 1BX**	**Case 2BX**
Set to 1997	Age dist.	Age dist. & Dom. MC	Age dist. & Dom. TC	Age dist. & Dom. MCTC	Age dist. & Ext. TC
RW_{pw}	1.0087	0.9856	0.9288	0.9134	0.9769
RW_{pew}	1.0046	0.9815	0.9250	0.9097	0.9729
$Gini^{Le}_{RW,pew}$	1.0014	0.9901	0.9941	1.0007	0.9546
$Gini^{Le}_{W,pew}$	0.9983	1.0025	1.0536	1.0708	0.9634
$Gini^{L}_{RW,pw}$	1.0007	0.9884	0.9931	0.9985	0.9541
$Gini^{L}_{W,pw}$	0.9975	1.0004	1.0528	1.0685	0.9628

Notes: All entries are normalized by Case 0. Each counterfactual sets the listed fundamentals to their 1997 calibrated values and keeps all other fundamentals at the 2017 benchmark. The upper block keeps the 2017 age distribution and changes only the listed friction blocks to their 1997 values. The lower block repeats the same exercises but also sets the age distribution to its 1997 value. "Age dist." denotes the population age distribution. "Dom. MC" denotes domestic migration costs, "Dom. TC" denotes domestic trade costs, "Dom. MCTC" denotes domestic migration costs and domestic trade costs jointly, and "Ext. TC" denotes external trade costs. RW_{pw} is aggregate real wage income per worker. RW_{pew} is aggregate real wage income per effective worker. $Gini^{Le}_{RW,pew}$ is the effective-labor-weighted Gini of real wage per efficiency unit across regions. $Gini^{Le}_{W,pew}$ is the effective-labor-weighted Gini of nominal wage per efficiency unit across regions. $Gini^{L}_{RW,pw}$ is the worker-weighted Gini of real wage income per worker across Chinese regions. $Gini^{L}_{W,pw}$ is the worker-weighted Gini of nominal wage income per worker across Chinese regions.

Table 8: Counterfactual Results: Combined Effects of Domestic and External Frictions

Set to 1997	Case 0 None	Case 1C Dom. MCTC & Ext. TC	Case 2C Dom. MC & Ext. TC	Case 3C Dom. TC & Ext. TC
RW_{pw}	1.0000	0.8711	0.9485	0.8842
RW_{pew}	1.0000	0.8711	0.9485	0.8842
$Gini^{Le}_{RW,pew}$	1.0000	0.9617	0.9438	0.9549
$Gini^{Le}_{W,pew}$	1.0000	1.0593	0.9691	1.0425
$Gini^{L}_{RW,pw}$	1.0000	0.9601	0.9426	0.9546
$Gini^{L}_{W,pw}$	1.0000	1.0573	0.9676	1.0420
Set to 1997	**Case X** Age dist.	**Case 1CX** Age dist. & Dom. MCTC & Ext. TC	**Case 2CX** Age dist. & Dom. MC & Ext. TC	**Case 3CX** Age dist. & Dom. TC & Ext. TC
RW_{pw}	1.0087	0.8780	0.9578	0.8905
RW_{pew}	1.0046	0.8744	0.9539	0.8868
$Gini^{Le}_{RW,pew}$	1.0014	0.9572	0.9460	0.9512
$Gini^{Le}_{W,pew}$	0.9983	1.0504	0.9672	1.0356
$Gini^{L}_{RW,pw}$	1.0007	0.9554	0.9445	0.9506
$Gini^{L}_{W,pw}$	0.9975	1.0484	0.9655	1.0348

Notes: All entries are normalized by Case 0. Each counterfactual sets the listed fundamentals to their 1997 calibrated values and keeps all other fundamentals at the 2017 benchmark. The upper block keeps the 2017 age distribution and changes only the listed friction blocks to their 1997 values. The lower block repeats the same exercises but also sets the age distribution to its 1997 value. "Age dist." denotes the population age distribution. "Dom. MC" denotes domestic migration costs, "Dom. TC" denotes domestic trade costs, "Dom. MCTC" denotes domestic migration costs and domestic trade costs jointly, and "Ext. TC" denotes external trade costs. $RW_{pw} = \sum_n W_n L^e_n / P_n \big/ \sum_n L_n$ is aggregate real wage income per worker. $RW_{pew} = \sum_n W_n L^e_n / P_n \big/ \sum_n L^e_n$ is aggregate real wage income per effective worker. $Gini^{Le}_{RW,pew}$ is the L^e_n-weighted Gini of W_n/P_n. $Gini^{Le}_{W,pew}$ is the L^e_n-weighted Gini of W_n. $Gini^{L}_{RW,pw}$ is the L_n-weighted Gini of $(W_n/P_n)L^e_n/L_n$. $Gini^{L}_{W,pw}$ is the L_n-weighted Gini of $W_n L^e_n / L_n$.

5.3 Provincial Effects and Spatial Reallocation

6 Conclusion

References

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

Anderson, James E. and Eric van Wincoop. 2003. “Gravity with Gravitas: A Solution to the Border Puzzle.” *American Economic Review* 93 (1):170–192.

Broda, Christian and David E. Weinstein. 2006. “Globalization and the Gains from Variety.” *Quarterly Journal of Economics* 121 (2):541–585.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro. 2019. “Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock.” *Econometrica* 87 (3):741–835.

Caliendo, Lorenzo and Fernando Parro. 2015. “Estimates of the Trade and Welfare Effects of NAFTA.” *Review of Economic Studies* 82 (1):1–44.

Department of Population and Employment Statistics, National Bureau of Statistics of China. 2016. *2015 nian quanguo 1% renkou chouyang diaocha ziliao*. Beijing: China Statistics Press. In Chinese: 20151%.

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

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. “The Next Generation of the Penn World Table.” *American Economic Review* 105 (10):3150–3182.

Groningen Growth and Development Centre. 2021. “Penn World Table Version 10.0.” Dataset. University of Groningen.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. “Commuting, Migration, and Local Employment Elasticities.” *American Economic Review* 108 (12):3855–3890.

OECD. 2021. “Inter-Country Input-Output Tables.” Dataset. OECD ICIO database.

Redding, Stephen J. and Esteban Rossi-Hansberg. 2017. “Quantitative Spatial Economics.” *Annual Review of Economics* 9 (1):21–58.

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

Tombe, Trevor and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109 (5):1843–1872.

Yamano, Norihiko, Joaquim Guilhoto, Ali Alsamawi, Colin Webb, Peter Horvát, Agnès Cimper, Carmen Zürcher, and Xue Han. 2021. "Development of the OECD Inter-Country Input-Output Database 2021." OECD Science, Technology and Industry Working Papers 2021/14, OECD Publishing.

Zhao, Q., C. Pan, Z. Zhang, X. C. Guo et al. 2024. "An Inter-Regional Input-Output Table Series of China from 1987–2017 with Integrated Carbon Emission Data." *Scientific Data* 11 (1):1335. URL https://www.nature.com/articles/s41597-024-04263-3.

Appendix

A Data

B Tables

Table B.1: Mainland Chinese Province-Level Administrative Regions

Code	**Region**	**Code**	**Region**	**Code**	**Region**	**Code**	**Region**
11	Beijing	31	Shanghai	42	Hubei	53	Yunnan
12	Tianjin	32	Jiangsu	43	Hunan	54	Tibet
13	Hebei	33	Zhejiang	44	Guangdong	61	Shaanxi
14	Shanxi	34	Anhui	45	Guangxi	62	Gansu
15	Inner Mongolia	35	Fujian	46	Hainan	63	Qinghai
21	Liaoning	36	Jiangxi	50	Chongqing	64	Ningxia
22	Jilin	37	Shandong	51	Sichuan	65	Xinjiang
23	Heilongjiang	41	Henan	52	Guizhou		

Notes: The table reports the 31 mainland Chinese province-level administrative regions in the benchmark calibration. Province-level administrative regions include provinces, autonomous regions, and municipalities directly under the central government.