Discussion Paper Series – CRC TR 224

Discussion Paper No. 189
Project B 06

Skill-Biased Imports, Skill Acquisition, and Migration

Jingting Fan [1]
Lei Li [2]

June 2023
(*First version: June 2020)*

[1] Pennsylvania State University, Email: jxf524@psu.edu
[2] University of Mannheim, Email: lei.li@uni-mannheim.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Skill-Biased Imports, Skill Acquisition, and Migration*

Jingting Fan †   Lei Li ‡

May 2022

Abstract

Imported capital goods, which embody skill-complementary technologies, can increase the supply of skills in developing countries. Focusing on China and using a shift-share design, we show that city-level capital goods import growth increases the local skill share and that both skill acquisition and migration play a role. We develop and quantify a spatial equilibrium model with these two mechanisms to examine the aggregate effects of capital goods imports, accounting for trade and migration linkages between cities. Counterfactual experiments suggest that the growth in capital goods imports in China between 2000 and 2010 led to a 3.7-8.9 million increase in the stock of college graduates, representing 5.7-13% of the total increase over this period. However, this growth disproportionately favored coastal regions, exacerbating existing spatial disparities.

JEL Classifications: F14, F16, F66, J24, J61

Keywords: Imported capital goods, capital-skill complementarity, skill acquisition, migration

*This paper subsumes an earlier working paper by Lei Li titled "Skill-Biased Imports, Skill Acquisition, and the Allocation of Talent". We are grateful to David Autor, Pietro Biroli, Lorenzo Casaburi, Gregory Crawford, Lachlan Deer, David Dorn, Harald Fadinger, Gordon Hanson, Liyan Han, David Hémous, Inga Heiland, Nir Jaimovich, Hyunseob Kim, Guilherme Lichand, Volker Nocke, Ralph Ossa, Dina Pomeranz, Hannes Schwandt, Jan Schymik, Tong Zhang, and seminar and workshop participants for their valuable comments and suggestions. Supports by the German Research Foundation (DFG) through CRC TR 224 (Project B06), the Swiss National Science Foundation (grant BSSGIO_155804), and Mannheim Centre for Competition and Innovation (MACCI) are gratefully acknowledged.

†Department of Economics, Pennsylvania State University, United States. E-mail: jxf524@psu.edu.

‡Department of Economics, University of Mannheim, Germany. E-mail: lei.li@uni-mannheim.de.

1 Introduction

Developing countries rely heavily on imports for their use of capital goods. As many of these goods, such as advanced computers and sophisticated production machinery, originate from developed countries and inherently embody technologies that exhibit a greater complementarity with high-skill labor than low-skill labor, their introduction into developing countries can trigger significant shifts in skill demand and increase wage disparities. Indeed, using empirical and quantitative methods, recent research has found that in many developing countries, the increase in capital goods imports led to an increase in the skill premium (e.g., Burstein, Cravino and Vogel, 2013; Parro, 2013; Raveh and Reshef, 2016; Fan, 2019; Koren and Csillag, 2019). In this paper, we build on this body of evidence and go one step further to study how—by increasing the demand for skill and pushing up the skill premium—imported capital goods change the pattern of skill acquisition and the geographic distribution of skills within a country.

Our rationale is intuitive. The use of imported capital goods increases the skill premium, which in turn motivates skill acquisition. Given that the usage is not uniformly distributed across regions, two implications arise. First, the increase in skill acquisition will vary across regions. Second, as workers can respond to the changing wage structure through migration, the uneven use of imported capital goods can lead to distinct migration patterns between high- and low-skill workers. With these mechanisms, the rise in capital goods imports observed in many developing countries in the recent past can influence not only the overall skill level in the importing country but also the spatial distribution of skills within it. Understanding such impacts is crucial given the fundamental role of human capital for aggregate and regional growth (Gennaioli, La Porta, Lopez-de Silanes and Shleifer, 2013).

In this paper, we make two contributions. First, we document causal evidence for the mechanisms discussed above, exploiting variations across locations in capital goods imports. These results, while informative about the existence of the mechanisms, do not map into general equilibrium effects. Our second contribution is to quantify a spatial equilibrium model to examine the impacts of these mechanisms on the aggregate and distribution of skills.

Our study focuses on China. The simultaneous increases in capital goods imports, the skill premium, and the number of college graduates in China over the past few decades make it a useful setting to test these mechanisms. As is well known, the Chinese economy experienced rapid integration with the global economy between 1978 and 2010. In particular, after Deng's 1992 Southern Tour, which reaffirmed the market reform policy, China's capital goods imports skyrocketed. This increase accelerated further in 2001 with China's WTO accession (Figure 1). Accompanying this surge were the increases in the college wage premium and the share of college graduates in the population (the college share).[1] The college wage premium rose from 14.3% in 1992 to 44.4% in 2009, as shown in Figure 2. Meanwhile, the college share quadrupled

[1] Throughout the paper, we define college shares based on people with any experience with college, including community college. We will use the terms skill premium and college wage premium interchangeably.

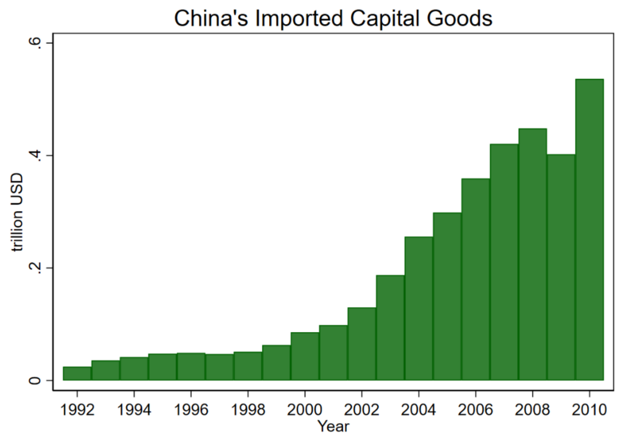


Figure 1 Capital Goods Imports in China

Note: The figure shows the pattern of China's capital goods imports (unit: 1 trillion US$). See Section 2 for the definition. Source of data: UN Comtrade.

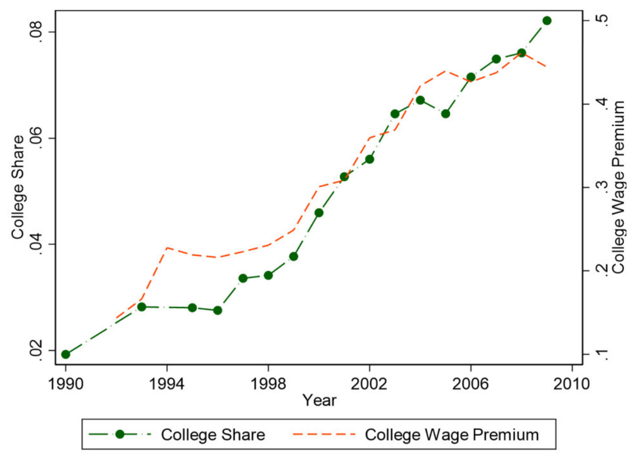


Figure 2 China's Rising College Share and College Wage Premium

Note: College share is the share of people above 15 years old with at least some college education. The college wage premium is the wage gap between people with at least some college education and those without, estimated using a Mincer-style OLS regression that controls for gender, working experience and its square term, employer ownership type, and industry fixed effects. Data: China Statistical Year Books (1990-2009) and the Urban Household Survey (1992-2009).

from 2% in 1990 to 8.2% in 2009, consistent with rising demand for skill driving both the college wage premium and the college share.

Our empirical analysis exploits the variations in the growth of capital goods imports across Chinese cities, drawing on several firm- and household-level datasets. In the data, cities with more rapid growth in capital goods imports over 2000-2010 experienced larger increases in college share. To establish a causal effect, we adopt a shift-share instrumental variable (IV) strategy, where we predict the growth in a city's capital goods imports using the inner product between the city's initial composition of imports across different capital goods and the (leave-one-out) national import growth rate in these goods over 2000-2010. In the IV regression, we view the national growth rate of the individual product import as the result of random shocks. We calculate a few diagnostic statistics following the recommendation of Borusyak, Hull and Jaravel (2022a). These statistics are generally supportive of the quasi-experimental national shock assumption. We also address the inference issues that arise in this setting (Adao, Kolesár and Morales, 2019).

We find that imported capital goods are an important force shaping the college share of Chinese cities. We estimate that between 2000 and 2010, an increase of 100 US dollars (USD) in a city's per-capita capital goods imports leads to a 1.23 percentage points increase in the city's college share. According to this estimate, capital goods imports can account for a quarter of the increase in college share and 12% of the dispersion in its increases across cities (measured by the interquartile range).

Our interpretation of the mechanism is that by pushing up the skill premium, capital goods imports incentivize skill acquisition and migration—a story of labor market demand. To provide suggestive evidence on this demand channel, we use firm- and city-level data to document the relationship between capital goods imports, wage, and skill employment. A plausible alternative mechanism is that, with imported technologies, cities become more productive, and local households and governments have more resources to invest in higher education—an income effect on skill acquisition. We account for this alternative mechanism by incorporating controls that speak directly to the income effects of households and city and provincial governments. The estimated effect of per capita capital goods import growth remains similar and statistically significant. Additional analysis shows that other explanations also cannot account for our findings.

We further examine the effect of capital goods imports on college share by age and by channel. We find that the college share among young workers responds much more strongly to capital goods imports than that among old workers. This is perhaps unsurprising, as young workers can adjust through skill acquisition and migration, but old workers respond mostly through migration. Using newly constructed data on city-to-city individual migration, we decompose the increase in the number of college graduates into three components: the number of college graduates among stayers, immigrants, and emigrants, respectively. The decomposition shows that while the estimate for migration is generally smaller than the estimate for skill acquisition, especially for young people—perhaps as a result of China's hukou restriction on migration—the role of migration is quantitatively important and statistically significant.

As is well recognized by the literature, in the presence of trade and, more importantly, migration linkages between cities, our reduced-form estimates do not recover the general equilibrium effect of capital goods imports. In the second part of the paper, we develop a simple spatial equilibrium model, with which we conduct counterfactual experiments. The model features capital-skill complementarity, endogenous skill acquisition, and costly trade and migration. To speak to the empirical patterns, we assume that there are two groups of people, young and mature. Young people decide whether to go to college and then make migration decisions; mature people also make migration decisions but treat their education as exogenous. In making these decisions, both groups consider the wage they will receive, which are determined in general equilibrium and can vary due to the cost of importing capital goods.

For transparency in estimation and counterfactual exercises, as well as to maintain a tight connection with the reduced-form exercises, we follow the Adao, Arkolakis and Esposito (2022) approach and consider a linearized model around the 2000 baseline equilibrium. Our estimation rationalizes the differential increases in capital goods imports across cities through heterogeneous reductions in importing costs. We then estimate the key elasticities of the model—those that govern workers' skill acquisition and migration decisions—using the same identifying assumption that is implicit in the reduced-form analysis.

We use the parameterized model for counterfactual experiments. We find that the decrease in capital goods import costs between 2000 and 2010 leads to between 3.7 million and 8.9 million increase in the stock of college graduates, accounting for between 5.7% and 13% of the 70 million increase in college graduates over this period—while not the primary driver of the increase in skill supply, the role of imported capital goods cannot be ignored either. Importantly, the increases in the college share are heterogeneous across locations. Dividing China into a coastal region and an interior region, we find that the coastal region experiences a 16%-18% increase in the college share. In contrast, the increase in the inland region is only modest (2-3%). Consistent with the reduced-form facts, we find that young workers play a much more important role in driving the skill divergence between regions. The stark contrast between these two regions highlights the importance of imported capital goods for not only skill acquisition but also skill distribution within a country.

This paper is related to a few strands of literature. First, an influential literature has documented the effect of skill-biased technological change on the wage premium of college workers (Katz and Murphy, 1992; Berman, Bound and Griliches, 1994; Goldin and Katz, 1998; Katz and Autor, 1999; Krusell, Ohanian, Ríos-rull and Violante, 2000; Acemoglu and Autor, 2011, 2012). An immediate implication of the findings in this literature for international trade—that importing capital goods embodying skill-biased technology can increase the wage premium in developing countries—has also been well studied. For example, Burstein et al. (2013), Parro (2013), and Fan (2019) incorporate capital-skill complementarity in trade models to study the impact of imported capital goods on the skill premium; Raveh and Reshef (2016), Li, Li and Ma (2022), and Koren and Csillag (2019) provide empirical support for this mechanism in China, Hungary, and other developing countries.[2] Our contribution relative to the existing literature is to test and quantify the importance of the implications of this mechanism on the acquisition and spatial distribution of skills within the importing country, along the way providing estimates for the key structural elasticities governing the response to such shocks.

Of course, we are not the first to study how trade affects either human capital accumulation or migration. For example, Atkin (2016), Blanchard and Olney (2017), and Li (2018) estimate the impacts of trade on human capital accumulation; Facchini, Liu, Mayda and Zhou (2019) estimate how trade affects migration in China. Relative to these works, which focus on either

[2] How trade affects wage inequality is the subject of a large literature. In addition to capital-skill complementarity, other channels that have been studied include (Feenstra and Hanson, 1996, 1999), skill upgrading (Bustos, 2011; Verhoogen, 2008), and more (see Goldberg and Pavcnik, 2007 for a review).

export opportunities or import competition, the main novelty of our analysis is that it focuses on a specific yet important type of trade—trade in capital goods. In addition, by combining causal estimates with a general equilibrium model, we quantify the aggregate effects of the mechanism.

Last but not least, in quantifying the mechanism, we build on recent development in spatial equilibrium models (Redding and Rossi-Hansberg, 2017). Our model is closest to that of Fan (2019), who himself builds on the models of Parro (2013) and Burstein et al. (2013). Our main departure from these existing models is that, motivated by the research question, we incorporate endogenous skill acquisition. Other recent works that consider skill acquisition to spatial equilibrium models include Ghose (2021) and Hsiao (2022).

The rest of the paper is organized as follows. Section 2 introduces the background and the data. Section 3 describes our empirical specification and reports the results. Sections 4 and 5 develop and quantify the model. Section 6 concludes.

2 Background and Data

In this section, we briefly describe the trade liberalization episode and the institutional background of the higher education sector in China. We then explain the data used in our analyses.

2.1 Background

China's capital goods imports surge. China had been isolated from the Western world since 1949, with its technology level lagging far behind. As part of the "reform and opening up" policy regime, the country started the "bring in advanced foreign technology" initiative in 1978. While technological transfer has been taking place in many forms, the importing of capital equipment that embodies advanced technologies has been particularly important for convenience and transparency. Hungry for advanced technology, the Chinese government gradually relaxed the constraints that impede such imports by cutting tariffs, removing import quotas, and abolishing trade licenses. These efforts accelerated in 1992, when Deng Xiaoping, through the Southern Tour, signaled to the country (inside and outside the party elite circle) and the world that China was going to continue the market-oriented reform (Vogel, 2011).

The result of these reforms is the surge in capital goods imports, shown in Figure 1. In making the figure and in the rest of this paper, we essentially define capital goods as durable machinery using the Broad Economic Classification (BEC). In particular, a Harmonized System (HS) product is defined as a capital good if it falls into BEC industry 41 (capital goods) or 521 (transportation equipment used for industrial purposes). As Figure 1 shows, between 1992 and 2000, capital goods imports grew by 16% per year; the growth accelerated to 20% in 2001 when China joined the WTO.

These imported capital goods play an important role in China's overall capital formation. Figure A.1 depicts the ratio of capital goods imports and the aggregate investment in equipment and machinery. This share increased from around 60% to the peak level of 90% in 2004, before it gradually decreased. The share stayed consistently above 40% between 2000 and 2010.[3]

Higher education in China. When the "reform and opening up" era began in 1978, the higher education sector in China was small and just beginning to recover from the disruptions caused by the Cultural Revolution. From 1978 through 1999, China saw a modest increase in the number of college students, with admissions rising from 0.4 million in 1978 to 1 million in 1998.

In part to meet the rising demand for skill and in part to alleviate the youth unemployment problem during a period of state-owned enterprises reform and bankruptcy, in 1998, the Chinese government launched a campaign to expand the higher education sector. Since then, college admission has been steadily increasing. According to the Education Statistical Yearbook of China, by 2010, the total college admission reached 6.6 million, marking an annual growth rate of 17% from 1998. About half of the increase took place through the establishment of new colleges, while the remainder occurred through the expansion of incumbent institutions.[4]

The implications of education policy for our study. Against the backdrop of this nationwide increase in college accessibility, we will investigate whether regions that use more imported capital goods see a larger increase in college shares. We will focus on the demand mechanism—rising demand for college graduates incentives more workers to acquire college education. A natural question is whether capital goods import growth can affect college shares through its impacts on the higher education policy. This concern is especially relevant, as Chinese universities are predominantly financed by the government.[5]

As the above discussion hopefully clarifies, the expansion of higher education is a nationwide policy. To the extent that capital goods import—by fueling the overall economic growth and boosting skill demand—is part of the reason for the expansion, given the important role of the provincial and central government in funding higher education institutions, most of such response should operate at the national or provincial levels, instead of at the city that our identification strategy is suitable for. Thus, we will not attempt to quantify the impacts of capital goods imports on the higher education policy. Instead, our investigation will focus on how, given the overall expansion in higher education, capital goods imports affect local college share through the demand channel.

Even with this focus, one might be concerned that if local governments' resources are affected by capital goods imports, then the increased higher education investment from the local government can partially explain our finding. In the empirical specification later, we will exploit the institutional setting of China's college admission and use additional controls to isolate the college seat supply mechanism.

[3]Our calculation does not adjust for quality. If foreign equipment is more advanced than domestic equipment, their true impact would be larger than implied by the ratio.

[4]Many of the newly established colleges are affiliated with incumbents as a way of circumventing red tape from the Department of Education regarding program offerings and size.

[5]According to China Educational Finance Statistical Yearbook, Chinese colleges at all tiers rely on governments, including central, provincial, and city governments, for more than half of their financing.

2.2 Sample and Data Sources

This subsection describes our regression sample and data sources.

Time horizon and regression sample. Our empirical analysis uses regional data from 2000-2010 to investigate the causal effects of capital goods imports on college shares. Following Topalova (2010), Autor et al. (2013), and Kovak (2013), we will adopt a local market approach. Our definition of local labor markets is prefecture cities. In the Chinese political hierarchy, prefectures are the administrative units between a province and a county. We choose prefectures as the units of analysis for both theoretical appeal and practical convenience. Specifically, prefectures are the closest counterpart to commuting zones often used in the studies of U.S. local labor markets, and they represent the most granular level at which some data are available. Moreover, within a prefecture, counties tend to be closely integrated economically through commuting ties and are governed by the same set of policies—many local government policies, such as those on hukou, land use, and investment, are made at the prefecture-city level.

There are 337 cities in mainland China in 2000, including four municipalities under the direct administration of the central government and 333 city-level prefectures. Without confusion, in the rest of this paper, we will refer to these units as cities. Over 2000-2010, there have been some changes in city boundaries through the reassignment of counties to cities. We account for such changes by constructing time-consistent county groups and matching cities across census years to the administrative division in 2000.[6]

Trade data. We construct prefecture-level capital goods imports from export and import transaction records provided by China's Customs Bureau, which covers 1998 to 2010. The data contain information not only on the product being traded but also on the location of the importing/exporting firm, which allows us to calculate import growth by city.

In accord with the aggregate patterns depicted in Figure 1, the summary statistics calculated from the microdata, reported in Appendix Table A.1, show considerable growth in per-capita capital goods imports during this period, with an average five-year (2000 to 2005, and 2005 to 2010) increase of 70 USD (for comparison, the average imports per capita in 2000 is 32 USD.) Importantly, the increase is highly heterogeneous across the population. For example, the people in the 90th percentile experience an average five-year increase of 146 USD, and the people in the 10th percentile experience a modest decrease.

We use product-level imports to implement our shift-share IV strategy. The raw data are at HS 8-digit level. As product codes change over time, we aggregate the data to the level of HS-4-digit products and calculate the national growth rate for these products as the "shifters". There are 159 HS-4-digit products across 16 HS-2-digit segments. To have a sense of what is a 4-digit product, Appendix A.2 lists the number of HS-4-digit products and gives an example of one such product within each HS-2-digit segment. As can be seen from the descriptions, an HS-4-digit code tends to contain products that serve a common, specialized function. This

[6]The construction of the county-level crosswalk is based on information on the administrative division changes published by the Ministry of Civil Affairs (www.mca.gov.cn/article/sj/xzqh//1980/).

corroborates our later assumption of quasi-random national shocks at the HS-4-digit level.

We construct the "shares" in the shift-share instrument using the city-level importing data provided by the Customs Bureau for 1998, the earliest year for which we could obtain the data.[7] We will fix the shares at this pre-period following Autor et al. (2013), but we note that under the maintained assumption introduced in the next section, our identification does not rely on the shares being exogenous.

Population census. We use the data from China's Population Census in 2000 and 2010 and the Population Survey in 2005 to measure regional college shares and migration. All these datasets are created by the National Bureau of Statistics of China. The census data cover the entire population, whereas the 2005 survey, also known as the Mini-Census, covers 1% of the population. The sample available to us covers 0.1%, 0.2%, and 0.3% of the Chinese population in the three years, respectively. We test the representativeness of these subsamples by comparing various city-level indicators calculated from the microdata to those reported in the official statistics based on the full sample. This test confirms that our samples are representative of the national data, with the exception of 7 cities in 2010, which we exclude in all analyses.[8]

We also rely on these three samples to construct migration flows, which we use in the quantitative exercises as well as in the empirical decomposition. There are two ways of measuring migration from our microdata: one based on the location of a person five years ago (through a retrospective question in the census/survey), and one based on the birthplace of the person. The challenge is that for both definitions, the census microdata reports only the origin province, not the origin city. We rely on the output of a new data harmonization effort by Dorn and Li (2023) that creates city-to-city migration flows under both definitions. In this data project, we use auxiliary information to map all individuals living in a province five years ago to cities in that province, ensuring the population statistics in origin cities are consistent with our microdata and other official statistics. For example, in defining previous-to-current-residence migration from 2000 to 2005, we map individuals to their previous resident cities in 2000 so that the number of residents in city d for a given cohort in 2000 estimated based on the 2005 data is consistent with the actual number of residents in city d reported in the 2000 census. When defining birth-to-current-residence migration flows, we attribute individuals to birth cities based on these cities' share of total births in their province. See the appendix for additional details.

Our micro data reveal that the growing national skill share depicted in Figure 2 went hand-in-hand with widening disparities between cities. As Table A.1 shows, the average five-year growth in college share for people aged above 15 years old over 2000-2010 is 3.61 p.p. (for comparison, the average college share in 2000 is 4.33%). For the people at the 90th percentile, the average five-year increase is 7.24 p.p.; for people at the 10th percentile, the increase is only 0.7 p.p.

Supplementary data. To provide auxiliary evidence for the mechanisms and for the construc-

[7]The data for 1998 is at HS-4 level, another reason why our shift-share instruments are built at this level.

[8]The seven cities include Cangzhou, Hengshui, Heze, Langfang, Liangshan, Zhoukou, and Zhumadian. They appear to be under-sampled in our 2010 microdata sample. These cities are all small and collectively account for less than 1% of China's total population.

tion of the variables used in quantification, we also use the following datasets.

We obtain information on wages from two sources. The first is the Urban Household Survey (UHS), available from 1992 to 2009. The survey provides detailed demographic, employment, and income information of urban residents. It is one of the most comprehensive household surveys in China and forms the basis for the wage and consumption information reported in the national statistical yearbooks. The full data cover 31 provinces; we have access to data from 18 of them. These 18 provinces are representative of the country in geographic location and economic development.[9] We will primarily use this dataset for measuring city-level wage and skill premiums to provide suggestive evidence for the mechanism. In addition, our quantitative analysis requires wage information for the entire country. We will use the 2005 Mini Census to supplement the UHS to construct this measure, a process we describe in quantification.[10]

To examine the relationship between imported capital goods use and firm skill intensity, we will use the 2004 industrial census, which covers the universe of industrial firms, and the Annual Survey of Industrial Firms (ASIF), which covers all state-owned enterprises (SOEs) and non-state firms in the mining, manufacturing, and public utility sectors with annual sales above 5 million Chinese yuan. The balance sheet and production information in these firm-level data allow us to construct the city-level use of capital and equipment, which we will use in model calibration.

In addition to the aforementioned datasets, we use various statistical yearbooks, such as China Education Statistical Yearbooks, to measure city characteristics.

2.3 A First Look from A Regional Perspective

Before proceeding to empirical analysis, we take a look at the cross-regional pattern in capital goods imports, the college share, and the skill premium.

Given the widespread spatial disparity in China in international trade participation, it is perhaps expected that the coastal region imports much more than the inland region. Figure 3 shows that coastal regions not only import more but also import more intensively in capital goods. For the decade starting in 2000, about 40% of the imports in the coastal region are in capital goods; in the inland region, this ratio is only around 30%. Figure 4 shows that over the same period, both the college share and the college premium increase faster in the coastal region. These spatial patterns align with our hypothesis that capital goods imports contribute to the increase in college share via capital skill complementarity. However, these patterns could also be influenced by confounding factors. In the next section, we will use a shift-share instrumental variable (IV) to estimate the causal effect of capital goods import growth on college shares.

[9] The 18 provinces include coastal provinces (Beijing, Guangdong, Jiangsu, Liaoning, Shandong, Shanghai, and Zhejiang) and inland provinces (Anhui, Chongqing, Gansu, Heilongjiang, Henan, Hubei, Jiangxi, Sichuan, Shaanxi, Shanxi, and Yunnan).

[10] Among the three population censuses/surveys, only the one in 2005 contains wage information.

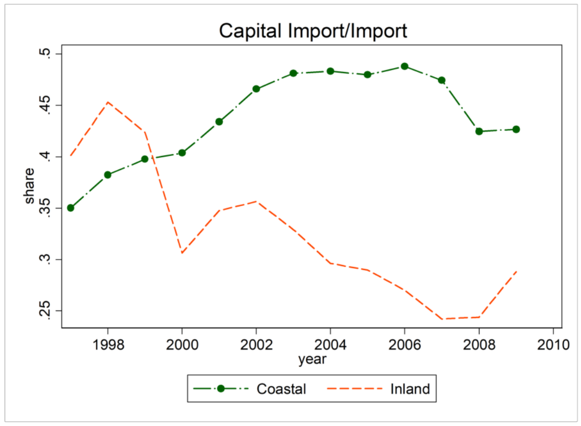


Figure 3 Capital Goods Imports in All Imports

Notes: Depicted in the figure are the share of capital goods imports in total imports for coastal and inland regions. *Data*: China General Administration of Customs, 1997-2009.

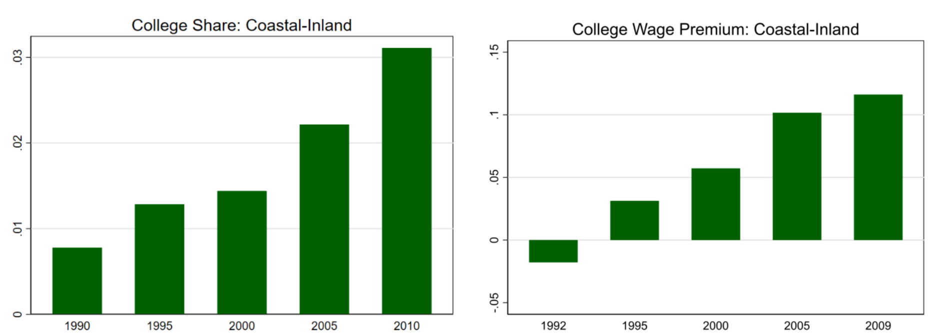


Figure 4 The Widening Regional Differences in College Share and College Wage Premium

Note: The college share is defined as the number of people with at least some college education as a share of the population. The college wage premium is estimated based on Mincer-style OLS regression after we control for gender, working experience and its square term, employer ownership type, and industry dummies.

3 Imported Capital Goods and College Share

3.1 Empirical Specification

Our regression specification is as follows:

$$\Delta Y_{it} = \beta_1 \Delta KIP_{it} + \alpha_{pt} + X_{it}^{'}\delta + \epsilon_{it}, \quad (1)$$

where ΔY_{it} is the change in regional outcome variables—such as the share of people with some college education—in city i between $t-1$ and t, and ΔKIP_{it} is the increase in per-capita capital goods imports in city i between $t-1$ and t. Having at hand the data for 2000, 2005, and 2010, we estimate equation (1) through stacked 5-year first differences, that is, the difference from 2000 to

2005 and from 2005 to 2010.[11] We weight cities by their start-of-period population shares.[12]

College share. Our primary outcome of interest is the change in the college share of a city. We count college graduates based on whether a person has any college education experience at all. As it usually takes four years to complete a college education, this definition allows us to correct the four-year lagged response of college education. Under this definition, we construct the outcome variables as follows:

$$\Delta Y_{it}^{g} = \frac{Skill_{it}^{g}}{L_{it}^{g}} - \frac{Skill_{it-1}^{g}}{L_{it-1}^{g}}, \tag{2}$$

where ΔY_{it}^{g} is the change between 2000 and 2005 or between 2005 and 2010 in the share of people of group g (certain birth cohort) in city i with some college education, $Skill_{i,t}^{g}$ is the number of people with college experience in city i in year t, and L_{it}^{g} is the number of people in age group g residing in city i in year t.

Capital goods import growth. Our primary variable of interest is per-capita capital goods import growth, defined as

$$\Delta KIP_{it} = \frac{M_{it} - M_{it-1}}{L_{it-1}} = \frac{\Delta M_{it}}{L_{it-1}}, \tag{3}$$

where ΔM_{it} is the change in capital goods imports in region i between $t-1$ and t, and L_{it-1} is the population in region i in year $t-1$.[13]

Control variables. Our first-difference specification removes time-invariant city-specific determinants of the outcome variables. To account for time-variant shocks, we include additional fixed effects and control variables. Specifically, to absorb variations across broad regions due to China's uneven spatial development and regional policies, we control for province $\times$ year dummies (α_{pt}). We accommodate the possibility that cities follow different time trends in both capital goods imports and college shares based on their (observable) features by controlling for a number of start-of-period city characteristics, including the minority population share, age structure (the fractions of people born in 1951-1960, 1961-1970, etc.), manufacturing employment share, and the shares of the cities' exports in the two major exporting industries of China between 2000 and 2010 (textile, and electronics & machinery).

Last but not least, we address potential biases in the measure of capital goods import growth arising from wholesale and transit trade. Recall that we map imports to cities using the location of the importing firm.[14] While for the goods considered in this paper, capital machinery, in most cases, the importer would be the user—or at least in the same city as the user—it is possible

[11] In Appendix A.3, we show that using 10-year-difference specification leads to similar findings. We choose the stacked five-year differences specification so that in the later decomposition, we can use the information on five-year migration from the census data.

[12] For regressions by birth cohorts, the weights are cities' start-of-period cohort population share.

[13] Given that some regions began with very low levels of imported capital goods, using size-normalized import increase as the explanatory variable is more appropriate than using the log change in imports.

[14] There are two locations associated with an import transaction. The first is the city where the goods clear the customs, the second is the location of the importer. We use the latter location.

that some firms act as a wholesaler and re-sell the imported goods to firms in other cities, creating non-classical measurement errors. To gauge the severity of such errors, we merge import transactions with the firm-level production data. We then calculate total imports as a share of a firm's total input use. Only 1% of the firms, most of them located in major ports, import more products than their reported input use. Although the data at hand does not allow us to directly correct for such biases, this exercise suggests that we can attempt to account for such biases by adding controls for 7 large port cities that showed exceptionally high levels of capital goods import growth.[15] Thus, in some specifications, we also control for the interaction between the indicators for these cities and period dummies.

3.2 Shift-Share IV, Diagnostics, and Inference

Despite the rich set of controls, the OLS estimate of equation (1) suffers from simultaneity and reverse causality biases. To establish a causal interpretation, we use a shift-share IV strategy (Bartik, 1991). We construct the IV as:

$$\Delta KIP_{it}^{Bartik} = \left[\sum_j \frac{M_{ij,1998}}{M_{i,1998}} \left(\frac{M_{jt}^{-i} - M_{jt-1}^{-i}}{M_{jt-1}^{-i}} \right) \right] \frac{M_{i,1998}}{L_{i,1998}}, \tag{4}$$

where M_{jt}^{-i} is China's import of capital goods in product j and year t, calculated excluding the province where city i locates, $M_{i,1998}/L_{it_{1998}}$ is the imported capital goods per capita of city i in the year 1998, and $M_{ij,1998}/M_{i,1998}$ is the share of an HS-4-digit product j in city i's capital goods imports in the year 1998, which captures a city's initial reliance on a certain type of capital equipment. As accurate city residential population is only available in census years, we proxy for the 1998 population using data from the 2000 census.

Equation (4) essentially predicts a city' per-capita capital goods import growth *rate* by combining its structure of imports across products in 1998 with the national growth in each product. As the predicted growth rate enters the IV in product with the per capita imports in 1998, we also include the interaction between the latter and period dummies to absorb the variations due to the pre-period per capita imports.

As illuminated by the recent methodological contributions on shift-share IVs, researchers can maintain one of the following two assumptions for this design to be valid: that the product-level import growth (the shifters) are generated by quasi-experimental shocks (Borusyak et al., 2022a), or that the shares in the pre-period are exogenous (Goldsmith-Pinkham et al., 2020). We will maintain the first assumption, viewing the shocks as generated by idiosyncratic foreign productivity growth/import cost reductions across small product categories.

[15] In the data, the top 7 cities in capital goods import growth import substantially more than the 8th city and all remaining cities. These cities are Dongguan, Guangzhou, Zhuhai, Shenzhen, Haikou, Suzhou, and Xiamen. The first four of these cities are in the Guangdong province and lie close to Hong Kong or Macao; Haikou is on the northern coast of the Hainan province; Suzhou is a city that borders Shanghai, and Xiamen is a city beside the Taiwan Strait. All 7 cities are international ports and also the earlier places to set up special economic zones.

Diagnostics. Following the recommendations of Borusyak et al. (2022a), we conduct three diagnostic tests on the validity of this assumption. First, we assess the extent to which the variations among HS-4-digit capital goods are between versus within HS-2-digit segments. Appendix Figure A.4 depicts the growth rate for each HS-4-digit capital goods category. The goods within the same HS-2-digit segment tend to occupy a wide range on the vertical axis, indicating high variability in import growth among similar products. This figure is supportive of the quasi-randomness assumption of the shifters. It also suggests that the shocks are unlikely to have a strong correlation with the technological/industrial characteristics of the products—if that were the case, we would have seen a strong correlation in growth within the same HS-2-digit segment.

Second, we calculate the expected HHI of average shock exposure (import share) to describe the effective sample size in Appendix Table A.4. The inverse HHI of the import shares is relatively high: 61 when exposure is measured at the HS-4-digit level, with the largest weight being 6.7%. Both statistics suggest we have a reasonably large effective sample size to exploit quasi-experimental variations in the shocks.

Third, we conduct regional balance tests in Table A.5. We regress city characteristics on the IV constructed according to equation (4). Given our focus on the change in college share as the outcome variable, cities' initial urbanization rate, age structure (i.e., the shares of the population born in 1941-1950, 1951-1960, 1961-1970, etc.), and the cultural value of education (proxied using minority share) are clearly relevant controls. We find that the IV is uncorrelated with most of these variables. On the other hand, we find the IV to be statistically significantly correlated with manufacturing employment share. Recall, however, that by design, our IV depends on per-capita imports in 1998, so this correlation could simply reflect the variations across cities in this variable. As we will include the initial manufacturing employment share and per-capita capital goods imports in our main specification, this correlation will not directly threaten our results. Overall, we find the test to be supportive of the assumption of quasi-random national shocks.

Inference. Under our identifying assumption, standard inference procedures can result in over-rejection due to correlation across cities in the shares (Adao et al., 2019). To assess the importance of this concern, we conduct a placebo analysis proposed in Adao et al. (2019) by running a regression using simulated shifters drawn randomly from a normal distribution with mean 0 and variance 5. Figure A.3 reports the results of this analysis. Interestingly—-perhaps because the inter-regional correlation in importing shares is relatively low—-conventional clustered (by province) standard errors turn out highly accurate.[16] The resulting placebo instruments' coefficients are significant at the 5% level in 5.7% of cases, and at the 1% level in 1.2% of cases. Given this finding, we will use the clustered standard error.

Table 1 Capital Goods Import Growth and College Share Increase

Dependent Variable: 100 × Δ (college share) (in % pts)					
Panel A: 2000-2010 Stacked First Differences					
	(1) OLS	(2) OLS	(3) IV	(4) IV	(5) IV
Δ Capital goods import per capita	0.32***	0.53***	0.47**	1.19***	1.23**
	(0.10)	(0.19)	(0.21)	(0.45)	(0.56)
Dummy for period 2000-2005	√		√		
Province × period		√		√	√
Dummies for large ports × period		√		√	√
Start-of-period controls		√		√	√
Import per capita in 1998 × period					√
Panel B: 2SLS First Stage Estimates					
Δ Predicted imported capital goods per capita			1.60***	0.82***	0.78***
			(0.21)	(0.28)	(0.22)
S.W. F statistics for the weak identification			56.5	8.9	12.1

Note: N=660. Regressions are weighted by city-level residence-based population in 2000. The start-of-period controls in 2000 and 2005 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) and their interactions with year are controlled in Columns (2), (4), and (5). Column (5) further controls for the import per capita in 1998 and its interaction with the period dummy. Robust standard errors clustered at province are shown in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

3.3 Baseline Results

Table 1 Panel A presents the baseline results. All regressions are weighted by the population in 2000. Columns (1) and (2) use the OLS specification. Both columns show a large and statistically significant correlation between the growth in capital goods imports and the change in college shares.

Columns (3) to (5) present the results from the IV specification, with the first-stage results reported in Panel B. Comparing the OLS estimate in Column (1) with the IV estimate in Column (3), we find that the coefficient of the IV estimation is slightly larger than that in the OLS estimation. This discrepancy can be explained either by attenuation bias or by households responding more to capital imports driven by national shifters than those driven by a city's idiosyncratic demand. In Column (4), when we include a richer set of controls, the coefficient increases to 1.19. Mechanically, this increase reflects the decrease in the first-stage regression coefficient with these controls.[17] In Column (5), we control for the interaction between each city's per-capita capital goods imports in 1998 and period dummies. The coefficient does not change materially.[18]

Column (5) is our preferred specification. The point estimate implies that a 100 USD increase

[16]The same capital goods are used in many sectors, and this might explain why the correlation in import shares can be lower than the correlation in employment shares in the setting considered in Adao et al. (2019).

[17]Upon closer inspection, we find that the main reason for the change in the first-stage regression is the terms involving large port dummies that control for biases in measurement due to transit trade.

[18]This indicates that most of the variations in 1998 per-capita capital goods imports are already absorbed by other controls, so they are not the main source of identification for our results.

Table 2 Alternative Mechanisms

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

	(1)	(2)	(3)	(4)	(5)	(6)
ΔCapital goods import per capita	1.10**	1.29***	1.60***	1.23**	1.63***	1.79***
	(0.49)	(0.46)	(0.54)	(0.55)	(0.56)	(0.52)
log(total admission at local colleges)	0.15***					0.14***
	(0.03)					(0.03)
Δln(GDP per capita)		-0.65				-0.58
		(1.08)				(0.92)
ΔImported inputs per capita			-2.15**			-2.18**
			(0.97)			(0.82)
ΔImport weighted tariff				1.12		1.97
				(1.04)		(1.31)
Employment share of state-owned enterprises					1.87***	2.35***
					(0.72)	(0.85)
Employment share of foreign-owned enterprises					−1.95	−1.07
					(1.53)	(1.81)
N	646	547	652	660	612	496
Dummy for period 2000-2005	√	√	√	√	√	√
Province ×period	√	√	√	√	√	√
Dummies for large ports ×period	√	√	√	√	√	√
Start-of-period controls	√	√	√	√	√	√
Import per capita in 1998 × period	√	√	√	√	√	√

Note: Regressions are weighted by city-level residence-based population in 2000. All columns are estimated using 2SLS. The start-of-period controls include the minority share, population shares by cohorts (e.g., people born before 1940, 1941-1950, 1951-1960, 1961-1970), manufacturing employment share, the export share of textiles, and the export share of electronics & machinery. The import per capita in 1998 and its interaction with the period dummy are also included. Dummies for major port cities (Dongguan, Guangzhou, Haikou, Jiayuguan, Shenzhen, Suzhou, and Xiamen and Zhuhai) and their interactions with year are controlled. Robust standard errors clustered at province are shown in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

in imported capital goods per capita leads to a 1.23 p.p. increase in local college share. Note that in the data, the interquartile range of five-year capital goods import growth and five-year change in college share are around 30 USD and 3.18 p.p., respectively. Our estimate implies that the interquartile range in the imported capital goods increase can account for 11.6% (1.23×0.3/3.18) of the interquartile range in the skill share increase. The average five-year changes for these two variables are, respectively, 70 USD and 3.61 p.p. If we were to use this estimate to calculate the aggregate effects of capital goods imports on college share, we would find that capital goods imports account for about a quarter of the college share increase (1.23×0.7/3.61).[19]

3.4 Suggestive Evidence and Alternative Mechanisms

Results in the previous subsection suggest that capital goods import growth has a large and statistically significant effect on city college share. Our preferred interpretation of the causal effect is that rising skill premiums motivate more workers to seek college education—a demand for skill channel. We now discuss some alternative explanations for the finding and present additional evidence to shed light on which one is more likely.

The first alternative explanation is that capital goods may enhance the productivity of cities,

[19]In our setting, where some of the estimated differential effects are due to migration, this back-of-envelop calculation can be biased. Indeed, this number turns out to be larger than the finding from the quantitative experiments.

bringing in additional revenue to local governments. As higher education in China relies heavily on government finance, more government investment in this sector can lead to an increase in the accessibility of college education, thereby increasing the local college share.

This is a very likely explanation for the *nationwide* increase in college share shown in Figure 2. However, given the institutional background of college admission in China, it is unlikely to be a major driver for our reduced-form estimate. In China, high school graduates take the National College Entrance Exam (commonly known as "gaokao"). Their scores in these exams determine what kind of college they qualify for. Importantly, as college seats are allocated to provinces, the qualification cutoffs for each college are also set at the provincial level. For example, students from Guangdong Province would be ranked within Guangdong to determine their eligibility and potential placement. Therefore, access to college varies largely by province, and these variations would be absorbed by our province×period fixed effects. As such, our estimates would not be affected by the concern that people *from* a province more exposed to capital goods have easier access to college and return to the home city after graduating from college.

We can not rule out that *within* a province, students from a city with more colleges are more likely to go to college simply because they can stay close to home while in college. Perhaps more important, students are more likely to go to a city for college (and stay after graduation) if that city has more seats available. To account for this channel, we control for the total admission by local colleges in 2000 on the basis of the specification in Table 1 Column (5).[20] The result is reported in the first column of Table 2. We find that the total admission by local colleges has a positive and statistically significant coefficient, but the coefficient for capital goods import growth does not change.

The second alternative explanation, conceptually similar to the first, is that with an influx of imported capital goods, local residents' income rises, thereby increasing the demand for higher education via an income effect. To account for this channel, we control for the change in log GDP per capita in the second column of Table 2. Our sample shrinks somewhat due to the availability of local GDP data, but the primary result remains the same. As local government resources are also tightly connected to the local GDP, this control also alleviates the concern that in response to an increase in government revenue, the local government increases investment in other local public goods—such as by increasing the spending on the local K-12 education—which improves the readiness of local students for college.

The third alternative explanation suggests that perhaps capital goods imports are complementary to other imported goods, and it is these other goods, rather than capital goods, that increase the college share. This complementarity could be direct—through production—or indirect; for example, cities importing more capital goods could lobby for lower tariffs on other goods

[20]Ideally, we would like to also control for the change in city-level admission, but we only have access to the city-level admission data in 2000. Using this as a control is still informative—as mentioned earlier, most of the increase in college admission is through either the expansion of incumbent institutions or the entry of new institutions that are affiliated with existing institutions. For this reason, initial college admission is likely to be highly correlated with the increase in college seats between 2000 and 2010.

they use. In Column (3), we control for the change in total imports per capita.[21] In Column (4), we control for the city-level average import tariffs. These controls do not diminish the estimate for capital goods imports.

Finally, during the decade 2000-2010, China carried out major reforms in the state-owned enterprises sector and attracted many foreign enterprises. If—for reasons unrelated to capital-skill complementarity—these firms have a higher demand for both capital goods and skills, then the increase in capital goods imports can increase local skill demand in cities where SOEs and foreign firms have a bigger presence. To account for this mechanism, in Column (5), we control city-level employment shares of state-owned and foreign firms. We find that the SOE employment share is positively correlated with the college share, but if any, the coefficient for capital goods imports becomes larger with these controls.

In the last column, we include all these controls simultaneously. After accounting for alternative explanations, the results suggest that capital goods imports remain strongly correlated with the college share, thereby supporting our proposed mechanism.

These exercises show that alternative stories are not the main driving force of our findings. In the appendix, we present two pieces of suggestive evidence on the capital-skill complementarity mechanism. First, using the UHS panel, we explore the impact of imported capital goods on the wage structure. We find that imported capital goods increase the skill premium, consistent with the findings of the literature. Additionally, in line with a gradual increase in college supply, the increases in skill premiums attenuate over time. Second, we use firm-level data to examine how a firm's production and demand for skill correlate with capital goods imports. We find that firms using imported capital are more productive, have a more skilled workforce, pay higher wages, and maintain a higher computer-worker ratio. These patterns are consistent with an increase in skill demand in response to an increase in imported capital goods.

3.5 Decomposition: Education vs. Migration

To inform the quantitative framework, we examine the response across birth cohorts and by adjustment margin (skill acquisition versus migration). To facilitate decomposition, we consider the following outcome variable:

$$
\begin{aligned}
\Delta Y_{it}^{'g} &= \frac{Skill_{it}^{g} - Skill_{it-1}^{g}}{L_{it-1}^{g}} \\
&= \frac{Skill\,(Unskill_{it-1})_{it}^{g} - Skill\,(Dead)_{it}^{g}}{L_{it-1}^{g}} + \frac{Skill\,(IM)_{it}^{g}}{L_{it-1}^{g}} - \frac{Skill\,(EM)_{it}^{g}}{L_{it-1}^{g}}
\end{aligned} \tag{5}
$$

where $\Delta Y_{it}^{'g}$ is the change in the number of people in cohort g with a college education between $t-1$ and t over the population of the people in cohort g in $t-1$.

[21] The shift-share instruments for non-capital goods imports are constructed similarly to the IV for the capital goods imports in equation (4).

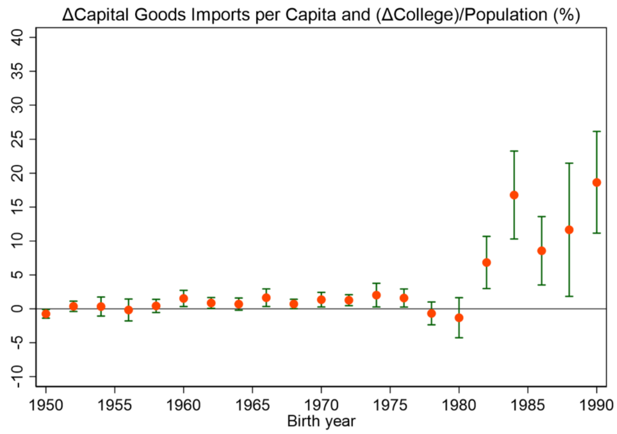


Figure 5 Imported Capital Goods and College Share: by Birth Cohort
Note: Coefficients from regression by cohort, with outcome variables defined by equation (5). Each cohort group includes people born in two adjacent years. Regression specification is the same as Column (5) of Table 1.

This outcome variable can be conveniently decomposed into three margins of adjustment: the change in the number of people with a college education among the stayers (the first term in the second line, which, in turn, can be decomposed into skill acquisition and death), among the immigrants (the second term), and among the emigrants (the third term). We measure these terms using the newly constructed city-to-city, five-year migration flows, defined based on residents' current location in time t and their residence five years ago $(t-1)$. As our data are repeated cross-sectional, we have no information on total death by skilled stayers, so we report the sum of the two terms for stayers.

Figure 5 reports the estimation results for the outcome variables defined in equation (5). The coefficients are positive and statistically significant for most cohorts born after 1965, but they are substantially larger for people born after 1980. This finding is intuitively sensible, given that most people attend college in China before they turn 30.[22] It is precisely this group of people that are likely to respond to capital goods import growth through *both* skill acquisition *and* migration—while the remaining groups respond primarily through migration.[23]

In Table 3, we decompose the estimates into the three components from equation (5) for a few relatively young cohorts, for whom the differential death rates by skill are unlikely to be significant. Except for the 1978-1981 cohort, the coefficients for both stayers and immigration are positive and mostly statistically significant, whereas the coefficients for emigration are generally small and insignificant.[24] To interpret these coefficients, note that the difference between Column

[22]See Appendix A.2 for the age distribution of students in higher education institutions, including colleges and postgraduate schools. It demonstrates that people older than 30 years old account for less than 2% of the enrollments. This implies that the response of skill supply is mainly due to the entry of new cohorts who receive higher education rather than skill upgrades by mature workers who are already in the labor market.

[23]This finding is also reassuring—if regional capital import shocks are correlated with other regional fundamentals that determine the higher-education choice and if such fundamentals are persistent, then one should expect that such shocks be correlated with education choice in the past. Of course, much has changed in China between when the old cohorts make college education choices and our sample period, so this check is best viewed as a sanity check.

[24]There might be a cohort-specific shock, but we do not have an explanation for what it might be.

Table 3 Decomposition by Birth Cohort

(1) = (2) + (3) − (4)	(1) Y = 100 × (no. people with some college or above) /population		(2) Local		(3) Immigration		(4) Emigration
		=		+		−	
Panel A. 1986-1989							
ΔCapital goods import per capita	13.51**		9.34**		3.22**		−0.95**
	(5.44)		(4.32)		(1.42)		(0.39)
Panel B. 1982-1985							
ΔCapital goods import per capita	10.14***		6.89***		2.82**		−0.43
	(2.38)		(1.82)		(1.22)		(0.61)
Panel C. 1978-1981							
ΔCapital goods import per capita	-0.19		−1.85		1.38***		−0.28
	(1.89)		(2.01)		(0.39)		(0.39)
Panel D. 1974-1977							
ΔCapital goods import per capita	2.23**		0.85		1.35***		-0.03
	(0.92)		(1.07)		(0.29)		(0.28)

Note: N=660. Regressions are weighted by city-level population in 2000. Regression specification is the same as Column (5) of Table 1. Robust standard errors clustered at province are shown in parentheses, *** p<0.01, ** p<0.05, * p<0.1

(2) and Column (4) tells us how many people residing in a city exposed to the shock in the previous period responded through skill acquisition, whereas the coefficients in Columns (3) tell us how many skilled people responded by migrating to that city. A general pattern revealed by this exercise is that younger workers respond primarily through skill acquisition, while older workers respond primarily through migration.

As noted by Borusyak et al. (2022b), in a spatial setting like ours, interpreting the coefficient for migration as capturing the true importance of migration might be misleading, as our exposure measure does not account for the true exposure through labor market linkages. Instead of an exact decomposition, our main takeaway from Table 3 is that migration could be an important mechanism for the change in regional college share. We will incorporate and quantify this mechanism through the model.

Summary. Taking stock, we find that city-level capital goods imports lead to an increase in skill shares. The most likely explanation for this finding is that facing an increase in skill premium, people respond by going to college or through migration, with both mechanisms playing a statistically significant role.

Given the spatial linkages between cities, these reduced-form estimates do not recover the general equilibrium effect of capital goods imports. To speak to such effects, we now turn to quantitative analysis.

4 The Model

We consider an environment with domestic and international trade, capital-skill complementary, and education and migration choices. The trade block of our model builds on Burstein et al. (2013) and Parro (2013), which extends Eaton and Kortum (2002) and Caliendo et al. (2019) to incorporate capital-skill complementarity; the migration block of the model builds on the recent quantitative spatial model (see Redding and Rossi-Hansberg, 2017 for a review). Our main departure from existing models is that we incorporate endogenous education choice, which allows the model to speak to the empirical patterns documented in Section 3.

4.1 Environment

Our model is static. There are $d = 1, ..., N$ locations, representing N Chinese cities and a synthetic rest of the World (the last location; hereafter RoW). Workers can migrate within China but not internationally. Chinese cities trade with each other and with the RoW.

There are three sectors, denoted by $s \in \{K, OT, NT\}$, which stand for capital goods (K), other tradable (OT), and non-tradable (NT), respectively. Within each sector, there is a continuum of intermediate goods, indexed by $\omega \in (0, 1)$. These intermediate goods are traded across regions (except for those in the NT sector, which are subject to infinite iceberg trade costs) and aggregated to make composite sector final goods. Composite final goods are not traded and used locally for consumption or as input for production.

In the rest of this section, we describe first production and trade decisions and then migration and education choices. Finally, we define the equilibrium and explain how we will bring the model to the data to quantify the impacts of imported capital goods (more precisely, the shock that drives the import growth) on skill acquisition and skill distribution within China.

4.2 Production and Trade

Consider an intermediate variety $\omega \in (0, 1)$ in sector $s \in \{K, OT, NT\}$. Each region d draws variety-specific productivity $z(\omega)$ from a Frechet distribution with location parameter T_d^s and dispersion parameter θ. The production technology for ω is given by:

$$y_d^s(\omega) = z(\omega) \cdot [q_d^{s,KHL}(\omega)]^{\gamma_d^{s,KHL}} \cdot [q_d^{s,OT}(\omega)]^{\gamma_d^{s,OT}} \cdot [q_d^{s,NT}(\omega)]^{\gamma_d^{s,NT}}, \tag{6}$$

where $q_d^{s,OT}(\omega)$ and $q_n^{s,NT}(\omega)$ are the use of the final good in sector OT and NT, respectively, in the production of ω; $q_d^{s,KHL}(\omega)$ is the use of "equipped labor"; $\gamma_d^{s,\tilde{s}}, \tilde{s} \in \{KHL, OT, NT\}$ is the output elasticity for these inputs.

Letting L_d^H and L_d^L be the number of high- and low-skill workers *employed* in d, the supply of "equipped labor", denoted by q_d^{KHL}, is given by

$$q_d^{KHL} = \Big[(\mu_d)^{\frac{1}{\rho_2}}[L_d^L]^{\frac{\rho_2-1}{\rho_2}} + (1-\mu_d)^{\frac{1}{\rho_2}}[q_d^{KH}]^{\frac{\rho_2-1}{\rho_2}}\Big]^{\frac{\rho_2}{\rho_2-1}} \tag{7}$$
$$q_d^{KH} = \Big[(1-\lambda_d)^{\frac{1}{\rho_1}} \cdot [L_d^H]^{\frac{\rho_1-1}{\rho_1}} + (\lambda_d)^{\frac{1}{\rho_1}} \cdot [q_d^K]^{\frac{\rho_1-1}{\rho_1}}\Big]^{\frac{\rho_1}{\rho_1-1}}, \quad \rho_2 > \rho_1 > 0.$$

This nested-CES production function follows that of Krusell et al. (2000), which has been applied to trade settings by Parro (2013) and Burstein et al. (2013). In the inner nest, capital equipment is combined with high-skill labor to produce "equipped skill" q_d^{KH}; in the outer nest, q_d^{KH} is then further combined with low-skill labor to produce "equipped labor" q_d^{KHL}. λ_d and μ_d are location-specific technology shifters that determine the share of each factor in equipped labor q_d^{KHL}. Under the assumption that $\rho_2 > \rho_1 > 0$, the model displays complementarity between capital and high-skill labor, so a decrease in the price of capital equipment increases the skill demand and the skill premium. The unit cost for "equipped skill" and "equipped labor", denoted by c_d^{KH} and c_d^{KHL}, respectively, are given by

$$c_d^{KH} = \Big[(1-\lambda_d)(w_d^H)^{1-\rho_1} + (\lambda_d)(P_d^K)^{1-\rho_1}\Big]^{\frac{1}{1-\rho_1}}$$
$$c_d^{KHL} = \Big[(\mu_d)(w_d^L)^{1-\rho_2} + (1-\mu_d)(c_d^{KH})^{1-\rho_2}\Big]^{\frac{1}{1-\rho_2}}.$$

Letting c_d^s be the production cost of a variety ω of a firm with unit productivity, w_d^e be the wage for workers of skill e, and P_d^s be the price for the sectoral final goods, equation (6) implies

$$c_d^s = \Big[\frac{c_d^{KHL}}{\gamma_d^{s,KHL}}\Big]^{\gamma_d^{s,KHL}} \cdot \Big[\frac{P_d^{OT}}{\gamma_d^{s,OT}}\Big]^{\gamma_d^{s,OT}} \cdot \Big[\frac{P_d^{NT}}{\gamma_d^{s,NT}}\Big]^{\gamma_d^{s,NT}}. \tag{8}$$

In each sector and each location, a representative sectoral final good producer combines intermediate varieties from the same sector using CES technology. The total sectoral final goods produced in sector s at location d, denoted by Q_d^s, is:

$$Q_d^s = \Big[\int_0^1 [Q_d^s(\omega)]^{\frac{\sigma}{\sigma-1}} d\omega\Big]^{\frac{\sigma}{\sigma-1}},$$

where $Q_d^s(\omega)$ is the use of variety ω by the final good producer. For each variety, the final good producer source from the country with the lowest shipping-cost inclusive price. The Frechet assumption implies that the trade shares and the sectoral price index are given by:

$$\pi_{do}^s = \frac{T_o^s(c_o^s\tau_{od}^s)^{-\theta}}{\Phi_d^s}, \text{ where } \Phi_d^s \equiv \Big[\sum_o T_o^s(c_o^s\tau_{od}^s)^{-\theta}\Big] \tag{9}$$
$$P_d^s = [\Gamma(\frac{\theta+1-\sigma}{\theta})]^{\frac{1}{1-\sigma}} \cdot (\Phi_d^s)^{-\frac{1}{\theta}},$$

where τ_{od}^s is the iceberg cost of shipping goods from o to d, and $\Gamma(\frac{\theta+1-\sigma}{\theta})$ is a constant that depends only on θ and σ.

4.3 Education, Migration, and Consumption Decisions

The labor supply in the RoW is fixed. In China, the supply of high- and low-skill labor to cities is the result of individuals' skill and migration choices. We assume that each origin city i is home to two cohorts of people, the young and the mature, indexed by Y and M with their quantity denoted by $\underline{L}_{i,Y}$ and $\underline{L}_{i,M}$. Young people decide whether to receive a college education. Mature people take their skills as given, an assumption motivated by the fact that during our sample period, people above the conventional college age do not respond much through skill acquisition. Both groups can migrate. Thus, in response to a change in the wage premium across cities, mature people respond through migration, whereas young people respond through both education and migration. In what follows, we describe first the education choice of the young and then the migration of both groups.

Education. Consider a young worker ν who is born in city i. Her decision about whether to receive a college education is described by:

$$\max\{\frac{1}{\delta_i} \cdot u^H_{i,Y} \cdot \zeta^H_i(\nu), \qquad u^L_{i,Y} \cdot \zeta^L_i(\nu)\},$$

In this problem, $u^H_{i,Y}$ and $u^L_{i,Y}$ are the expected utility of young workers from location i with skill H or L, to be defined below. $\zeta^H_i(\nu)$ and $\zeta^L_i(\nu)$ are individual ν's idiosyncratic tastes for the two skill levels, drawn from a Frechet distribution with shape parameter ξ. δ_i represents the utility cost of a college education, encapsulating in a reduced-form way the effort, time, and monetary commitments involved in prepping for the National College Entrance Exam ('gaokao') and then pursuing a college degree. While the monetary and time cost of pursuing a college degree might not vary significantly across locations, the level of effort needed for college admission can vary markedly. In particular, as universities allocate more seats to the home province, the difficulty of qualifying for college depends on one's home location. We thus allow δ_i to be specific to i.

The fractions of young workers from i that chooses H and L, denoted by l^H_i and l^L_i are

$$l^H_i = \frac{(u^H_{i,Y}/\delta_i)^\xi}{(u^H_{i,Y}/\delta_i)^\xi + (u^L_{i,Y})^\xi}, \qquad l^L_i = \frac{(u^L_{i,Y})^\xi}{(u^H_{i,Y}/\delta_i)^\xi + (u^L_{i,Y})^\xi}.$$

The number of high- and low-skill young workers from i are $\underline{L}_{i,Y} \cdot l^H_i$ and $\underline{L}_{i,Y} \cdot l^L_i$, respectively.

Migration. Having made their education choice, young workers choose the location of work to maximize their utility, along with mature workers. The utility a worker in age group $a \in \{Y, M\}$ and skill group $e \in \{H, L\}$ obtains by moving from origin i to destination d depends on three components: the indirect utility in the destination city, the "iceberg" migration cost from i to d, and the idiosyncratic preference draw of the worker for d. Formally, the location choice of worker ν is represented by:

$$u^e_{i,a}\Big(\epsilon^e_1(\nu),\epsilon^e_2(\nu),...,\epsilon^e_N(\nu)\Big) = \max_d\{\frac{v^e_{d,a}}{\kappa^e_{id,a}}\cdot\epsilon^e_d(\nu)\},$$

where $v^e_{d,a}$ is the indirect utility workers of type (a,e) derive from residing in d; $\kappa^e_{id,a}$ is the migration cost from i to d, which can differ by both skill and age; $\epsilon^e_d(\nu)$ is a location-specific taste draw that is independent across individuals and locations. We assume $\epsilon^e_d(\nu)$ follows a Frechet distribution with a common location parameter across d and a shape parameter η^e.[25]

Workers choosing a city derive utility from local amenities and the consumption of sectoral final goods according to the following preference:

$$A^e_{d,a}\cdot(C^K_d)^{\alpha^K}\cdot(C^{OT}_d)^{\alpha^{OT}}\cdot(C^{NT}_d)^{\alpha^{NT}}.$$

$A^e_{d,a}$ is the amenities in location d for a worker of type (a,e); C^s_d is the consumption of sectoral final good in sector s. By allowing $A^e_{d,a}$ to depend on (a,e), we capture that amenities that appeal to young low-skill workers might not be as attractive to mature high-skill people, and vice versa. This preference gives an indirect utility of

$$v^e_{d,a} = \frac{A^e_{d,a}\cdot w^e_d}{P_d},$$

where w^e_d is the earning in d and

$$P_d = (\frac{P^K_d}{\alpha^K})^{\alpha^K}\cdot(\frac{P^{OT}_d}{\alpha^{OT}})^{\alpha^{OT}}\cdot(\frac{P^{NT}_d}{\alpha^{NT}})^{\alpha^{NT}}. \tag{10}$$

Under the Frechet assumption, the fraction of group (a,e) from i choosing d for work is

$$\lambda^e_{id,a} = \frac{(\frac{A^e_{d,a}\cdot w^e_d}{P_d\cdot\kappa^e_{id,a}})^{\eta^e}}{\sum_{d=1}^N(\frac{A^e_{d,a}\cdot w^e_d}{P_d\cdot\kappa^e_{id,a}})^{\eta^e}}. \tag{11}$$

The expected utility among workers in this group is

$$u^e_{i,a} = \tilde{\gamma}\left[\sum_{d=1}^N(\frac{A^e_{d,a}\cdot w^e_d}{P_d\cdot\kappa^e_{id,a}})^{\eta^e}\right]^{\frac{1}{\eta^e}}, \tag{12}$$

[25]Systematic differences in preference across cities will be captured in amenities.

with $\tilde{\gamma}$ being a constant. The supply of young and mature workers with skill e to location d is

$$L^e_{d,Y} = \sum_{i=1}^{N} \underline{L}_{i,Y} \cdot l^e_i \cdot \lambda^e_{id,Y},$$
$$L^e_{d,M} = \sum_{i=1}^{N} \underline{L}_{i,M} \cdot \lambda^e_{id,M}$$

The total supply of workers in d with skill e is

$$L^e_d = L^e_{d,Y} + L^e_{d,M}.$$

Discussion on model assumptions. In our model, young workers make education choices before migration decisions and the realization of their location-specific preference draws ϵ^e_d. This setup might seem inconsistent with the observation that some people attend college in other cities and stay there after graduation. As noted earlier, a major component of college education cost, which we capture with δ_i, is the time commitment required to prepare for the entrance exam. This process starts long before college admission—see Ash (2016) for a description of the process and the stress it entails. During much of the preparation process, most students do not have a clear idea about which college they will get into, let alone where they will reside after graduation. In this sense, the education choice in our model is best thought of as the decision of whether to start the years-long preparation process. This interpretation, and the fact that college seats are allocated by province, also motivate an origin-specific—rather than destination-specific—cost for college education, introduced via δ_i.

4.4 The Definition of Equilibrium

The primitives of the model include the structural elasticities governing production, trade, and household choices $\{\theta, \sigma, \gamma^{s,s'}, \alpha^s, \rho_1, \rho_2, \xi, \eta^e\}$, CES production function share parameters $\{\lambda_d, \mu_d\}$, fundamental shifters including $\{T^s_d, \tau^s_{od}, A^e_{d,a}, \delta_d, \kappa^e_{od,a}\}$, and labor endowment $\{\underline{L}_{d,a}\}$. Given these primitives, a competitive equilibrium of the model is a set of prices and allocations so that i) goods and labor markets are clear; ii) all workers make optimal location choices; young workers make optimal education choices; and iii) all firms are price takers and choose quantities to maximize profit. We describe the conditions that characterize the equilibrium in Appendix B.1.

4.5 First-Order Approximation for Estimation and Counterfactuals

We will use the model to examine how the acquisition and distribution of skills in China are affected by a change in the fundamentals of the economy, focusing on the cost of importing capital goods observed over 2000-2010. To this end, we first recover model primitives that govern the response of individual decisions to changing wage structures, η^e and ζ, among other model parameters. We use first-order approximations to write the changes in models' endogenous outcomes as linear combinations of the change in model fundamentals. As articulated by Adao et al. (2019) and discussed below, this approach will enable us to be precise about the identification assumption imposed while ensuring consistency with the previous reduced-form analysis.

To simplify the exposition, we use bold letters to denote vectors and matrices obtained from stacking variables by location (and sector). For example, We define $\boldsymbol{T}_{(N\times 3\text{ matrix})}$ to be $\left[\boldsymbol{T}^{K}_{(N\times 1)}, \boldsymbol{T}^{OT}_{(N\times 1)}, \boldsymbol{T}^{NT}_{(N\times 1)}\right]$, where the subscripts indicate the dimension of variables. Similarly, $\boldsymbol{\tau}_{(N\times 3N)} \equiv \left[\boldsymbol{\tau}^{K}, \boldsymbol{\tau}^{OT}, \boldsymbol{\tau}^{NT}\right]$, $\boldsymbol{A}_{(N\times 4)} \equiv \left[\boldsymbol{A}^{H}_{Y}, \boldsymbol{A}^{H}_{M}, \boldsymbol{A}^{L}_{Y}, \boldsymbol{A}^{L}_{M}\right]$. We use $\bar{x}$ to indicate a variable x of a baseline competitive equilibrium and $\hat{x}$ to indicate the percentage change in variable x between two competitive equilibria. Lastly, we use $vec(\boldsymbol{x})$ to denote the vector formed by stacking matrix $\boldsymbol{x}$ by column. We have the following proposition.

Proposition 1. *Up to the first order, the percentage change from a baseline equilibrium to a counterfactual equilibrium in any outcome of the model—such as the skill share or total employment of a city—due to changes in the model's fundamentals can be expressed as a linear combination of the percentage change in these fundamentals, with the weight depending solely on outcomes in the baseline equilibrium and the models' structural elasticities. That is,*

$$\begin{aligned}\hat{\boldsymbol{y}}_{(N\times 1)} =& \boldsymbol{B}^{y,A}_{(N\times 4N)} \cdot vec(\hat{\boldsymbol{A}}) + \boldsymbol{B}^{y,T}_{(N\times 3N)} \cdot vec(\hat{\boldsymbol{T}}) + \boldsymbol{B}^{y,\tau}_{(N\times 3N^2)} \cdot vec(\hat{\boldsymbol{\tau}}) \\ &+ \boldsymbol{B}^{y,\delta}_{(N\times 2N)} \cdot vec(\hat{\boldsymbol{\delta}}) + \boldsymbol{B}^{y,\kappa}_{(N\times 4N^2)} \cdot vec(\hat{\boldsymbol{\kappa}}),\end{aligned} \tag{13}$$

where $\hat{\boldsymbol{y}}_{N\times 1}$ is an N-by-1 vector representing a change in outcome y in each location, $vec(\hat{\boldsymbol{A}})$, $vec(\hat{\boldsymbol{T}})$, $vec(\hat{\boldsymbol{\tau}})$, $vec(\hat{\boldsymbol{\delta}})$, and $vec(\hat{\boldsymbol{\kappa}})$ are vectors of fundamental changes, and matrices $\boldsymbol{B}^{y,f}, f \in \{A, T, \tau, \delta, \kappa\}$ are non-linear functions of the baseline equilibrium and structural elasticities.

This proposition expresses the changes in model outcomes in a counterfactual experiment as a linear function of the changes in fundamentals. $\boldsymbol{B}^{y,f}, f \in \{A, T, \tau, \delta, \kappa\}$ in equation (13) summarize the general equilibrium elasticities of endogenous outcome y with respect to the change in each of the fundamentals. For example, if $\boldsymbol{y}$ is the vector of high-skill worker employment $\{L^{H}_{d}\}$, the $(1, 1)^{th}$ element of $\boldsymbol{B}^{y,T}$ is the elasticity of outcome L^{H}_{1} with respect to the productivity in the capital industry in location 1. Such elasticities are general equilibrium, as they encapsulate the effects of a fundamental on y both direct and indirect through all endogenous prices and allocations.

According to Proposition 1, conditional on the structural elasticities and the outcomes from the baseline equilibrium, the level of fundamentals in the baseline economy is not necessary for counterfactuals. This result bears a resemblance to the "exact-hat algebra" approach pioneered by Dekle et al. (2008). The difference here is that the impacts of the change in different model fundamentals are linearly separable, a result of the first-order approximation. As discussed below, this linear separability has implications for using Proposition 1 for both identification and counterfactuals.

Implication for identification of structural parameters. Suppose we observe the outcome of the model in two periods and have measured the changes in model fundamentals between these two periods (e.g., $\hat{\boldsymbol{\tau}}$), then equation (13) can be used to estimate the structural elasticities of the model. In particular, as the GE elasticity matrices $\boldsymbol{B}^{y,f}, f \in \{A, T, \tau, \delta, \kappa\}$ are determined solely by the baseline equilibrium outcomes (the data) and structural elasticities, we can pick the structural elasticities by minimizing the differences between the actual change in an outcome variable and the model-predicted change according to equation (13).

A complication that arises is that, given the dimension of the model, researchers rarely observe the changes in *all* fundamentals.[26] If only a subset of the changes in fundamentals is observed, then the changes in $\hat{\boldsymbol{y}}$ that is attributed to the unobserved components enter as a structural error in the estimation, resulting in a potential endogeneity problem. Following Adao et al. (2019), we will use an identification assumption that is consistent with the reduced-form analysis for structural estimation.

Let $\hat{\tau}^K_{Ni}$, the $(N,i)^{th}$ element of $\boldsymbol{\tau}^K$, denote the change in capital import cost of a Chinese location i from the RoW, and let $b^{y,\tau}_{(d,i\times N)}$ denote the $(d, i\times N)^{th}$ element of $\boldsymbol{B}^{y,\tau}$, which is the elasticity of y_d with respect to $\hat{\tau}^K_{Ni}$, with y being any endogenous outcomes of the model, then $\hat{y}_d$ can be expressed as:

$$\hat{y}_d = a + \sum_{i=1}^{N-1} b^{y,K}_{(d,i\times N)} \cdot \hat{\tau}^K_{Ni} + \iota^y_d, \tag{14}$$

Here, a is the mean difference in $\hat{\boldsymbol{y}}$ between the two periods (i.e., $\frac{1}{N-1}\sum_{d=1}^{N-1}\hat{y}_d$), $\sum_{i=1}^{N-1} b^{y,K}_{(d,i\times N)} \cdot \hat{\tau}^K_{Ni}$ is the effect on outcome y in location d of the change in capital goods importing costs across the country, and ι^y_d is the structural residual that captures the general equilibrium effects of the changes in all fundamentals other than $\{\tau^K_{Ni}\}_{i=1}^{N-1}$.

We assume that conditional on the outcomes of the baseline equilibrium $\bar{\boldsymbol{Y}}$, the covariance between $\{\hat{\tau}^K_{Ni}\}_{i=1}^{N-1}$ and the structural error $\{\iota^y_d\}$ are zero. Formally, $\forall i, d, Cov(\hat{\tau}^K_{Ni}, \iota^y_d|\bar{\boldsymbol{Y}}) = 0$. A sufficient condition for this to hold is that the shocks to capital goods importing costs are uncorrelated with shocks to other fundamentals of the economy. While this assumption can be strong, it is consistent with the identification imposed for a causal interpretation of many recent studies applying a shift-share design to a spatial equilibrium setting, including the one presented in our empirical section. To see this, recall that in the empirical section, we predict the capital import growth of a city using the inner product between the city' composition of capital goods imports across product categories and the (leave-one-out) national import growth in these categories, viewing the source of variations as coming from random national shocks to product-level capital import costs that are independent of changes in other fundamentals. In the model, there is only one capital goods sector, so we view $\hat{\tau}^K_{Ni}$ as the weighted average capital import cost change across product categories and will measure it as such.[27] If the shocks at the

[26] Under assumed values for structural elasticities, one can "invert" the data to obtain the changes in fundamentals. However, changes obtained through this inversion will satisfy equation (13) by design when structural elasticities take the assumed values, rendering equation (13) useless for estimation.

[27] We chose to incorporate only one capital in the model because when we move to the quantification, the data on domestic trade is not as granular as for international trade, precluding us from accurately measuring domestic trade

product category level are uncorrelated with other shocks, an assumption we have maintained and provided supported for, so are their weighted averages—that is, $\hat{\tau}^K_{Ni}$ are uncorrelated with other changes to fundamentals. This motivates the following moment condition

$$\mathbb{E}\left[\iota^y_d \cdot (\hat{\tau}^K_{Nd} - \bar{\hat{\tau}}^K_{Nd})\right] = 0, \tag{15}$$

where $\bar{\hat{\tau}}^K_{Nd} \equiv \frac{1}{N-1}\sum_{d=1}^{N-1} \hat{\tau}^K_{Nd}$ is the mean capital import cost shock.[28] In the next section, we will use this condition for identification.

Implications for counterfactuals. Proposition 1 also shows that the impact of $\{\tau^K_{Ni}\}_{i=1}^{N-1}$ on equilibrium outcomes do not depend on the change in other fundamentals. In particular, as discussed in Section 2, China went through a higher-education expansion, which can have major impacts on skill acquisition and distribution. Proposition 1 shows that to understand the impact of the change in capital goods importing cost, we do not have to back out the change in the other fundamentals—and our finding will have first-order accuracy.

Two caveats to this approach are noteworthy. First, one may think that part of the driving force of higher-education expansion, captured in our model by $\hat{\delta}_i$, is endogenous. For instance, the increase in capital goods imports could contribute to growth in household income and local government revenue. The former could make college more affordable, whereas the latter could increase the supply of college seats. While these two channels may be both important at the national level, our empirical evidence suggests that they do not seem to interact with capital goods imports at the local level, so in carrying out counterfactuals, we will assume that $\{\tau^K_{Ni}\}_{i=1}^{N-1}$ does not affect $\hat{\delta}_i$—in line with the previously imposed identification assumption. To the extent that capital goods imports lead to a decrease in δ_i common to all locations by increasing national income and central government revenue, our counterfactual findings can be viewed as a lower bound for the full effects of capital goods imports.

Second, although first-order approximation simplifies the analysis—and has been shown by recent studies (e.g., Kleinman et al., 2020) to be accurate for trade models—one might still be concerned that the higher-education expansion and other reforms (such as the hukou reform) carried out during this period can have a big impact on the economy, thereby interacting with $\{\tau^K_{Ni}\}_{i=1}^{N}$ nonlinearly. For example, when the cost of going to college is lower, the same increase in skill premium will have a bigger effect on skill acquisition than our model implies. To provide a bound for the importance of such nonlinear mechanisms, we will carry out a counterfactual analysis around the 2010 education and migration outcomes, which would account for the lower education/migration costs in the post-reform economy.

in narrowly defined capital goods in the baseline economy.

[28] Adao et al. (2019) shows that under the orthogonality condition that we invoke, one can use a more general form of the specification, $\mathbb{E}\left[\iota^y_d \cdot \sum_i h_{di}(\hat{\tau}^K_{Ni} - \bar{\hat{\tau}}^K_{Ni})\right] = 0$, choosing h_{di} appropriately to increase estimation efficiency. We use an identity matrix both for simplicity and because we lack accurate first-hand data on trade between Chinese locations.

5 Quantification

5.1 Data

We use the data on China's spatial economy and international trade around the year 2000 to construct the outcomes in the baseline competitive equilibrium, which will be denoted by letters with an upper bar.

Employment and migration data. The demographic and migration data are constructed based on a 1% individual-level sample of the 2000 population census. We classify individuals aged between 18 and 60 into two groups, young (18-35) and mature (36-60), denoted by Y and M, respectively. For each individual, we observe their education and current location, which allows us to construct regional employment $\{\bar{L}^e_{d,a}\}$, $\forall a \in \{Y, M\}, e \in \{H, L\}$. We use the city-to-city migration flows from Dorn and Li (2023). As our model focuses on a one-shot long-term migration decision, the migration measure we will use is the one based on birthplace.

Factors shares in production and consumption. We assume that the shares of equipped labor and other final goods in the production of intermediate goods ($\gamma^{s,s'}$ in equation (6)) and the shares of household expenditures over final goods differ between China and the RoW but not between cities within China. We set these parameters using the corresponding values in Fan (2019), who calculates the values for China using the Chinese input-output table and the values for the synthetic RoW as the average values of major countries based on WIOD.

Equipped labor contains value added from capital equipment and high- and low-skill workers. We obtain the shares of capital in equipped labor as follows. For Chinese locations, we measure the capital share in value-added using the Annual Survey of Industrial Firms, calculating this share for each city by aggregating across firms there. For the synthetic RoW, we assume that 40% of the value added comes from capital, with the remaining from labor, a division close to that of the U.S. around 2000.

We then measure the relative share of high- and low-skill workers in equipped labor. For Chinese cities, these shares are calculated from the employment and wages of high- and low-skill workers using three sources of information: employment data from the population census and wage data from two sources—the 2002 UHS, which covers 181 cities (earlier UHS covers only half as many cities), and the mini census from 2005, which covers all Chinese cities. For cities covered by the UHS, we use the UHS to measure the average wage for high- and low-skill workers; for cities not covered by it, we infer their wages in 2002 using the wage in the 2005 mini census and the wage growth rate of similar cities between 2002 and 2005.[29] For the synthetic RoW, we assume skilled workers account for about 73% of the labor value added, a share calculated based on the U.S. college share and skill premium around 2000.

Trade and production. We measure city-sector level trade shares and value-added using the model-based imputation of Fan (2019), who estimates a model of domestic and international

[29]Using the 2005 mini census average wage for all cities leads to materially the same quantitative findings.

trade with capital-skill complementarity using China's inter-provincial input-output table and city-sector level import, export, and value-added data. As the production of the model there is similar to the one adopted in this paper,[30] the values from Fan (2019) are consistent with what we would have obtained if we have repeated the same structural estimation process. This step gives us $\{\bar{\pi}^s_{od}\}$ and $\{\bar{X}_d\}$.

Measuring the change in capital goods import cost. To operationalize the model for counterfactuals, we need to take a stand on what drives the growth in capital goods imports—e.g., foreign capital goods sector productivity, changes in capital goods import costs, or other reasons. Given our focus on the outcomes within China, and considering that our model incorporates only one capital good sector, we assume that the import increase is driven by a location-specific decrease in importing costs. Thus, $\forall i = 1, ..., N-1$, we measure $\hat{\tau}^K_{Ni}$ as

$$\hat{\tau}^K_{Ni} = -\frac{1}{\theta} \cdot \widehat{KIP}_i^{IV}, \tag{16}$$

where $\widehat{KIP}_i^{IV}$ is the *percentage growth rate* in capital goods imports predicted by the shift-share formula (i.e., the term inside the square bracket of equation (4), not including the initial imports-per-capita term), θ is trade elasticity, and $\hat{\tau}^K_{Ni}$ is the change in import cost necessary to account for $\widehat{KIP}_i^{IV}$.[31] As city-level capital goods import growth rates show a wide range of variations, the inferred $\hat{\tau}^K_{Ni}$ is widely dispersed. To reduce the impact of outliers, we winsorize $\hat{\tau}^K_{Ni}$ at -100 log points, which corresponds to approximately 64% decrease in capital goods importing cost (or a four-time increase in capital goods imports) over 2000-2010. Figure 6 depicts the histogram of $\hat{\tau}^K_{Ni}$.

5.2 Structural Elasticities

In addition to the parameters in production functions, the model has the following structural parameters: ρ_1 and ρ_2, which shape capital-skill complementarity; θ, the trade elasticity; ζ, which governs the dispersion in the taste draws in skill choice; η^e, $e \in \{H, L\}$, which govern the dispersion in workers' taste draws preferences for locations. Many of these parameters have been estimated in the literature, so we assign their values based on external estimates. In particular, we set $\theta = 4$ following Eaton and Kortum (2002). We set $\rho_1 = 0.67$ and $\rho_2 = 1.67$ following

[30] The main difference is that the 'OT' sector in the present paper is split into agricultural and other manufacturing in Fan (2019). We use this estimate because the readily available inter-provincial input-output table is only at the provincial-pair level, but our reduced-form analysis and model quantification are at the city level, so model-based imputation is necessary.

[31] In a model that explicitly incorporates many capital goods varieties, the partial equilibrium change in capital goods imports in the region i can be written as $\widehat{KIP}_i^{IV} = -\theta \cdot im^j_i / \sum_j im^j_i \cdot \hat{\tau}^j_{Ni}$, where im^j_i is the baseline import of capital goods variety j, and $\hat{\tau}^j_{Ni}$ is the change in the importing cost of that variety. Thus, $\hat{\tau}^K_{Ni}$ in our model can be viewed as the weighted average of the changes in the cost of importing individual goods in a model with many capital goods.

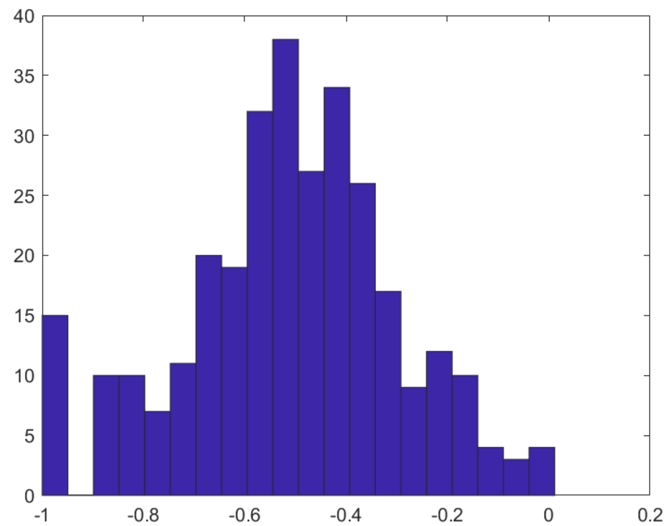


Figure 6 The Inferred Values of $\hat{\tau}^K_{Ni}$

Note: This figure plots the frequency distribution of $\{\hat{\tau}^K_{Ni}\}_{i=1}^{N-1}$, defined in equation (16).

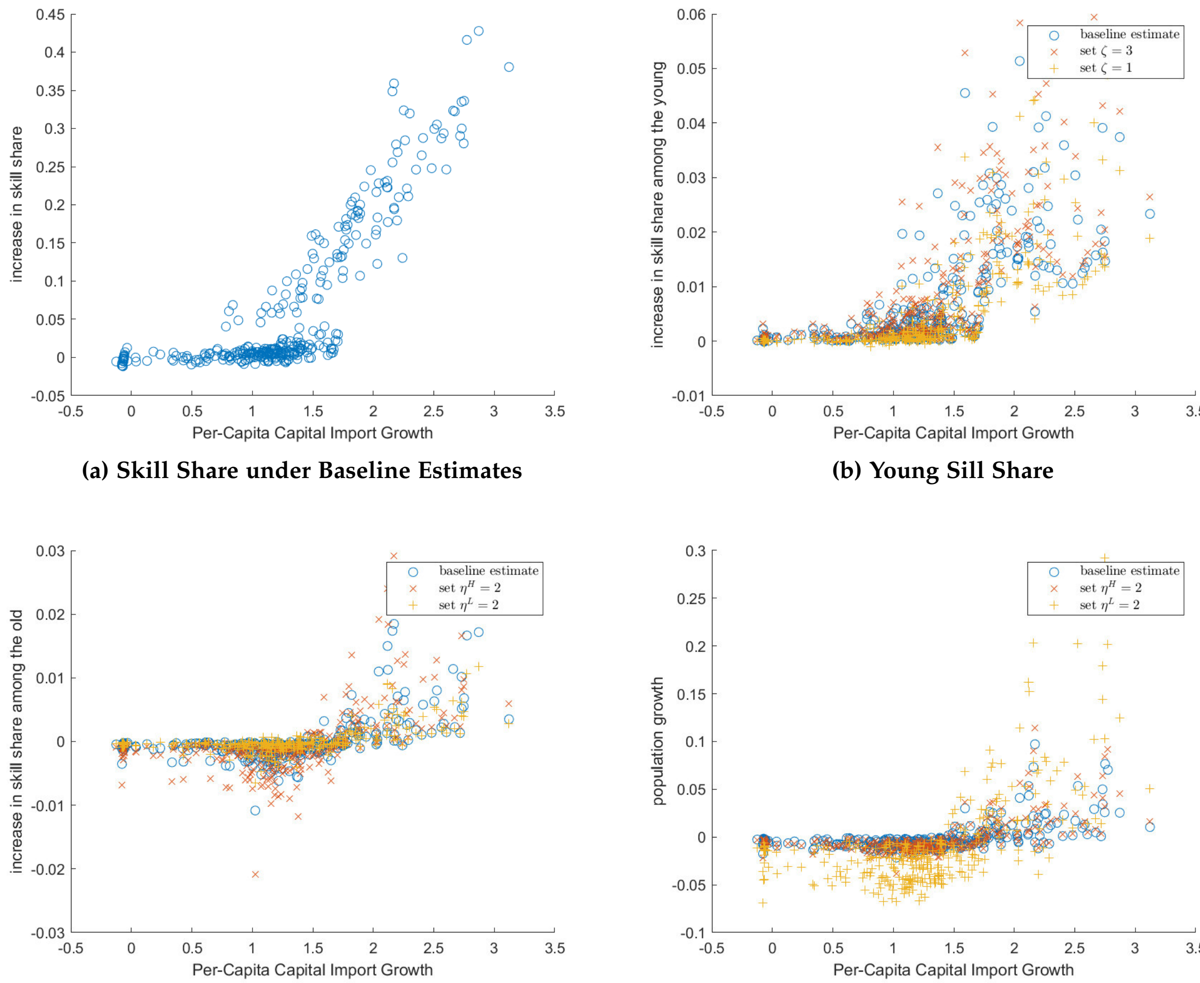
Krusell et al. (2000).[32]

The remaining elasticities, ζ, η^H, and η^L, are closely tied to the outcomes we seek to explain—the acquisition and spatial distribution of skills—and lack well-established estimates. As a result, we calibrate these internally. We identify these parameters using the changes in six outcome variables between 2000 and 2010: skill share, skill share among the city's stayers, skill share among young workers, skill share among young stayers, skill share among mature workers, and population. Intuitively, the change in a city's overall skill share represents the net effect of both skill acquisition and migration. By targeting the changes in population and skill share simultaneously, we can differentiate the impact of acquisition from migration. Moreover, we can further distinguish the migration parameters of high- and low-skill workers by matching the skill shares among different subgroups, where the two margins play varying roles.

Let $\hat{y}_d$ denote the change in these variables in the data, and let $\hat{y}_d(\Theta, \hat{\boldsymbol{\tau}}^K, \bar{\boldsymbol{Y}})$ be the model-implied change under the measured $\hat{\tau}^K_{Ni}$ and the baseline equilibrium data $\bar{\boldsymbol{Y}}$, and a particular set of structural parameters Θ. Our moment condition is that the de-meaned difference between $\hat{y}_d$ and $\hat{y}_d(\Theta, \hat{\boldsymbol{\tau}}^K, \bar{\boldsymbol{Y}})$ being orthogonal to the de-meaned τ^K_{Nd}.

Identification of the elasticities. We illustrate how various outcome variables assist in identifying the structural elasticities. Panel 7a of Figure 7 shows that, under the baseline estimate

[32] The estimates of Krusell et al. (2000) are based on an aggregate production function. One may be concerned that in our setting with many cities, these aggregate elasticities capture both the elasticities in a city-level production function and the substitution between cities with different factor shares, and therefore do not map into the elasticities in our city-level production function. See Oberfield and Raval (2021) for a discussion of this difference in the context of using firm-level data to estimate the aggregate elasticity. In the appendix, we exploit cross-region variations to estimate ρ_1 and ρ_2 internally at the city level. We find that $\rho_1 = 0.425$ and $\rho_2 = 1.175$, implying slightly weaker capital-skill complementarity. We show that using these internally estimated values leads to similar counterfactual findings.



(a) Skill Share under Baseline Estimates

(b) Young Sill Share

(c) Mature Skill Share

(d) Population Growth

Figure 7 Identification of Structural Parameters

Note: The figures illustrate how city outcomes vary with key structural elasticities. In all panels, the horizontal axis is the per-capita capital goods import growth. The vertical axis is, respectively, the change in skill share (Panel a), the change in skill share among young workers (Panel b), the change in skill share among mature workers (panel c), and the log difference in city population (Panel d).

(which we will discuss later), the increase in skill share is positively correlated with the growth of a city's per-capita capital goods imports. The cities with the most significant increase in imports see a rise in skill share by 3.5 percentage points, approximately two-fifths from their baseline levels. The changes in skill shares result from several decisions—the education decision of young workers, and the migration decision of both the young and mature. Since these decisions are governed by different parameters, we leverage the relative importance of these factors for identification.

Panel (7b) depicts the change in the skill share of the young, under both the baseline estimate ($\zeta = 1.95$) and alternative education choice parameters ($\zeta = 1$ or 3). As expected, the skill share among the young also increases most in cities experiencing the biggest growth in capital goods

Table 4 Summary of Structural Parameters

Parameters	Descriptions	Source/Target	Value
A. Calibrated externally			
ρ_1	capital-skill complementarity, lower nest	Krusell et al. (2000)	0.67
ρ_2	capital-skill complementarity, upper nest	Krusell et al. (2000)	1.67
θ	trade elasticity	Eaton and Kortum (2002)	4
σ	demand elasticity	-	5.5
B. Determined internally			
ζ	taste draws for education	equation (15)	1.95
η^H	tastes draws for locations: high-skill workers	equation (15)	0.91
η^L	tastes draws for locations: low-skill workers	equation (15)	0.34

Notes: Panel A reports the parameters that are calibrated externally. Panel B reports the parameters that are calibrated internally.

imports. The increase becomes more pronounced with higher ζ. Panel (7c) illustrates the change in skill shares among mature workers. In the most heavily exposed cities, skill shares in this age group increase by about 2 percentage points. A higher migration elasticity for skilled workers increases this number, while a higher migration elasticity for low-skilled workers decreases it. Thus, skill shares provide insight into the *differential* mobility between high- and low-skill workers. To pinpoint the *level* of mobility, we rely on population growth. Panel (7d) shows that increasing the migration elasticity for either high- or low-skill workers boosts the population growth in the most exposed cities.

Calibration Results. Panel A of Table 4 summarizes the sources and values of externally calibrated parameters. The lower panel summarizes the estimation results. We find that $\zeta = 1.95$, which means that a one-percent increase in college return $u^H_{i,Y}$ leads to a 1.95% increase in $\log(\frac{l^H_i}{l^L_i})$. In the baseline economy, 5.6% of the people included in the analysis received college education, so the 1.95% increase in log odds ratio translates to around 0.37 p.p. increase in college share. This effect might appear large, but note that it is estimated in a context with a low stock of college graduates and using a static model with a one-shot education choice. Thus, this number is best viewed as a long-run semi-elasticity. Our estimate of ζ also fall in the same ballpark as Ghose (2021), who estimates a 1.6 dispersion parameter for education choice in the context of India.

We find the dispersion parameters for location tastes to be 0.91 for skilled workers, and 0.34 for low-skill workers. With these estimates, our model implies that idiosyncratic tastes are less important in education choice than in migration decisions, echoing the finding of Kennan (2015). To place the estimated migration parameter in the literature, the absolute values of these parameters tend to be lower than Fan (2019) and Tombe and Zhu (2015), both of which focus on China. However, our finding that the migration decisions of high-skill workers tend to be more responsive to destination wages than that of low-skill workers is consistent with the finding of Fan (2019) based on the dispersion of the earnings distribution. In the appendix, we report a sensitivity analysis of alternative migration parameters.

Comparison to the data. Before turning to counterfactuals, we assess, in two ways, the model's ability to account for the differential changes in skill shares across cities. In the first assessment, we regress the model-predicted changes in skill share on the model-implied increase in per-capita capital goods imports due to the $\{\hat{\tau}^K_{Nd}\}_{d=1}^{N-1}$ shock, and then compare the regression coefficient to the reduced-form estimates reported in Table 1. This comparison is informative about whether the model can produce the spatial correlation between these two variables in the data. We estimate a coefficient of 0.64 (s.e. 0.094) from model-generated data. This coefficient falls between the baseline IV estimate reported in Column (3) of Table 1 and our preferred specification reported in Column (5). So under our parameterization, the model accounts for most of the spatial correlation between the two variables.

In the second assessment, as advocated by Adao et al. (2022), we regress the actual change in skill shares between 2000 and 2010 on the model-predicted changes. Under the null hypothesis that the model is correctly specified and parameterized and that the orthogonality condition holds, this regression should yield a coefficient of one. We find a coefficient of 1.28 (s.e. 0.35), unable to reject the null. This shows that our model yields a good *in-sample* fit for the key outcome of the empirical section.

Based on this analysis, we conclude that the parameterized model can generate the differential effects of capital goods import growth on college share across locations observed in the data.

5.3 The Impact of Capital Goods Import on Skill Acquisition and Distribution

We use the model to quantify how the increase in capital goods import affects skill acquisition in China and the distribution of skills across regions.

As shown in Table 5, there were about 42 million college graduates in China in 2000, accounting for 5.6% of the people between the ages of 18 and 60. After applying the inferred $\{\hat{\tau}^K_{Ni}\}_{i=1}^{N}$ shock, we see an increase of 3.7 million college graduates, which represents an 8.8% increase from the 2000 level. While significant, this number accounts for only a small fraction of the 70.3 million increase in college graduates seen from 2000 to 2010. Other factors, such as domestic productivity growth, reforms in the hukou system, and (especially) the broad expansion of higher education, likely had more substantial impacts.

In order to understand how these other changes shaped the effect of capital goods import on skill acquisition, we conducted an analysis based on the demographic data from 2010, when the total skill count is 112 million, accounting for 13.15% of the population. More formally, we use data on migration, education, and age distributions from 2010, instead of those from 2000, to conduct the counterfactual evaluations. By using the data from 2010, we implicitly incorporate the possibility that, due to productivity growth, reforms, and higher education expansion, the cost of acquiring a college education and/or migrating between regions are lower than 2000.[33] Because these changes unfolded gradually over the decade, counterfactuals from the baseline equilibrium with 2010 demographics can be seen as an upper bound on the impact of imported capital when migration and education costs are gradually lowered by other reforms.

We find that from this baseline, the $\{\hat{\tau}^K_{Ni}\}_{i=1}^{N-1}$ shock leads to an 8.9 million increase in college graduates, more than twice the level of the baseline estimate, accounting for approximately 13%

[33]These changes are reflected in, for example, higher college and migration shares in 2010 than in 2000.

Table 5 Counterfactual Outcomes

	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	45.36 million		120.84 million	
increase from baseline (%)	8.80%		7.91%	
Panel B: spatial distribution of skills				
	Coastal	**Inland**	**Coastal**	**Inland**
% increase in skill	18.23%	3.40%	16.17%	2.39%
% accounted for by				
Y, stayer	71.11%	114.96%	73.59%	120.80%
Y, migrant	23.12%	2.78%	21.48%	1.45%
M, stayer	2.45%	-10.21%	0.93%	-18.07%
M, migrant	3.32%	-7.53%	4.01%	-4.18%

Notes: Panel A reports the number of high-skill workers and their share in the population in the baseline equilibrium (the first two rows), and the effect of the counterfactual change (the third and fourth rows). The lower panel reports the total increase in the number of high-skill workers in coastal and inland regions, and the contribution of four margins, young stayers, young migrants, mature stayers, and mature migrants, to this increase. The total shares of these four margins sum up to 100%. For each panel, we report the results separately for when 2000 or 2010 demographics (education, migration, young and old group size) are used as the baseline economy.

of the 70.3 million increase in college graduates over the decade. Importantly, we find that the increase in college graduates as a share of the baseline economy (7.91%) is similar to the percentage increase inferred using 2000 demographics, so the larger estimates based on the 2010 demographics are due to a larger base number of college graduates (or alternatively, a lower implied college education cost), not a larger percentage response to the same shock.

We also examine the spatial distribution of skills in response to the shock. We find a much larger increase in skill share in coastal regions (18.2%) than in the hinterland (3.4%), which reflects that the coastal region generally experienced a larger increase in capital goods imports over the period. We decompose the increase into four groups: young stayers, young migrants, mature stayers, and mature migrants. According to this decomposition, young people from the coast (Y stayer) account for more than 70% of the total increase in college graduates on the coast; young migrants from the non-coastal region account for approximately 20%. Mature workers born on the coast and hinterland account for 2.45% and 3.3%, respectively. Recalling that mature workers do not make education choices, this change reflects that they disproportionately choose to stay in (or migrate to) the coastal region after capital import costs are reduced. This movement of mature workers contributes negatively to the skill share in the hinterland—the 3.4% increase in college graduate share there is entirely driven by young workers. We obtain similar findings when focusing on the 2010 demographics as the baseline economy.

Taking stock, our model-based analysis suggests that, depending on at which data point we evaluate the model, the increase in capital goods imports leads to between 3.7 to 8.9 million increase in college graduates, accounting for between 5.3% and 13% of the total increase in college graduates over the period. We also find that because the increase is concentrated in the coastal region, it leads to a substantially larger increase in skill shares there than in the hinterland, potentially contributing to China's spatial disparities. Moreover, these findings are robust across a range of parameters, as we show in the sensitivity analysis reported in the appendix.

6 Conclusion

Human capital is a fundamental determinant of long-run economic growth. In the globalized economy, understanding how it is shaped by international trade is important. In this paper, using empirical and quantitative analysis, we show that imported capital goods, a significant avenue for advanced technology adoption in developing countries, can increase the supply of high-skill people and lead to a reallocation of such people across space. Both mechanisms imply that capital goods imports can have long-run impacts on the spatial economy of the importing countries.

Several exciting opportunities present for future research. First, in this study, we have focused on imported capital goods both because it is relatively easy to measure and because, compared to home-grown technologies in developing countries, it is likely to be more complementary to skills. Future work should assess the similarities and differences between domestic and foreign imported capital goods. Second, motivated by both the importance of the Chinese economy and the salience of the mechanisms there, our analysis focuses on China. It is important to acknowledge that China differs from other developing countries in that it has a larger manufacturing sector, where capital-skill complementarity might be more important, and that it went through a rapid higher education sector expansion, which increased the accessibility of college education. Understanding whether the findings hold in other countries is an important next step.

References

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, *4* (PART B), 1043–1171.

_ and _ , "What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology," *Journal of Economic Literature*, 2012, *50* (2), 426–463.

Adao, Rodrigo, Costas Arkolakis, and Federico Esposito, "General Equilibrium Effects in Space: Theory and Measurement," 2022.

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

Ash, Alec, "Is China's Gaokao the World's Toughest School Exam," *The Guardian*, 2016, *12.*

Atkin, David, "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review*, 2016, *106* (8), 2046–2085.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–2168.

Bartik, Timothy J, "Who Benefits from State and Local Economic Development Policies?," *National Tax Journal*, 1991, *45* (4).

Berman, Eli, John Bound, and Zvi Griliches, "Changes in the Demand for Skilled Labor within US Manufacturing: Evidence from the Annual Survey of Manufactures," *The Quarterly Journal of Economics*, 1994, *109* (2), 367–397.

Blanchard, Emily J and William W Olney, "Globalization and Human Capital Investment: Export Composition Drives Educational Attainment," *Journal of International Economics*, 2017, *106*, 165–183.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, Jan 2022, *89* (1), 181–213.

_, Rafael Dix-Carneiro, and Brian Kovak, "Understanding Migration Responses to Local Shocks," *Available at SSRN 4086847*, 2022.

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

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 2011, *101* (1), 304–340.

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

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Global rebalancing with gravity: Measuring the burden of adjustment," *IMF Staff Papers*, 2008, *55* (3), 511–540.

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

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

Facchini, Giovanni, Maggie Y Liu, Anna Maria Mayda, and Minghai Zhou, "China's Great Migration: The Impact of the Reduction in Trade Policy Uncertainty," *Journal of International Economics*, 2019, *120*, 126–144.

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

Feenstra, Robert C and Gordon H Hanson, "Globalization, Outsourcing, and Wage Inequality," *The American Economic Review*, 1996, *86* (2), 240–245.

Feenstra, Robert C. and Gordon H. Hanson, "The Impact of Outsourcing and High-technology Capital on Wages: Estimates for the United States, 1979-1990," *Quarterly Journal of Economics*, 1999, *114* (3), 907–940.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Human Capital and Regional Development," *The Quarterly Journal of Economics*, 2013, *128* (1), 105–164.

Ghose, Devaki, "Trade, Internal Migration, and Human Capital: Who Gains from India's IT Boom?," *Working Paper*, 2021.

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

Goldin, Claudia and Lawrence F Katz, "The Origins of Technology-Skill Complementarity," *The Quarterly Journal of Economics*, 1998, *113* (3), 693–732.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, When, Why, and How," *American Economic Review*, 2020, *110* (8), 2586–2624.

Hsiao, Allan, "Educational Investment in Spatial Equilibrium: Evidence from Indonesia," 2022.

Katz, Lawrence and David Autor, "Changes in the Wage Structure and Earnings Inequality," *Handbook of Labor Economics*, 1999, *3*, 1463–1555.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Kennan, John, "Spatial Variation in Higher Education Financing and the Supply of College Graduates," 2015.

Kleinman, Benny, Ernest Liu, and Stephen J Redding, "International Friends and Enemies," 2020.

Koren, Miklós and Marton Csillag, "Machines and Machinists: Importing Skill-biased Technology," *mimeo, Central European University*, 2019.

Kovak, Brian K, "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?," *American Economic Review*, 2013, *103* (5), 1960–1976.

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.

Li, Bingjing, "Export Expansion, Skill Acquisition and Industry Specialization: Evidence from China," *Journal of International Economics*, 2018, *114*, 346–361.

Li, Hongbin, Lei Li, and Hong Ma, "China's Skill-biased Imports," *China Economic Review*, 2022, *74*, 101809.

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

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

Raveh, Ohad and Ariell Reshef, "Capital Imports Composition, Complementarities, and the Skill Premium in Developing Countries," *Journal of Development Economics*, 2016, *118*, 183–206.

Redding, Stephen J and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9*, 21–58.

Tombe, Trevor and Xiaodong Zhu, "Trade, Migration and Productivity: A Quantitative Analysis of China," *Manuscript, University of Toronto*, 2015.

Topalova, Petia, "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India," *American Economic Journal: Applied Economics*, 2010, *2* (4), 1–41.

Verhoogen, Eric A, "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *The Quarterly Journal of Economics*, 2008, *123* (2), 489–530.

Vogel, Ezra F, *Deng Xiaoping and the Transformation of China*, Vol. 10, Belknap Press of Harvard University Press Cambridge, MA, 2011.

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.

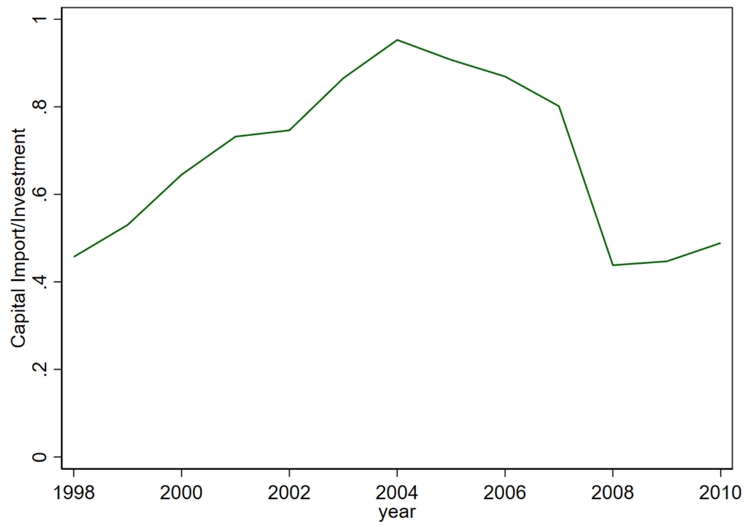


Figure A.1: Capital Goods Import/Investment

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%.

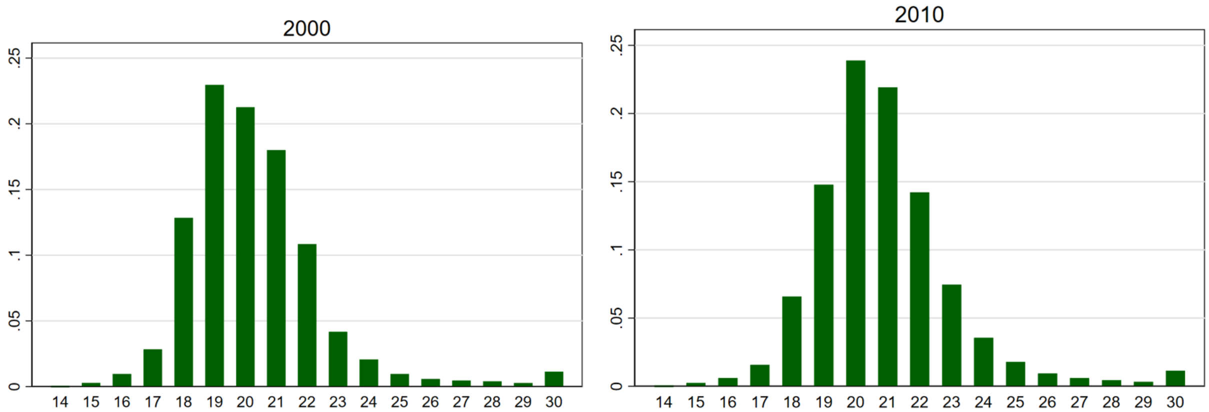


Figure A.2: Age Distribution of College Students

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: $100 \times \Delta$ (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.

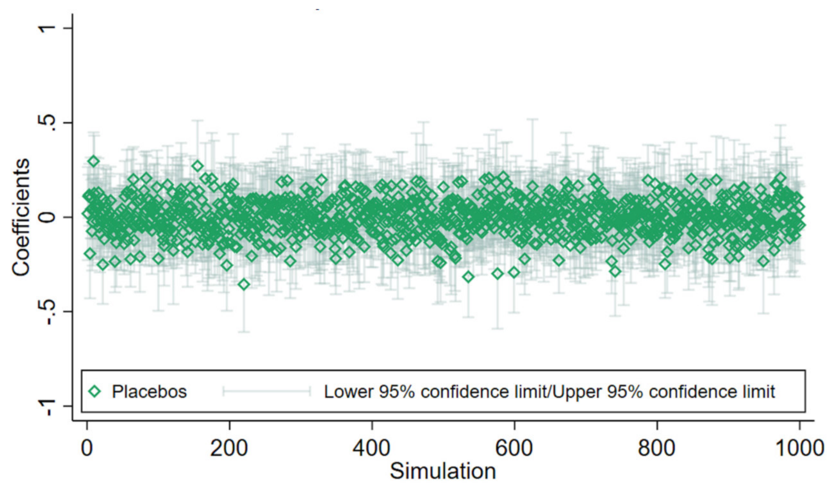


Figure A.3: Placebo import shocks

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.

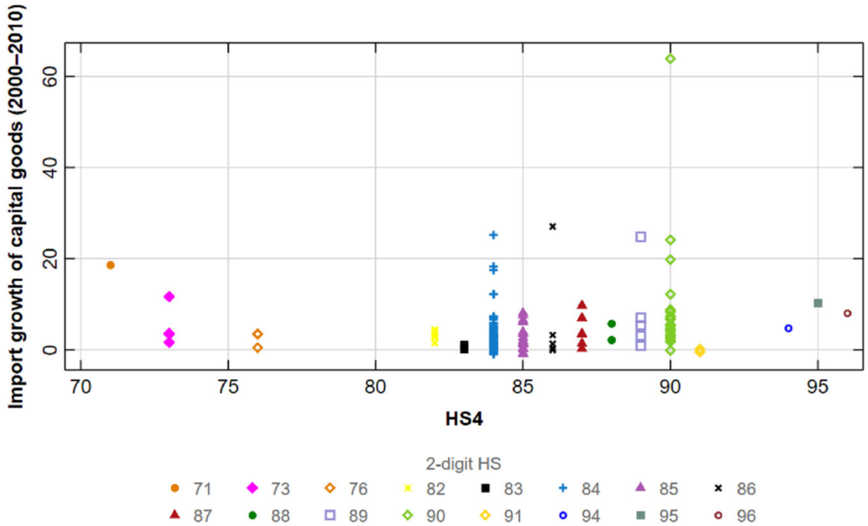


Figure A.4: Import Growth of Capital Goods by Products

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.

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

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 $\{w_i^L, w_i^L, P_i, P_i^s, c_i, c_i^{KH}, u_{i,a}^e, \text{ etc.}\}$ and allocations $\{X_i^s, \pi_{id}^s, \beta_i^{s,s'}, \lambda_{id,a}^e, l_i^e, L_{i,a}^e, L_i^e, \text{ etc.}\}$, 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

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

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

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

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

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

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.

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

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:

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

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 $\Big[\frac{\bar{P}_d^K \bar{q}_d^K}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H L_d^H} \cdot (-\frac{1}{\theta} \bar{\pi}_{dN}^K \cdot \widehat{KIP}_d^{IV}) + \frac{\bar{w}_d^H \bar{L}_d^H}{\bar{P}_d^K \bar{q}_d^K + \bar{w}_d^H \bar{L}_d^H} \cdot \Delta \ln w_d^H - \Delta \ln(w_d^L)\Big]$ 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.