Appendix For Online Publication

Skill-Biased Imports, Skill Acquisition, and Migration

Jingting Fan
Pennsylvania State University

Lei Li
University of Mannheim

Contents

A Data and Empirics

A.1 Summary Statistics and Additional Information

This appendix provides additional information and descriptive statistics for the main variables.

Summary Statistics. Table A.1 reports the summary statistics. To be consistent with our main specifications, all the statistics are weighted by city-level population in 2000. The average five-year import growth in capital goods per capita across the population is 0.7 (measured in units of 100 USD), with an interquartile range of 0 to 0.3, implying substantial skewness. The average five-year change in the share of people with a college education is 3.61 percentage points, with an interquartile range of 1.54 to 4.72 percentage points.

Table A.1: Summary Statistics

	mean	std	10th	25th	50th	75th	90th	N
Panel A: College share (%)								
College share, 2000	4.33	3.38	1.65	2.24	3.09	5.06	9.58	330
ΔCollege share, 00-05	2.20	1.88	0.43	0.98	1.76	2.97	4.59	330
ΔCollege share, 05-10	5.01	3.12	2.07	3.17	4.17	6.11	8.92	330
ΔCollege share, 00-10	3.61	2.93	0.70	1.54	3.05	4.72	7.24	660
(ΔCollege)/population, 00-10	3.90	3.82	0.56	1.42	2.84	4.88	8.93	660
Panel B: Capital goods imports (100 USD)								
Imported capital goods per capita (K), 2000	0.32	0.89	0.00	0.00	0.03	0.15	0.92	330
ΔImported capital goods per capita (ΔK), 00-05	0.74	2.66	0.00	0.01	0.04	0.23	1.28	330
ΔImported capital goods per capita (ΔK), 05-10	0.66	1.91	-0.03	0.00	0.06	0.33	1.73	330
ΔImported capital goods per capita (ΔK), 00-10	0.70	2.31	-0.02	0.00	0.05	0.30	1.46	660
ΔPredicted imported capital goods per capita (IV)	0.36	1.10	0.00	0.00	0.03	0.20	0.86	660

Note: The statistics are weighted by city-level residence-based population in 2000.

The importance of imports for capital formation. Figure A.1 plots China's total capital goods imports as a share of the aggregate investment in capital goods. The aggregate investment data are obtained from China Statistical Yearbooks. Because some of the investment is made in non-capital goods, such as buildings and structures, we adjust the aggregate investment using the machinery investment/aggregate investment ratio (33.7%) calculated from China's input-output table. According to this calculation, imported capital goods account for 46% of the aggregate investment in capital goods in 1998. This ratio reaches its peak in 2004 and then declines gradually afterward.

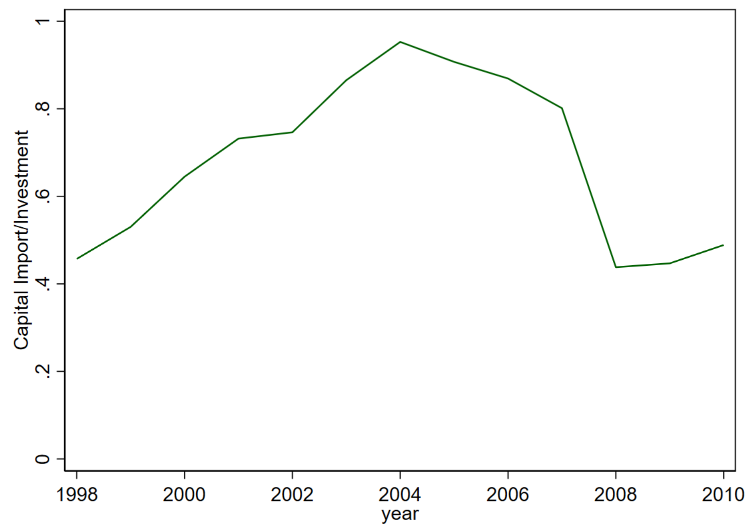

Figure A.1: Capital Goods Import/Investment

Note: See Section 2 of the text for the definition of capital goods import. The aggregate investment data are from China Statistical Yearbooks. We adjust the aggregate investment using the machinery investment/aggregate investment ratio (33.7%), calculated from China's input-output table in 2007. *Data Source*: China General Administration of Customs (1998-2010) and China Statistical Yearbooks

Example of HS-4-digit products. Table A.2 summarizes the number of HS-4-digit products within each HS-2-digit segment and gives an example for one such product.

There are a total of 16 HS-2-digit, each containing on average 10 HS-4-digit products. Take the 2-digit product 84, which has the largest number of HS 4-digit products, as an example. The descriptions for a typical product are, "Automatic data-processing machines and units thereof; magnetic or optical readers, machines for transcribing data onto data media in coded form and machines for processing such data, n.e.s. " (product code 8471). As we can see from the descriptions, products in the same HS-4-digit code tend to serve a common, specialized function, which corroborates our assumption of random growth shocks at the HS-4-digit level.

Table A.2: Examples of HS-4-Digit Products

	HS 2-digit	No. of HS 4-digit products	Example	Product Description
1	71	1	7115	Articles of precious metal or of metal clad with precious metal, n.e.s.
2	73	4	7309	Reservoirs, tanks, vats and similar containers, of iron or steel, for any material "other than compressed or liquefied gas", of a capacity of > 300 l, not fitted with mechanical or thermal equipment, whether or not lined or heat-insulated
3	76	3	7612	Casks, drums, cans, boxes and similar containers, incl. rigid or collapsible tubular containers, of aluminum, for any material (other than compressed or liquefied gas), of a capacity of <= 300
4	82	6	8201	Hand tools, the following: spades, shovels, mattocks, picks, hoes, forks and rakes, of base metal; axes, billhooks and similar hewing tools, of base metal; poultry shears, secateurs and pruners of any kind, of base metal; scythes, sickles, hay knives, hedge shears, timber wedges and other tools of a kind used in agriculture, horticulture or forestry, of base metal
5	83	2	8303	Armoured or reinforced safes, strongboxes and doors and safe deposit lockers for strongrooms, cash or deed boxes and the like, of base metal
6	84	76	8471	Automatic data-processing machines and units thereof; magnetic or optical readers, machines for transcribing data onto data media in coded form and machines for processing such data, n.e.s.
7	85	18	8542	Electronic integrated circuits as processors and controllers, whether or not combined with memories, converters, logic circuits, amplifiers, clock and timing circuits, or other circuits
8	86	8	8601	Rail locomotives powered from an external source of electricity or by electric accumulators
9	87	6	8701	Tractors
10	88	2	8801	Balloons and dirigibles, gliders, kites and other non-powered aircraft
11	89	5	8901	Cruise ships, excursion boats, ferry boats, cargo ships, barges, and similar vessels for the transport of persons or goods
12	90	23	9011	Optical microscopes, incl. those for photomicrography, cinephotomicrography or microprojection (excl. binocular microscopes for ophthalmology and instruments, appliances and machines of heading 9031)
13	91	2	9106	Time of day recording apparatus and apparatus for measuring, recording or otherwise indicating intervals of time, with clock or watch movement or with synchronous motor, e.g., time-registers and time recorders
14	94	1	9402	Medical, surgical, dental or veterinary furniture, e.g., operating tables, examination tables, hospital beds with mechanical fittings and dentists' chairs; barbers' chairs and similar chairs having rotating as well as both reclining and elevating movement; parts thereof
15	95	1	9508	Travelling circuses and travelling menageries; amusement park rides and water park amusements; fairground amusements, including shooting galleries; travelling theatres (e.g., motion simulators)
16	96	1	9618	Tailors' dummies and other lay figures, automata and other animated displays used for shop window dressing

Note: This table lists an example of a HS-4-digit product under each HS-2-digit category and reports the number of HS-4-digits products in each HS-2-digit category.

Age distribution for the enrollees in higher education institutions. Figure (A.2) depicts the distributions of college students in 2000 and 2010, calculated from the population census microdata. Most students are aged between 18 and 22 years old. The share of students aged above 30 is less than 2%.

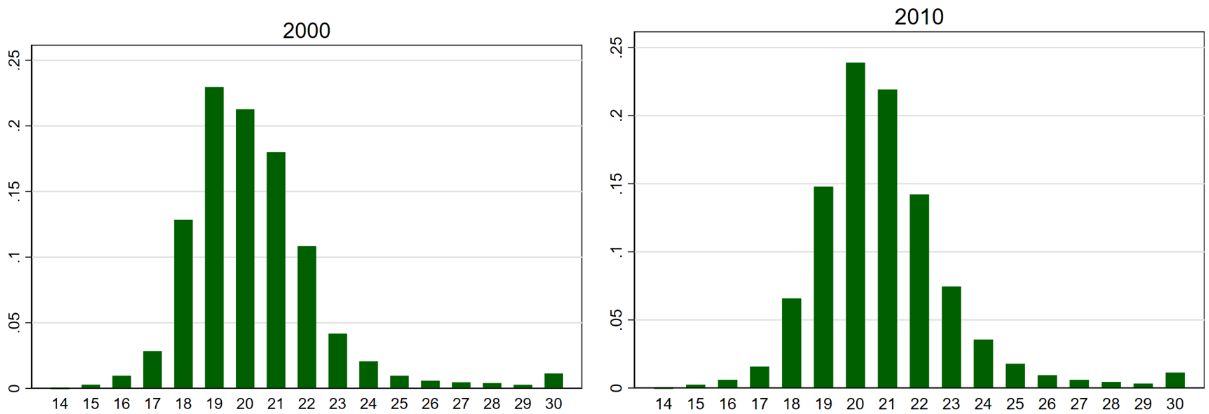

Figure A.2: Age Distribution of College Students

Note: People aged below 14 are merged with people aged 14 and people aged above 30 are merged with people aged 30. Data source: population census in 2000 and 2010.

A.2 Constructing City-to-City Migration Flows

In this subsection, we describe how we use census data to construct city-to-city migration flows. The construction is carried out in a separate data project described in Dorn and Li (2023). We summarize the main steps below.

Recall that our analysis uses two types of migration flows. For empirical analysis, we use five-year migration flows in the decomposition exercise to be consistent with the stacked five-year differences specification; for quantification, to be consistent with the model, where individuals make a one-time location choice, we use birth-to-current-location migration flows. Under both definitions, in the census microdata, we see individuals' origin province (i.e. the province five years ago and the birth province), but not the origin city. We use additional information to generate city-to-city migration flows

Specifically, for birth-city-to-current-location flows, we probabilistically allocate individuals to cities within their birth province based on each city's share of total births in their province among all local non-movers. For five-year migration flows, additional information on individuals' hukou registration locations and whether an individual's current and five-year-prior locations are the same as the hukou registration location is available. As hukou status is informative about recent moves, we will use these two pieces of information to improve the accuracy of origin city allocation for a subset of people.

To explain how we construct five-year city-to-city migration flows, we use the following definition. Let c_{it}^{-5} denote an individual i's location five years ago; h_{it} denote i's hukou registration location; for an individual whose current residence is not the same as the hukou city, we also observe whether he/she left hukou city five years ago or more recently (whether $l_{it} > 5$); finally, let p_{it}^{-5} be the province where i lived five years ago.

We categorize individuals into four categories, and in each case, we infer their location five years ago to minimize the required number of movements while ensuring consistency with variable aspects of micro- and city-level data. The first and most frequent case includes individuals whose geographic information does not indicate any past mobility. For such individuals, we set $c_{it}^{-5} = c_{it}$. The second case includes movers who left their hukou city less than five years ago. If the hukou city is in the individual's province of residence five years ago, we assume that these individuals lived in hukou city five years ago (i.e., $c_{it}^{-5} = h_{it}$). The third case is individuals who either left their hukou city more than five years ago, or they left the hukou city less than five years ago but the hukou city is not in the individual's province of residence five years ago. For these individuals, if hukou is in the current province of residence, we set by default that they lived in the current city five years ago.[1] The fourth case refers to all the remaining individuals that remain uncategorized after the first three steps. We probabilistically allocate them to cities of their resident provinces five years ago.

[1] If this assumption yields notable imbalances in the evolution of city populations within provinces, we adjust the assumption so a share of these individuals come from other cities in their current province during the last five years, with the share chosen to balance city population from five years ago.

A.3 Robustness: A Long-Difference Specification

In this subsection, we report the results from a long-difference (2000 to 2010) specification. The results are reported in Table A.3. It replicates the specifications in Table 1, with the only difference being here we use the 10-year difference instead of the stacked five-year differences. The points estimates are qualitative similar to that from Table 1, and all economically meaningful and statistically significant.

Table A.3: Capital Goods Import Growth and College Share Increase (10-Yr Difference Specification)

Dependent Variable: 100 × Δ (college share) (in % pts)

Panel A: Ten-Year Difference Specification					
	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV
Δ Capital goods import per capita	0.57***	1.03***	0.94**	3.92***	2.91***
	(0.16)	(0.15)	(0.44)	(1.14)	(0.93)
Province Dummies		√		√	√
Dummies for large ports		√		√	√
Start-of-period controls		√		√	√
Import per capita in 1998					√
Panel B: 2SLS First Stage Estimates					
Δ Predicted imported capital goods per capita			1.42***	0.92**	0.81**
			(0.21)	(0.39)	(0.36)
S.W. F statistics for the weak identification			45.3	5.5	5.2

Note: N=330. Regressions are weighted by city-level residence-based population in 2000. The start-of-period controls in 2000 include the minority share, population shares by cohorts (e.g., people born before 1940, 1941-1950, 1951-1960, 1961-1970, 1971-1980, 1981-1990, 1991-2000), manufacturing employment share, the export share of textiles, and the export share of electronics & machinery. Dummies for major port cities (Dongguan, Guangzhou, Haikou, Jiayuguan, Shenzhen, Suzhou, and Xiamen and Zhuhai) are controlled in Columns (2), (4), and (5). Column (5) further controls for the import per capita in 1998. Robust standard errors clustered at province are shown in parentheses, *** p<0.01, ** p<0.05, * p<0.1

A.4 Diagnostics on the Shift-Share IV

In this subsection, we present the results from the diagnostics tests for our identifying assumption, discussed in the text. These tests are reported in Figure (A.3), Figure (A.4), Table (A.4) and Table (A.5). See the text for descriptions.

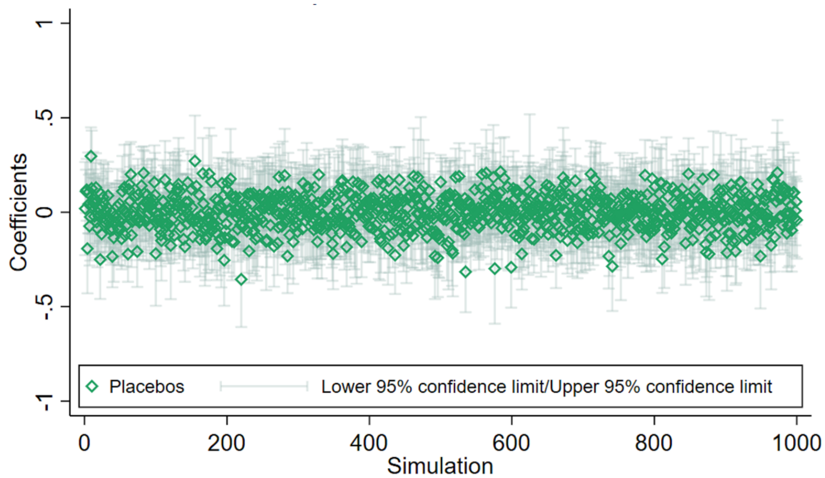

Figure A.3: Placebo import shocks

Note: This figure plots the coefficient on placebo shocks in 1,000 separate regressions following Adao et al. (2019). The dependent variable is the changes in college share, the same one as the dependent variable in Table 1. The unit of observation is a city. The placebo shock is defined as the interaction between capital goods import share in each city and a normally distributed random variable with mean 0 and variance 5. The controls are the same as in the main specification in Table 1.

Table A.4: Shift-Share Instrumental Variable: Shock Summary Statistics

	(1)	(2)
Mean	0.49	0
Standard deviation	0.58	0.53
Interquartile range	0.60	0.48
Residualizing on period FE		√
Effective sample size (1/HHI of s_{nt} wights)		
Across HS4 and periods	60.93	60.93
Across HS2 groups	1.93	1.93
Largest s_{nt} weight		
Across HS4 and periods	0.067	0.067
Across HS2 groups	0.69	0.69
Observation counts		
# of HS4-period shocks	318	318
# of HS4 groups	159	159
# of HS2 groups	16	16

Note: This table reports summary statistics for the import growth rates of HS-4-digit products. The first column uses the raw data; the second column uses period-fixed effects residualized data.

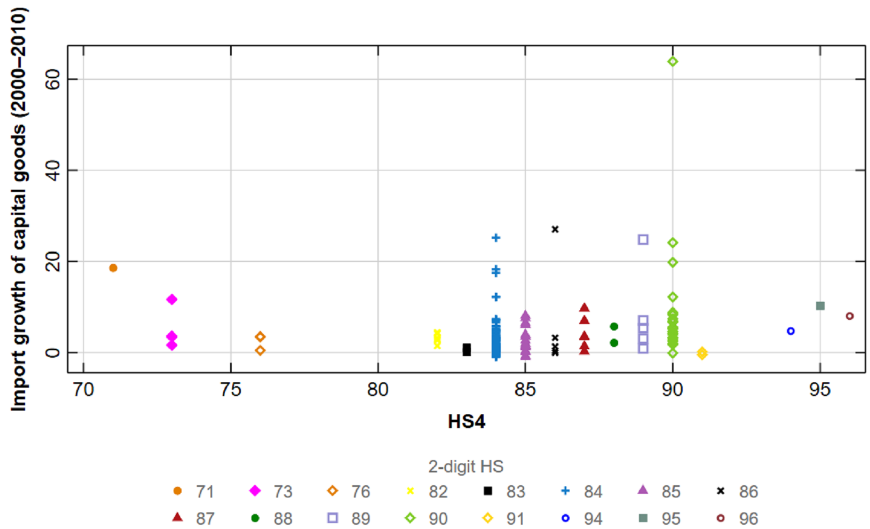

Figure A.4: Import Growth of Capital Goods by Products

Note: The plot displays the capital goods import growth by HS-4-digit product. Products within the same HS-2-digit product have the same color. 71 refers to precious metals, 72 refers to articles of iron or steel, 76 refers to tools of base metal, 83 refers to miscellaneous articles of base metal, 84 refers to machinery, 85 refers to electronics, 86 refers to railway and trainway, 87 refers to vehicles, 88 refers to aircraft or spacecraft, 89 refers to ships or floating structures, 90 refers to instruments or apparatus, 91 refers to clocks or watches, 94 refers to furniture, 95 refers to toys, and 96 refers to miscellaneous manufactured articles.

Table A.5: Shift-Share Instrument: Regional Shock Balance Tests

Regional balance variable	Coefficient	S.E.
Urbanization rate	0.173	0.105
Population share: 1980-1990	-0.017	0.039
Population share: 1970-1980	0.426	0.097
Population share: 1960-1970	-0.044	0.049
Population share: 1950-1960	-0.080	0.122
Population share: 1940-1950	-0.148	0.083
Minority share	-0.092	0.038
Manufacturing share	0.427	0.052

Note: N=660. This table reports the coefficients of unit-variate regressions between the IV and various confounding variables. To ease the comparison, all variables involved are normalized to have a variance of one and a mean of zero. Standard errors are clustered by province.

A.5 Suggestive Evidence on the Mechanism

Wage structure. We provide suggestive evidence on the underlying mechanism through which imported capital goods affect skill supply, focusing their impacts on the skill premium. We find that imported capital goods increase skill premiums and that this positive effect attenuates over time, consistent with skill supply responding to higher skill premiums.

We present the benchmark results of the 2SLS regressions in Table A.6. In both panels, we regress the log change in skill premium on the increase in per-capita capital goods imports, controlling for province-year fixed effect and a set of start-of-period controls. The skill premium is calculated using two methods. The first method is to calculate the wage difference between skilled workers and unskilled workers as a percentage of the wage of unskilled workers (Panel A). The second method is to estimate the Mincer-style OLS regression after we control for gender, working experience and its square term, employer ownership type, and industry dummies (Panel B). When taking the one-year first difference, we find that a city with a $100 rise in imported capital goods per capita increased 2.3 percentage points in the skill premium. The positive effects, however, decline over time. The gradual equalization of skill premium across regions is consistent with gradually increasing skill supply in regions more exposed to capital goods import growth.

Table A.6: Growth in Imported Capital Goods and Growth in Wage

Dependent variables	(1) 2003	(2) 2004	(3) 2005	(4) 2006	(5) 2007	(6) 2008	(7) 2009
Panel A: Δ (Skill premium: raw data)							
ΔCapital goods import per capita	0.023***	0.012***	0.004***	0.001***	−0.003	−0.000	−0.003
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Panel B Δ (Skill premium: Mincer)							
ΔCapital goods import per capita	0.008*	0.008***	0.003***	0.002***	0.000	0.001	−0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	181	181	181	181	180	180	180
S.W. F statistics for the weak identification	24.76	19.28	21.96	41.48	63.08	9.108	11.18

Note: We use the wage information from the Urban Household Survey (2002-2009). We choose the year 2002 as the base year and calculate the wage growth and skill premium growth between the year 2002 and year t, where t ranges from 2003 to 2009. Year 2002 is set as the base year because the number of cities covered in the UHS is 50% smaller for years before 2002. The unit of skill premium is 100%. Wages are measured in Chinse Yuan and deflated to the 1992 level. The sample sizes and first-stage results are the same for Panel A and B. All regressions control for province-year fixed effects and start-of-period characteristics and are weighted by city-level population in 2000. Standard errors clustered at province are shown in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Firm Evidence. We supplement our main analysis by documenting the relationship between the use of imported capital goods and a number of firm characteristics/choices, including labor productivity, employment structure, computer usage, and profit rate. We use the Annual Survey of Industrial Firms (ASIF), a large survey covering all state-owned enterprises and all large private enterprises in China. The ASIF data span a long period and allows us to control for time-invariant firm fixed effects.[2] Our regressions take the following form:

$$y_{it} = \beta_1 K_{it} + X_{it}\delta + \mu_i + \gamma_t + \varepsilon_{it}$$

where y_{it} refers to firm i's log wage, log labor productivity, the share of workers with a college education, number of computers per worker, and profit rate in year t, K_{it} is the ratio of imported capital goods over firm i's total imports, X_{it} is a set of firm-level controls, μ_i and γ_t are firm fixed effects and year fixed

[2]Most of the variables are reported by firms annually except for employment structure and computer usage, which are only reported in the census year (2004).

effects, respectively. For regressions in which the dependent variables are only available in 2004, the census year, we control for city-industry fixed effects instead (Appendix Table A.7 Columns 3-6).

We find that heavy capital goods importers pay higher wages (Column 1 of Table A.7). From 2000 to 2007, a one percentage point increase in capital goods import intensity is associated with a 2.5 percent increase in wages.[3] As suggested by Column (2), capital goods importers have higher labor productivity. In Column (3), we further explore the employment structure using the 2004 census data. With only one snapshot, instead of firm-fixed effects, we control for city-industry fixed effects and a number of firm characteristics. We find that firms with a higher capital goods import intensity have a higher share of skilled laborers (as measured by the share of workers with a college degree or above). Column (4) shows that capital goods importers use more computers, which are generally considered to be complementary to skilled workers. In Columns (5) and (6), we use two different ways to measure the profit rate and obtain the consistent finding that capital goods importers are more profitable. Although we cannot rule out the possibility of endogeneity in the OLS regressions, the results are consistent with imported capital goods being complementary to skills and contributing to an increase in the demand for skills.

Table A.7: Imported Capital Goods and Firm Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	100*Ln (Wage)	100*Ln (Value-added per Worker)	100*Share of Workers with College Degree	100*Computer per Worker	100*Profit /Sales	100*Operation Profit/Sales
	2000-2007	2000-2007	2004	2004	2004	2004
Imported Capital Goods/Imports	2.47***	6.25***	5.31***	6.28***	1.03***	0.93***
	(0.36)	(0.56)	(0.24)	(0.25)	(0.28)	(0.24)
Export/Sales	3.04***	1.41***	−1.89***	−0.75***	−0.63**	−0.56**
	(0.34)	(0.52)	(0.11)	(0.10)	(0.26)	(0.24)
Import/Inputs	4.66***	14.79***	3.07***	5.01***	0.12	0.21
	(0.86)	(1.39)	(0.37)	(0.37)	(0.74)	(0.67)
Foreign-owned firm indicator	2.66***	0.98*	4.18***	4.12***	0.46***	0.67***
	(0.39)	(0.59)	(0.11)	(0.10)	(0.13)	(0.12)
State-owned firm indicator	−1.59***	−5.63***	3.49***	1.29***	−4.32***	−4.62***
	(0.51)	(0.76)	(0.23)	(0.17)	(0.35)	(0.32)
Ln(Employment)	−14.79***	−42.33***	−0.88***	−1.56***	0.20***	0.11**
	(0.19)	(0.27)	(0.04)	(0.03)	(0.05)	(0.04)
City-Industry(4-digit) Fixed Effects			√	√	√	√
Firm Fixed Effects	√	√				
Year Fixed Effects	√	√				
Mean: Imported Capital Goods/Imports	0.03	0.03	0.03	0.03	0.03	0.03
Mean: Dependent Variable	246.90	401.29	11.48	7.48	3.48	3.34
Observations	1,482,241	1,482,241	216,932	216,932	216,995	216,995

Note: We use the data from the Survey of Industrial Production and the China General Administration of Customs. A skilled worker is defined as someone with a college degree or above. Imported capital goods intensity is defined as the share of imported capital goods out of capital stock. Reported standard errors are robust and are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1

[3] Because the national firm survey does not have wage data by education, we are unable to examine the relationship between capital goods imports and skill premiums.

B Theory

B.1 Definition of Equilibrium

We describe the conditions that characterize the competitive equilibrium.

Goods and labor market clear. Let X_i^s denote the value of the final good in sector s location i, which are aggregated from intermediate goods in the same sector by the local representative producer. As the sectoral final goods are non-tradable, their production must be equal to their local use—for consumption and for the production of intermediate goods in all sectors:

$$X_i^s = \underbrace{\alpha^s (w_i^H L_i^H + w_i^H L_i^H)}_{\text{consumption use}} + \underbrace{\beta_i^{K,s} \sum_j \pi_{ji}^K X_j^K + \beta_i^{NT,s} \sum_j \pi_{ji}^{NT} X_j^{NT} + \beta_i^{OT,s} \sum_j \pi_{ji}^{OT} X_j^{OT}}_{\text{production use by local intermediate goods producer}}, \tag{B.1}$$

where $\beta_i^{s',s}, s' \in \{K,,OT,NT\}$ are the share of sector s final goods in the value of sector s' intermediate goods, to be defined below. This condition implicitly imposes that the intermediate goods produced by location i are equal to the total demand for these intermediate goods.

Labor market clearing conditions are:

$$\begin{aligned} w_i^L \cdot L_i^L &= \beta_i^{K,L} \sum_j \pi_{ji}^K X_j^K + \beta_i^{NT,L} \sum_j \pi_{ji}^{NT} X_j^{NT} + \beta_i^{OT,L} \sum_j \pi_{ji}^{OT} X_j^{OT} \\ w_i^H \cdot L_i^H &= \beta_i^{K,H} \sum_j \pi_{ji}^K X_j^K + \beta_i^{NT,H} \sum_j \pi_{ji}^{NT} X_j^{NT} + \beta_i^{OT,H} \sum_j \pi_{ji}^{OT} X_j^{OT}, \end{aligned} \tag{B.2}$$

where $\beta_i^{s',H}, s' \in \{K,,OT,NT\}$ and $\beta_i^{s',L}, s' \in \{K,,OT,NT\}$ are the shares of high- and low-skill workers in the value of sector s' intermediate goods. These shares are given by

$$\begin{aligned} &\forall s \in \{K, OT, NT\} \\ &\beta_i^{s,OT} = \gamma_i^{s,OT} \\ &\beta_i^{s,NT} = \gamma_i^{s,NT} \\ &\beta_i^{s,K} = \gamma_i^{s,KHL} \cdot \frac{(1-\mu_i) \cdot (c_i^{KH})^{1-\rho_2}}{(1-\mu_i) \cdot (c_i^{KH})^{1-\rho_2} + \mu_i \cdot (w_i^L)^{1-\rho_2}} \cdot \frac{(\lambda_i)(p_i^K)^{1-\rho_1}}{(1-\lambda_i)(w_i^H)^{1-\rho_1} + (\lambda_i)(P_i^K)^{1-\rho_1}} \\ &\beta_i^{s,L} = \gamma_i^{s,KHL} \cdot \frac{\mu_i \cdot (w_i^L)^{1-\rho_2}}{(1-\mu_i) \cdot (c_i^{KH})^{1-\rho_2} + \mu_i \cdot (w_i^L)^{1-\rho_2}} \\ &\beta_i^{s,H} = \gamma_i^{s,KHL} \cdot \frac{(1-\mu_i) \cdot (c_i^{KH})^{1-\rho_2}}{(1-\mu_i) \cdot (c_i^{KH})^{1-\rho_2} + \mu_i \cdot (w_i^L)^{1-\rho_2}} \cdot \frac{(1-\lambda_i) \cdot (w_i^H)^{1-\rho_1}}{(1-\lambda_i)(w_i^H)^{1-\rho_1} + (\lambda_i)(P_i^K)^{1-\rho_1}}, \end{aligned} \tag{B.3}$$

where $c_i^{KH} = \left[(1-\lambda_i)(w_i^H)^{1-\rho_1} + (\lambda_i)(P_i^K)^{1-\rho_1}\right]^{\frac{1}{1-\rho_1}}$. Note that the share of K, H, and L are endogenous due to the CES production function in producing 'equipped labor', see equation (7).

Trade in intermediate goods.

$$\begin{aligned} \pi_{io}^s &= \frac{T_o^s (c_o^s \tau_{oi}^s)^{-\theta}}{\Phi_i^s}, \text{ where } \Phi_i^s \equiv \left[\sum_o T_o^s (c_o^s \tau_{oi}^s)^{-\theta}\right] \\ P_i^s &= [\Gamma(\frac{\theta + 1 - \sigma}{\theta})]^{\frac{1}{1-\sigma}} \cdot (\Phi_i^s)^{-\frac{1}{\theta}} \propto (\Phi_i^s)^{-\frac{1}{\theta}}. \end{aligned} \tag{B.4}$$

Household decisions. First, young people make optimal education choices given by

$$l_i^H = \frac{(u_{i,Y}^H/\delta_i)^\xi}{(u_{i,Y}^H/\delta_i)^\xi + (u_{i,Y}^L)^\xi}, \qquad l_i^L = \frac{(u_{i,Y}^L)^\xi}{(u_{i,Y}^H/\delta_i)^\xi + (u_{i,Y}^L)^\xi}, \tag{B.5}$$

with $u_{i,Y}^H$ and $u_{i,Y}^L$ defined by equation (12). Second, the fraction of individual age group a with skill e from i that chooses to migrate to d is given by

$$\lambda_{id,a}^e = \frac{\left(\frac{A_{d,a}^e \cdot w_d^e}{P_d \cdot \kappa_{id,a}^e}\right)^{\eta^e}}{\sum_{d=1}^N \left(\frac{A_{d,a}^e \cdot w_d^e}{P_d \cdot \kappa_{id,a}^e}\right)^{\eta^e}}, \tag{B.6}$$

with P_d defined by equation (10)

The supply of young and mature workers with skill e to location d is

$$L_{d,Y}^e = \sum_{i=1}^N \underline{L}_{i,Y} \cdot l_i^e \cdot \lambda_{id,Y}^e, \tag{B.7}$$

$$L_{d,M}^e = \sum_{i=1}^N \underline{L}_{i,M} \cdot \lambda_{id,M}^e$$

The total supply of workers in d is

$$L_d^e = L_{d,Y}^e + L_{d,M}^e. \tag{B.8}$$

Definition 1. *Given structural elasticities and fundamental parameters, the competitive equilibrium of the model is characterized by a set of prices* $\{w_i^L, w_i^L, P_i, P_i^s, c_i, c_i^{KH}, u_{i,a}^e,$ *etc.*$\}$ *and allocations* $\{X_i^s, \pi_{id}^s, \beta_i^{s,s'}, \lambda_{id,a}^e, l_i^e, L_{i,a}^e, L_i^e,$ *etc.*$\}$*, so that equations (B.1) to (B.8) hold.*

B.2 Proof of Proposition I

We derive first-order perturbation of the model around a baseline competitive equilibrium, characterized by equations (B.1) to (B.8), to obtain the change in equilibrium outcomes after a change in model fundamentals. We use variables with a bar (e.g., $\bar{w}_i^L$) to denote to indicate outcomes in the baseline equilibrium. We use a variable with a hat (e.g., $\hat{w}_i^L$) to denote the log change between the baseline equilibrium and a counterfactual equilibrium.

Deviation in prices and trade shares. We first derive changes in prices and trade shares as a function the of changes in wages. Differentiating equations (8), (9), and (10) , we obtain

$$\hat{P}_i = \alpha^K \hat{P}_i^K + \alpha^{NT} \hat{P}_i^{NT} + \alpha_i^{OT} \hat{P}_i^{OT}$$

and $\forall s \in \{K, OT, NT\}$ we have the following

$$\hat{P}_i^s = -\frac{1}{\theta} \cdot \hat{\Phi}_i^s = [\sum_o \bar{\pi}_{io}^s (-\frac{1}{\theta} \hat{T}_o^s + \hat{\tau}_{oi}^s + \hat{c}_o^s)]$$

$$\hat{c}_o^s = \beta_o^{s,OT} \cdot \hat{P}_o^{OT} + \beta_o^{s,NT} \cdot \hat{P}_o^{NT} + \bar{\beta}_o^{s,K} \cdot \hat{P}_o^K + \bar{\beta}_o^{s,H} \cdot \hat{w}_o^H + \bar{\beta}_o^{s,L} \cdot \hat{w}_o^L \tag{B.9}$$

$$\begin{aligned}
\hat{\pi}_{ij}^s &= \hat{T}_j^s - \theta \hat{\tau}_{ij}^s - \theta \hat{c}_j^s - \hat{\Phi}_i^s \\
&= \hat{T}_j^s - \theta \hat{\tau}_{ij}^s - \theta \hat{c}_j^s - [\sum_o \bar{\pi}_{io}^n (-\hat{T}_o^s + \theta \hat{\tau}_{oi}^s + \theta \hat{c}_o^s)] \\
&= \hat{T}_j^s - \theta \hat{\tau}_{ij}^s - \theta \hat{c}_j^s - [\sum_o \bar{\pi}_{io}^s (-\hat{T}_o^s + \theta \hat{\tau}_{oi}^s)] - [\sum_o \bar{\pi}_{io}^s (\theta \hat{c}_o^s)] \\
&= \hat{T}_j^s - \theta \hat{\tau}_{ij}^s - \theta [\beta_j^{s,OT} \cdot \hat{P}_j^{OT} + \beta_j^{s,NT} \cdot \hat{P}_j^{NT} + \bar{\beta}_j^{s,K} \cdot \hat{P}_j^K + \bar{\beta}_j^{s,H} \cdot \hat{w}_j^H + \bar{\beta}_j^{s,L} \cdot \hat{w}_j^L] - [\sum_o \bar{\pi}_{io}^s (-\hat{T}_o^s + \theta \hat{\tau}_{oi}^s)] \\
&\quad - \theta [\sum_o \bar{\pi}_{io}^s (\beta_o^{s,OT} \cdot \hat{P}_o^{OT} + \beta_o^{s,NT} \cdot \hat{P}_o^{NT} + \bar{\beta}_o^{s,K} \cdot \hat{P}_o^K + \bar{\beta}_o^{s,H} \cdot \hat{w}_o^H + \bar{\beta}_o^{s,L} \cdot \hat{w}_o^L)].
\end{aligned}$$

Deviation in market clearing conditions. Linearize the market clearing condition for final goods (B.1) to obtain

$$\begin{aligned}
&\forall s \in \{K, NT, OT\} \qquad \qquad \text{(B.10)} \\
\hat{X}_i^s =& \frac{\alpha^s \cdot \bar{w}_i^L \bar{L}_i^L}{\bar{X}_i^s} (\hat{w}_i^L + \hat{L}_i^L) + \frac{\alpha^s \cdot \bar{w}_i^H \bar{L}_i^H}{\bar{X}_i^s} (\hat{w}_i^H + \hat{L}_i^H) + \sum_{s' \in \{K,OT,NT\}} \sum_{j=1}^N \Big[\frac{\bar{\beta}_i^{s',s} \cdot \bar{\pi}_{ji}^{s'} \bar{X}_j^{s'}}{\bar{X}_i^s} (\hat{\pi}_{ji}^{s'} + \hat{X}_j^{s'}) \Big] \\
&+ \mathbf{I}(s == K) \sum_{s' \in \{K,OT,NT\}} \sum_{j=1}^N \Big[\frac{\bar{\pi}_{ji}^{s'} \bar{X}_j^{s'}}{\bar{X}_i^s} \cdot \hat{\beta}_i^{s',s} \Big],
\end{aligned}$$

in which $\hat{\beta}_i^{s,s'}$ is obtained by differentiating equation (B.3), which gives us

$$\begin{aligned}
&\forall s \in \{K, OT, NT\} \\
&\hat{\beta}_i^{s,OT} = 0 \\
&\hat{\beta}_i^{s,NT} = 0 \\
&\hat{\beta}_i^{s,L} = (1-\rho_2)\hat{w}_i^L - (1-\rho_2) \frac{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H} + \bar{\beta}_i^{s,L}} [\frac{\bar{\beta}_i^{s,K}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{P}_i^K + \frac{\bar{\beta}_i^{s,H}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{w}_i^H] - (1-\rho_2) \frac{\bar{\beta}_i^{s,L}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H} + \bar{\beta}_i^{s,L}} \hat{w}_i^L \\
&\hat{\beta}_i^{s,K} = (1-\rho_2)[1 - \frac{(\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H})}{\gamma_i^{s,KHL}}] \cdot [\frac{\bar{\beta}_i^{s,K}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{P}_i^K + \frac{\bar{\beta}_i^{s,H}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{w}_i^H] - (1-\rho_2) \frac{(\bar{\beta}_i^{s,L})}{\gamma_i^{s,KHL}} \hat{w}_i^L \\
&\qquad + (1-\rho_1)\hat{P}_i^K - (1-\rho_1) \frac{\bar{\beta}_i^{s,K}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{P}_i^K - (1-\rho_1) \frac{\bar{\beta}_i^{s,H}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{w}_i^H \qquad \qquad \text{(B.11)} \\
&\hat{\beta}_i^{s,H} = (1-\rho_2)[1 - \frac{(\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H})}{\gamma_i^{s,KHL}}] \cdot [\frac{\bar{\beta}_i^{s,K}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{P}_i^K + \frac{\bar{\beta}_i^{s,H}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{w}_i^H] - (1-\rho_2) \frac{(\bar{\beta}_i^{s,L})}{\gamma_i^{s,KHL}} \hat{w}_i^L \\
&\qquad + (1-\rho_1)\hat{w}_i^H - (1-\rho_1) \frac{\bar{\beta}_i^{s,K}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{P}_i^K - (1-\rho_1) \frac{\bar{\beta}_i^{s,H}}{\bar{\beta}_i^{s,K} + \bar{\beta}_i^{s,H}} \hat{w}_i^H.
\end{aligned}$$

Labor supply. Differentiating equations (B.7) and (B.8) gives us:

$$\hat{L}^e_d = \frac{\bar{L}^e_{d,Y}}{\bar{L}^e_{d,Y} + \bar{L}^e_{d,M}} \hat{L}^e_{d,Y} + \frac{\bar{L}^e_{d,M}}{\bar{L}^e_{d,Y} + \bar{L}^e_{d,M}} \cdot \hat{L}^e_{d,M} \tag{B.12}$$

$$\hat{L}^e_{d,M} = \frac{\underline{L}_{i,M} \cdot \bar{\lambda}^e_{id,M}}{\bar{L}^e_{d,M}} \cdot \hat{\lambda}^e_{id,M}$$

$$\hat{L}^e_{d,Y} = \frac{\underline{L}_{i,Y} \cdot \bar{l}^e_i \cdot \bar{\lambda}^e_{id,Y}}{\bar{L}^e_{d,Y}} \cdot [\hat{l}^e_i + \hat{\lambda}^e_{id,Y}],$$

where $\hat{l}^e_i$ and $\hat{\lambda}^e_{id,a}$, $e \in \{H, L\}, a \in \{Y, M\}$ are the changes in agents migration and education choice, obtained from differentiating equations (B.5), (B.6), and (12)

$$\hat{\lambda}^e_{id,a} = \eta^e(\hat{A}^e_{d,a} + \hat{w}^e_d - \hat{P}^e_d - \hat{\kappa}^e_{id,a}) - \eta^e \sum_j \bar{\lambda}^e_{ij,a}(\hat{A}^e_{j,a} + \hat{w}^e_j - \hat{P}^e_j - \hat{\kappa}^e_{ij,a}) \tag{B.13}$$

$$\hat{l}^H_i = \zeta \cdot (\hat{u}^H_{i,Y} - \hat{\delta}_i) - \zeta \cdot [\bar{l}^H_i \cdot (\hat{u}^H_{i,Y} - \hat{\delta}_i) + \bar{l}^L_i \cdot (\hat{u}^L_{i,Y})]$$

$$\hat{l}^L_i = \zeta \cdot (\hat{u}^L_{i,Y}) - \zeta \cdot [\bar{l}^H_i \cdot (\hat{u}^H_{i,Y} - \hat{\delta}_i) + \bar{l}^L_i \cdot (\hat{u}^L_{i,Y})]$$

$$\hat{u}^e_{i,a} = \sum_d \bar{\lambda}^e_{id,a} \cdot (\hat{A}^e_{d,a} + \hat{w}^e_d - \hat{P}^e_d - \hat{\kappa}^e_{id,a}).$$

Labor market clearing condition. Differentiate equations (B.2) to obtain

$$\forall e \in \{H, L\}$$

$$\hat{w}^e_i + \hat{L}^e_i = \sum_{s' \in \{K,OT,NT\}} \sum_{j=1}^{N} \Big[\frac{\bar{\beta}^{s',e}_i \cdot \bar{\pi}^{s'}_{ji} \bar{X}^{s'}_j}{\bar{X}^s_i} (\hat{\pi}^{s'}_{ji} + \hat{X}^{s'}_j) \Big] + \sum_{s' \in \{K,OT,NT\}} \sum_{j=1}^{N} \Big[\frac{\bar{\pi}^{s'}_{ji} \bar{X}^{s'}_j}{\bar{X}^s_i} \cdot \hat{\beta}^{s',e}_i \Big]. \tag{B.14}$$

Linearized system of equations. Equations (B.9) to (B.14) thus characterize the first-order changes in equilibrium outcomes in response to exogenous changes in model fundamentals, denoted $\{\hat{\kappa}^e_{ij,a}, \hat{A}^e_{d,a}, \hat{\delta}_i, \hat{T}^s_i, \hat{\tau}^s_{oi}\}$. The inputs to this system of equations are the changes in model fundamentals, the model's structural elasticities, and the equilibrium objects of the baseline equilibrium; the output of the system of equations are the changes in all endogenous variables of the model. Importantly, note that this system of equations is linear in both the changes in model fundamentals and the changes in the endogenous variables. Thus, it follows that the endogenous variables can be written as linear functions of the changes in fundamentals, with the weights being nonlinear functions of structural elasticities and the objects in the baseline equilibrium.

Although there is an analytical solution to these linear functions (equation (13) of the text), in quantification, we rely on conventional fixed point algorithms to solve the model. This eases computational burden as in our model with many cities and input-output linkages, constructing the matrices in equation (13) turns out to be computationally costly. Thus, instead of pursuing a full characterization of the linear functions, below we only sketch key steps needed for such a characterization.

Steps for analytical characterization of the first-order solution.

i Take the changes in wages $\{\hat{w}^H_i\}$ and $\{\hat{w}^L_i\}$ as given, equation (B.9) can be use to derive $\{\hat{\pi}^s_{ij}\}$ and $\{\hat{P}^s_i\}$ as linear function of $\{\hat{w}^H_i\}$ and $\{\hat{w}^L_i\}$ and the changes in model fundamentals.

ii Plug the output of step (i) into equations (B.12) and (B.13), and we can write $\{\hat{L}^e_i\}$ as a linear function of $\{\hat{w}^H_i\}$ and $\{\hat{w}^L_i\}$, as well as the changes in model fundamentals.

iii Plug the output from step [i] and [ii] into equation (B.10) to solve for $\{\hat{X}^s_i\}$ as a linear function of $\{\hat{w}^H_i\}$ and $\{\hat{w}^L_i\}$, as well as the changes in model fundamentals.

iv Plug all above into equation (B.14). By now all endgenous variables in (B.14) are linear function of $\{\hat{w}_i^H\}$, $\{\hat{w}_i^L\}$, and the changes in model fundamentals. It follows that the solution to equation (B.14) gives us $\{\hat{w}_i^H\}$ and $\{\hat{w}_i^L\}$ as linear functions of the changes in model fundamentals.

v Plug the output from (iv) to the output from step (i)-(iii) to express all endogenous variables as linear functions of the changes in model fundamentals.

B.3 Internal Estimation of Capital-Skill Complementarity

In the baseline calibration, we calibrate the parameters that govern the strength of capital-skill complementarity using the estimates of Krusell et al. (2000), which have been widely used and corroborated by studies in other settings (see, e.g., Burstein et al., 2013). In this appendix, we use a city-level wage panel, constructed from the UHS data, to estimate these parameters internally, exploiting variations across local labor markets in relative prices. It turns out that our own estimates imply a slightly weaker capital-skill complementarity, mostly because the substitution between capital and unskill labor is weaker. However, the quantitative implications remain similar.

Note first that equation (7) implies the following relationship between factor shares (the left-hand side) and factor prices and weight parameters (the right-hand side).

$$\frac{P_d^K q_d^K}{w_d^H L_d^H} = (\frac{P_d^K}{w_d^H})^{1-\rho_1} \cdot \frac{\lambda_{d,t}}{1-\lambda_{d,t}} \tag{B.15}$$

We allow $\lambda_{d,t}$ to differ across locations and over time. Take the log of the equation and then take the first difference by city d gives us

$$\Delta \ln(\frac{P_d^K q_d^K}{w_d^H L_d^H}) = (1-\rho_1) \cdot [\Delta \ln(P_d^K) - \Delta \ln(w_d^H)] + \underbrace{[\Delta \ln(\lambda_{d,t}) - \Delta \ln(1-\lambda_{d,t})]}_{\equiv \epsilon_{d,1}^K}. \tag{B.16}$$

The left-hand side is the relative shares of capital goods over high-skill labor. We construct the numerator of this ratio by aggregating from the firm-level data; we construct the denominator by first aggregating the firm-level wage bill and then multiplying it by the share of high-skill workers in the city's total wage bill that is calculated from the UHS. On the right-hand $\ln(P_d^K)$ in this specification is not observed, but equation (B.9) implies that up to the first order,

$$\begin{aligned} \Delta \ln(P_d^K) &= \sum_o \bar{\pi}_{do}^K(-\frac{1}{\theta}\hat{T}_o^K + \hat{\tau}_{od}^K + \hat{c}_o^K) \\ &= \bar{\pi}_{dN}^K(-\frac{1}{\theta}\hat{T}_N^K + \hat{\tau}_{Nd}^K + \hat{c}_N^K) + \sum_{o \neq N}(-\frac{1}{\theta}\hat{T}_o^K + \hat{\tau}_{od}^K + \hat{c}_o^K) \\ &\equiv \bar{\pi}_{dN}^K(-\frac{1}{\theta}\hat{T}_N^K + \hat{\tau}_{Nd}^K + \hat{c}_N^K) + \epsilon_{d,2}^K, \end{aligned} \tag{B.17}$$

where $\epsilon_{d,2}^K \equiv \sum_{o \neq N}(-\frac{1}{\theta}\hat{T}_o^K + \hat{\tau}_{od}^K + \hat{c}_o^K)$. This equation decomposes the change in the price of capital goods into two terms: the decrease in the price of foreign capital goods, which enters with a weight $\bar{\pi}_{dN}^K$, and changes in the price of capital goods purchased from domestic sources. Assuming that the increase in imported capital goods is driven by the decrease in the price of foreign capital goods (as assumed in the rest of our calibration), in partial equilibrium we can then proxy the term $(-\frac{1}{\theta}\hat{T}_o^K + \hat{\tau}_{od}^K + \hat{c}_o^K)$ using $-\frac{1}{\theta} \cdot \widehat{KIP}_i$, where $\widehat{KIP}_i$ is the percentage growth in capital goods imports. Plugging this proxy into

Table B.1: Estimating ρ_1 and ρ_2 Internally

	$1-\rho_1$		$1-\rho_2$	
Outcome: relative factor shares	(1)	(2)	(3)	(4)
Δ Relative price	0.565	0.575	-0.139	-0.175
	(0.118)	(0.17)	(0.07)	(0.08)
Year fixed effects	$\surd$	$\surd$	$\surd$	$\surd$
Province $\times$ year		$\surd$	$\surd$	$\surd$
Dummies for large ports $\times$ year		$\surd$		$\surd$

Note: Estimated using first-differenced yearly data over the period 2002-2009, over which we have the larger UHS sample that covers 180 cities. The regressions are weighted based on the 2000 city population. The first two columns estimate $1-\rho_1$, whereas the last two columns estimate $1-\rho_2$. Robust standard errors clustered at province are shown in parentheses. See Appendix section B.3 for the specification and the instrumental variable used.

equation (B.16) gives us

$$\Delta\ln\left(\frac{P_d^K q_d^K}{w_d^H L_d^H}\right) = (1-\rho_1)\cdot\underbrace{\left[-\frac{1}{\theta}\bar{\pi}_{dN}^K\cdot\widehat{KIP}_d - \Delta\ln(w_d^H)\right]}_{\text{the change in relative price}} + \underbrace{\epsilon_{d,1}^K+\epsilon_{d,2}^K}_{\text{structural residual}}. \tag{B.18}$$

Two endogeneity problem arises in the estimation of equation (B.18). First, local technology shocks $\epsilon_{d,1}^K$ might be correlated with the local relative price, especially through $\Delta\ln(w_d^H)$. Second, shocks to other domestic locations can be correlated with $\widehat{KIP}_d$. For example, if the reason location d imports more foreign capital is because nearby locations are becoming less productive at producing capital goods, then $\widehat{KIP}_d$ would be correlated with $\epsilon_{d,2}^K$.

We address the first concern through controls. Note that with a first-difference specification, we already remove the difference in the average *level* of $\lambda_{d,t}$ across locations. We will further control for province-year fixed effects, maintaining the assumption that the change in the capital bias parameter $\lambda_{d,t}$ is common across cities within a province. We address the second concern through instrumental variables. In particular, we instrument for the relative price in equation (B.18) using $[-\frac{1}{\theta}\bar{\pi}_{dN}^K\cdot\widehat{KIP}_d^{IV} - \Delta\ln(w_d^H)]$, where $\widehat{KIP}_d^{IV}$ is the capital goods import growth predicted by the shift-share design.[4]

The first two columns of Table B.1 report the estimates. Both columns use $[-\frac{1}{\theta}\bar{\pi}_{dN}^K\cdot\widehat{KIP}_i^{IV} - \Delta\ln(w_d^H)]$ as the instrumental variable. Our preferred specification is the second column, which implies $\rho_1 = 0.425$, i.e., capital goods are complementarity to high-skill labor. It is worth noting that the two columns give essentially the same estimate. This is reassuring as if the assumption stated in footnote (4) of this appendix is violated. We should expect the province-year fixed effect to pick up the correlation between $\Delta\ln(w_d^H)$ and $\epsilon_{d,2}^K$, thus changing the estimate substantially.

We derive an analogous estimation specification for ρ_2. Recall that the production function implies

$$\frac{c_d^{KH}q_d^{KH}}{w_d^L L_d^L} = \left(\frac{c_d^{KH}}{w_d^L}\right)^{1-\rho_2}\cdot\frac{1-\mu_{d,t}}{\mu_{d,t}}. \tag{B.19}$$

[4] For this regression to be valid, we impose that conditional on the controls, $\Delta\ln(w_d^H)$ is uncorrelated with not only $\epsilon_{d,1}^K$, but also $\epsilon_{d,2}^K$. This is a reasonable assumption given that we control for province-year fixed effects, which absorb non-foreign sources of $\hat{P}_i^K$ shock that could be correlated with the local skill wage.

Take the log of the equation and then take the first difference by city d gives us

$$\Delta \ln(\frac{c_d^{KH} q_d^{KH}}{w_d^L L_d^L}) = (1-\rho_2) \cdot [\Delta \ln(c_d^{KH}) - \Delta \ln(w_d^L)] + [\Delta \ln(1-\mu_{d,t}) - \Delta \ln(\mu_{d,t})]. \tag{B.20}$$

The left-hand side of this equation can be constructed using the firm-level survey and the UHS. On the right-hand side, c_d^{KH} is unobserved. As before, we approximate $\Delta \ln(c_d^{KH})$ using first-order approximation.

$$\Delta \ln(c_d^{KH}) = \frac{\bar{P}_d^K \bar{q}_d^K}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot \Delta \ln P_d^K + \frac{\bar{w}_d^H \bar{L}_d^H}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot \Delta \ln w_d^H,$$

where the letters with an upper bar correspond to variables in the baseline equilibrium. Combining this with equation (B.17) and (B.20), we arrive at the following approximation:

$$\begin{aligned}\Delta \ln(\frac{c_d^{KH} q_d^{KH}}{w_d^L L_d^L}) =& (1-\rho_2) \cdot \overbrace{\Big[\frac{\bar{P}_d^K \bar{q}_d^K}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot (-\frac{1}{\theta} \bar{\pi}_{dN}^K \cdot \widehat{KIP}_d) + \frac{\bar{w}_d^H \bar{L}_d^H}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot \Delta \ln w_d^H - \Delta \ln(w_d^L)\Big]}^{\text{the change in relative price}} \\ &+ \underbrace{\tilde{\epsilon}_{d,1}^K + \tilde{\epsilon}_{d,2}^K}_{\text{structural residual}},\end{aligned} \tag{B.21}$$

where $\tilde{\epsilon}_{d,1}^K$ is the change in the local factor bias in production $\Delta \ln(1-\mu_{d,t}) - \Delta \ln(\mu_{d,t})$; $\tilde{\epsilon}_{d,2}^K$ captures the variations in $\hat{P}_d^K$ that are due to domestic factors. We control for $\Delta \ln(1-\mu_{d,t}) - \Delta \ln(\mu_{d,t})$ using fixed effects; we account for the correlation between $\widehat{KIP}_d$ through an IV-strategy, using $\Big[\frac{\bar{P}_d^K \bar{q}_d^K}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot (-\frac{1}{\theta} \bar{\pi}_{dN}^K \cdot \widehat{KIP}_d^{IV}) + \frac{\bar{w}_d^H \bar{L}_d^H}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot \Delta \ln w_d^H - \Delta \ln(w_d^L)\Big]$ as an IV for the change in relative price in equation (B.21), maintaining that conditional on the fixed effects, $\Delta \ln w_d^H$ and $\Delta \ln(w_d^L)$ are uncorrelated with $\tilde{\epsilon}_{d,2}^K$.

Columns (3) and (4) of Table B.1 report the results. Our preferred specification, reported in Column (5), suggests that $\rho_2 = 1.175$, so low-skill labor is substitutable for high-skill labor and capital goods. Reassuringly, the two columns give essentially the same estimates again.

To summarize, using cross-city variations, we estimate $\rho_1 = 0.425$, $\rho_2 = 1.175$. Under these estimates, capital goods are complementarity to high-skill workers but substitutable to low-skill workers. These estimates are qualitatively in line with the estimates of Krusell et al. (2000) ($\rho_1 = 0.67; \rho_2 = 1.67$). In terms of level, both elasticities are lower in my setting. This could be due to the fact that my estimates are city-level elasticities, whereas Krusell et al. (2000) recover a macro-level elasticity, which captures not only the substitution within a city but also substitution between cities with different factor shares (see Oberfield and Raval, 2021 for a discussion of the relationship between micro and macro elasticities.) In the next subsection, we show that using our own estimates to carry out counterfactual analysis yields very similar findings.

B.4 Sensitivity to Alternative Parameterizations

In this subsection, we report the findings under three alternative sets of parameters. In the first exercise, we alter the migration elasticity parameters η^H and η^L; in the second exercise, we alter the responsiveness of education choice; in the third exercise, we use our internally estimated value for capital-skill complementarity—which turns out to be weaker than implied by the estimates of Krusell et al. (2000)—for analysis. These exercises suggest that the prediction of the model is robust across a reasonable range of parameters.

Higher migration elasticity. Table B.2 reports the results when we use higher migration elasticities. In particular, we set $\eta^H = 4$ and $\eta^L = 3$. These values fall between the calibrated value in this paper and the values in Fan (2019), who use the coefficient of variation of the earnings distribution to pin down the parameters governing migration elasticities.

Using these values, we find that the overall increase in skill supply due to capital goods import growth is similar to the baseline calibration. However, the increase is even more concentrated in the coastal region. The importance of various channels for the change in skill shares in the coastal and inland regions are also different. In particular, high-skill migration from inland to the coast now plays a more important role in driving the spatial disparities in skill.

Table B.2: Counterfactual Outcomes with *Higher Migration Elasticities*

	2000 demographics as basline		**2010 demographics as baseline**	
Panel A: overall skill acquisition				
baseline skill count	41.69 million		111.98 million	
skill share in population	5.63%		13.15%	
counterfactual skill count	45.27 million		120.81 million	
increase from baseline (%)	8.59%		7.89%	
Panel B: spatial distribution of skills				
	Coastal	**Inland**	**Coastal**	**Inland**
% increase in skill	25.89%	-1.18%	23.86%	-2.78%
% accounted for by				
Y, stayer	45.67%	-118.11%	47.72%	5.80%
Y, migrant	40.95%	70.43%	40.49%	26.54%
M, stayer	5.88%	82.69%	2.32%	54.34%
M, migrant	7.49%	64.99%	9.47%	13.32%

Notes: See notes under Table 5 for descriptions. Results here are calculated for the 'high-migration-elasticity scenario', with $\eta^H = 4$ and $\eta^L = 3$.

Higher education elasticity. Table B.3 reports the results when we use a larger education elasticity parameter ($\zeta = 4$). The overall increase in skill supply in response to the shock is larger than under the baseline calibration, as expected. The spatial disparities in skill share change in a similar way as under the baseline calibration. Overall, main finding from the baseline calibration remains robust.

Table B.3: Counterfactual Outcomes with a *Higher Education Elasticity*

	2000 demographics as basline		**2010 demographics as baseline**	
Panel A: overall skill acquisition				
baseline skill count	41.69 million		111.98 million	
skill share in population	5.63%		13.15%	
counterfactual skill count	46.41 million		123.51 million	
increase from baseline (%)	11.33%		10.29%	
Panel B: spatial distribution of skills				
	Coastal	**Inland**	**Coastal**	**Inland**
% increase in skill	22.32%	5.02%	20.18%	3.7%
% accounted for by				
Y, stayer	76.79%	104.70%	73.59%	107.96%
Y, migrant	18.96%	6.16%	21.48%	4.92%
M, stayer	1.76%	-6.36%	0.93%	-10.54%
M, migrant	2.49%	-4.51%	4.01%	-2.33%

Notes: See notes under Table 5 for descriptions. Results here are calculated for the 'high-education-elasticity scenario', with $\zeta = 4$.

Internally estimated capital-skill complementarity. In the last sensitivity analysis, we consider alternative parameters characterizing the strength of capital-skill complementarity. Recall that our baseline calibration uses the estimates of Krusell et al. (2000) for ρ_1 and ρ_2. We now explore the sensitivity of the findings to values of ρ_1 and ρ_2, using our own estimates of these parameters, $\rho_1 = 0.425$, $\rho_2 = 1.175$. See Section B.3 of this appendix for detail about the estimation.

Table B.4 reports the findings under this parameterization. Noting that in this parameterization, $\rho_2 - \rho_1 = 0.75$, which is smaller than $\rho_2 - \rho_1$ under the baseline calibration. Thus, this alternative calibration implies weaker capital-skill complementarity. In accordance with this observation, Table B.4 shows a slightly weaker effect of capital goods import growth on skill acquisition. The overall finding, and the effect of capital goods on spatial disparities and the importance of various margins for such disparities, remain similar.

Table B.4: Counterfactual Outcomes with *Weaker Capital Skill Complementarity*

	2000 demographics as basline		**2010 demographics as baseline**	
Panel A: overall skill acquisition				
baseline skill count	41.69 million		111.98 million	
skill share in population	5.63%		13.15%	
counterfactual skill count	44.93 million		119.90 million	
increase from baseline (%)	7.77%		7.06%	
Panel B: spatial distribution of skills				
	Coastal	**Inland**	**Coastal**	**Inland**
% increase in skill	16.03%	3.03%	14.39%	2.18%
% accounted for by				
Y, stayer	70.11%	116.31%	72.84%	121.56%
Y, migrant	23.88%	1.92%	22.08%	0.85%
M, stayer	2.58%	-10.41%	0.96%	-18.17%
M, migrant	3.43%	-7.82%	4.12%	-4.24%

Notes: See notes under Table 5 for descriptions. Results here are calculated for weaker capital-skill complementarity scenarios, which are estimated using city-level data, with $\rho_1 = 0.425$, $\rho_2 = 1.175$.

References

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-share Designs: Theory and Inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

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

Dorn, David and Lei Li, "Accounting for Population Flows in the Chinese Microcensus," *Work in progress*, 2023.

Fan, Jingting, "Internal Geography, Labor Mobility, and the Distributional Impacts of Trade," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–288.

Krusell, Per, Lee Ohanian, José víctor Ríos-rull, and Giovanni L Violante, "Capital Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 2000, *68* (5), 1029–1054.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.