China Economic Review 73 (2022) 101791

Contents lists available at ScienceDirect

China Economic Review

journal homepage: www.elsevier.com/locate/chieco

Heterogeneous firm responses to increases in high-skilled workers: Evidence from China's college enrollment expansion[☆]

Shuaizhang Feng[a], Xiaoyu Xia[b,*]

[a] Institute for Economic and Social Research, Jinan University, 601 Huangpu W. Ave, Guangzhou 510632, China.
[b] Department of Economics, University of Essex. Wivenhoe Park, Colchester CO4 3SQ, United Kingdom.

ARTICLE INFO

Keywords:
Local labor market
Firm-level adjustment
College education
Chinese economy

ABSTRACT

Over the past several decades, the returns to college education have steadily increased in many countries of the world despite an increased supply of college graduates. In this paper, using local-labor market data on the composition of the labor force combined with detailed firm-level data covering the period of a large-scale expansion of college enrollment in China, we seek to identify within-firm adjustments to labor market changes. The empirical work is guided by a model in which there are two types of production technologies, characterized by two different types of capitals, one skill-biased and the other labor-biased. The empirical results, consistent with the model and the observed trends in schooling and rates of return, indicate that there were significant adjustments in capital and R&D within-firms in response to an enlarged college-educated labor force.

1. Introduction

Over the past several decades, the returns to college education have steadily increased in many countries of the world despite an increased supply of college graduates.[1] While an expansion of world trade and skill-biased technology change have been proposed to explain this phenomenon, it remains unclear whether the observed wage trends are caused by sectoral shifts in the composition of firms or within-firm input mix adjustment. Knowledge about the form of demand-side adjustments is critical for designing policies targeted at labor markets.[2] For example, if the increased demand for high-skill workers is due to domestic-industry adjustments rather than to trade shocks, then altering trade policies may be an ineffective means of aiding the affected workers.

In this paper, using local-labor market data on the composition of the labor force combined with detailed firm-level data covering

☆ For insightful suggestions, we thank Enrico Moretti, David Card, Giovanni Peri, Lars Lefgren, Yifan Zhang, Pauline Leung, and various seminar participants at the University of California, Berkeley; National University of Singapore; University of Technology Sydney; Tsinghua University. We also thank Gaojie Tang, Peng Xu and Minzhi Liu for their excellent research assistance. All errors are our own. Feng acknowledges financial support from the Chinese National Science Foundation for distinguished young scholars (Project Number: 71425005). Xia acknowledges financial support from the Research Grants Council of Hong Kong (ECS Grant 24500416).

* Corresponding author.
E-mail address: xiaoyu.xia@essex.ac.uk (X. Xia).

[1] See Bekman, Bound, and Machin (1998), Card and Lemieux (2001), Acemoglu (2002), Autor, Katz, and Kearney (2008), and Walker and Zhu (2008) for example.

[2] Example of discussions on the link between trade and wages include Krugman (2000) and Falvey, Greenaway, and Silva (2010), and examples of studies on skill-biased technology change include Katz and Murphy (1992), Krueger (1993), Bekman et al. (1998), and Autor et al. (2008)).

https://doi.org/10.1016/j.chieco.2022.101791
Received 11 August 2021; Received in revised form 1 February 2022; Accepted 28 March 2022
Available online 6 April 2022
1043-951X/

the period of a large-scale expansion of college-educated workers in China, we seek to identify within-firm adjustments to labor market changes. By merging data at the firm level with local labor-market data, we rule out the possibility that shifts in the demand for high-skill workers are solely driven by the relative expansion of capital-intensive sectors or export-intensive industries. More specifically, we propose a model that characterizes heterogeneous firms with differing production technologies, which leads to their differential adjustments in the use of capital and skilled-labor when the skill-mix exogenously change in the labor market. We then use firm level data to examine the model predictions on the use of capital and skilled labor, as well as other factors of production. Our model is in line with previous work that has shown that within-firm adjustments are the main mechanism through which labor supply shocks are absorbed by the economy (see Dustmann & Glitz, 2015; Lewis, 2013). As we examine consequences of increases in high-skill labor, our study also complements existing literature that mainly use immigration waves as the source of exogenous increases in low-skill labor (e.g. Lewis, 2011).

The Chinese college enrollment expansion program was implemented by the Chinese central government in 1999 and became one of the largest expansions of post-secondary education in the world.[3] By 2009, annual post-secondary school enrollment increased six-fold, from 1 million students to 6 million, and consequently, the share of college educated population doubled by 2010 (Table 1). The college wage premium remained stable despite the sharp and persistent increase in the relative supply of college-educated workers (Figs. 1 and 2). This suggests that a rightward shift occurred in the labor demand curve and motivates us to consider a model that incorporates firm-level endogenous adjustments in labor demand when local labor force skill mix changes.

We develop a framework with the coexistence of two types of production technologies, characterized by two different types of capitals, one skill-biased and the other labor-biased. Our model is an extension of the models in Beaudry, Doms, and Lewis (2010) and Lewis (2013), which include two types of capital (skill-biased and labor-biased) and two types of labor (high-skill and low-skill).[4] In our model, firms employ both high- and low-skill workers and choose either type of capital to produce the same output across two types of firms. To achieve a new equilibrium, firms adjust every input when the local labor market skill-mix changes. As high-skill labor becomes more abundant, firms using skill-biased capital invest more in capital and hire more workers, while firms using labor-biased capital reduce capital investment and workforce size as the relative supply of low-skill workers decreases. Therefore, our model suggests that endogenous technological change, in the aggregate, would have similar prediction to that of endogenous technological adoptions, in the sense that skill-biased technology becomes more prevalent as skilled-labor increase, even though the underlying mechanisms can be different.

Our model predicts that college wage premiums would not change as labor market skill mix changes, as the endogenous shifts in labor demand exactly compensate for the increases in skilled labor. Note that the Rybczynski theorem of the Heckscher-Ohlin trade model also predicts that changes in the relative supply of college graduates have no effect on city-specific relative wages, as long as firms change their output mix to fully absorb any endowment shock. However, the mechanism described in our model does not requires adjustments in output. Our mechanism for labor market adjustment differs and complements the Rybczynski-type output adjustments, yet which mechanisms are more important is an empirical question and the answer may vary by context.

To test our model empirically, we treat cities as individual labor markets and use variations in the college enrollment expansion of each city to identify shocks in the supply of college-educated workers. More specifically, we construct an instrumental variable for the local skill mix based on a city's historical, pre-policy share of college enrollment. The validity of our instrumental-variable strategy depends on this variable being a good predictor of future skill-mix of the labor force and its being uncorrelated with future labor demand shocks. With respect to the first point, the initial percentage of college enrollment of a city strongly predicts the local college-educated population share after a ten-year implementation of college expansion program, as the new generation of college graduates heavily affect the skill-mix in the labor force. Second, the cross-city variations we use to construct the instrument are based on city-level average college enrollment share before the policy shock from 1995 to 1997, and these initial enrollment shares are conditionally uncorrelated with city-level skill mix at the base year, largely due to the centralized higher education system in place till mid-1990s.[5]

Using the firm-level panel data on all large manufacturing firms in China from 1998 to 2008, we show that firms hiring more high-skill labor, which are more likely to use skill-biased capital in production, are more likely to increase their capital, R&D spending, and employment when the local share of skilled labor increases. We control for confounding factors such as state-owned enterprise (SOE) reform, rural-to-urban migration, and industry-specific time trends that might have been influenced by China's joining the World Trade Organization (WTO) in 2001. Using worker-level wage data from the Urban Household Survey (UHS), we verify that the estimate for the slope of the relative labor demand, or the inverse elasticity of substitution between college- and high-school-educated workers, is not statistically different from zero, consistent with our model's prediction.

Previous evaluations of China's college enrollment expansion program mainly use a difference-in-differences design that cannot rule out industry-specific time trends. For example, Li, Whalley, and Xing (2014) and Li, Ma, Meng, Qiao, and Shi (2017) describe how the college expansion program changes the return to education; Che and Zhang (2018) study the growth of firms' total factor productivity after the college enrollment expansion program. Our study differs from Che and Zhang (2018) mainly because we use

[3] We calculate that China's college expansion program generates an average increase of 0.6 year of education for cohorts born in 1990 compared with cohorts born in 1980.

[4] We use a nested CES production function which is a more general function form than that in Beaudry et al. (2010). Nevertheless, all predictions of the model are unchanged if we use exactly the same functional form as that in Beaudry et al. (2010).

[5] Till mid-1990s, jobs for college graduates are allocated by the MOE and related government organizations. The majority of college graduates had to return to work in their original province regardless of where they attended college, which explains the lack of correlation between a college's enrollment share of a city and the percentage of college-educated workers in that city.

Table 1
Percentage of college-educated population in Chinese Census.

Year	Age 20–40	Age 15–64
	(1)	(2)
All cities		
	6.9%	5.2%
	11.1%	7.7%
	18.9%	11.9%
By city type		
Municipality		
2000	14.5%	11.5%
2005	25.9%	17.8%
2010	42.5%	29.4%
Provincial capital		
2000	14.5%	11.4%
2005	19.4%	14.24%
2010	33.8%	23.0%
Other cities		
2000	5.2%	3.8%
2005	7.8%	5.3%
2010	14.3%	8.9%
By region		
Eastern		
2000	7.9%	6.0%
2005	14.2%	9.8%
2010	21.5%	13.7%
Middle		
2000	6.4%	4.8%
2005	9.2%	6.4%
2010	16.9%	10.6%
Western		
2000	5.8%	4.3%
2005	8.5%	5.9%
2010	16.3%	10.2%

Note: Our calculation is based on micro samples from Census 2000, 2010 and micro-Census 2005. The sample size in terms of population share is 0.95% for 2000, 0.2% for 2005 and 0.1% for 2010.

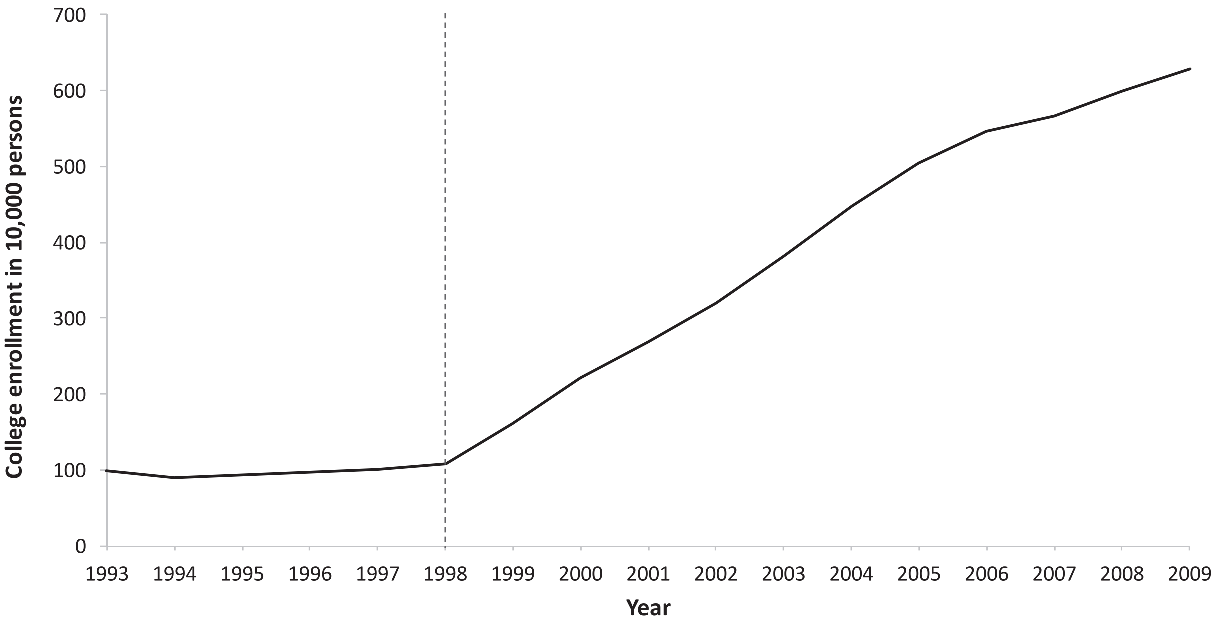

Fig. 1. Trend of college enrollment in China from 1993 to 2009. The data is from the annual statistical yearbook published by the NBS.

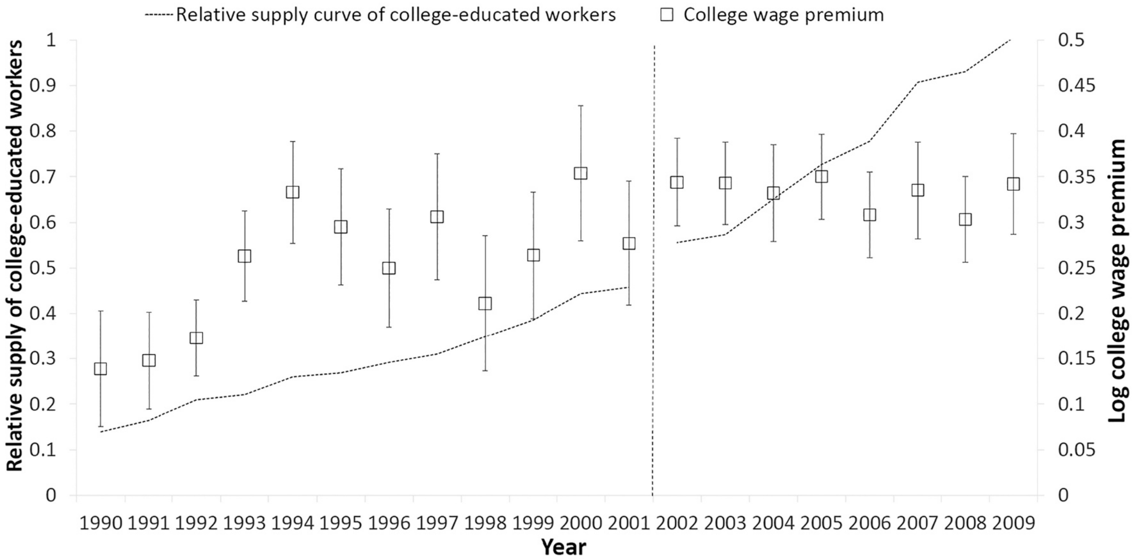

Fig. 2. The trend of college wage premium of workers age between 20 and 40 after the increase in relative labor supply. We calculate the college wage premium using individuals' wage and education achievement data in the UHS.

regional variations in college enrollment and firm level measures of skill intensity.

By examining one of the largest educational promotion programs, our study adds more micro-foundations to the "human capital engine" theory (Lucas, 1988). Instead of the cross-country comparison approach, we adopt a local labor market approach that is less limited by reverse causality.[6] In particular, to explain China's phenomenal growth in the past several decades, existing literature emphasizes drivers such as the reallocation of resources from the public sector to the private sector (Song, Storesletten, & Zilibotti, 2011), access to international trade (Brandt, Van Biesebroeck, Wang, & Zhang, 2017), capital formation (Chow, 1993), and institutional reform (Xu, 2011). We suggest a human capital channel that has been overlooked previously. Our results suggest that relaxing the supply constraints of college-educated workers could have triggered firm-level technology adjustment and economic growth.[7]

Our study also contributes to the understanding of the general equilibrium effects of educational investment. While a vast amount of literature estimates the returns to education in various countries (see a survey in Card, 1999), there are far fewer studies that discuss the general equilibrium effects of education and even fewer that study countries other than the United States (studies on human capital spillovers include Moretti, 2004a, 2004b, Ciccone & Peri, 2006, and Iranzo & Peri, 2009). Our results show that endogenous labor demand shifts may offset the wage effects of the supply increase, which provides an alternative channel to understanding the general equilibrium effects of educational promotion programs.

The remainder of this paper is organized as follows. Section II discusses the policy background and examines aggregate college wage premium trends before and after the college enrollment expansion. Section III outlines a stylized model that captures a firm's endogenous labor demand adjustment. Sections IV and V discuss data and our empirical strategies, followed by our empirical results in Section VI. Section VII concludes our paper.

2. Policy background

Each higher-education institution's enrollment in China is managed by the Ministry of Education (MOE), a branch of the Chinese government that falls below the State Council. However, a policy change originating in the State Council in 1999 started the college expansion program, which deviated from the MOE's five-year plan for 1996–2000 (and since the State Council manages the MOE, the State Council's policy took precedence over the MOE plan).[8] To achieve the ambitious goal set by the State Council, the MOE modified its college education enrollment plan abruptly. On June 24, 1999, only two weeks before the national college entrance exam, the MOE

[6] Examples of cross-country comparisons include Benhabib and Spiegel (1994), Barro and Lee (1994), Gregorio and Lee (2002), Barro, Caselli, and Lee (2013), and Gennaioli, Porta, Lopez-de-Silanes, and Shleifer (2013). Critiques of the cross-country comparison approach include Bils and Klenow (2000), Chevalier, Harmon, Walker, and Zhu (2004), and de la Fuente and Doménech (2006).

[7] From Chinese import competition, there can also be technology spillover effects in other countries as studied by Bloom, Draca, and Van Reenen (2016).

[8] Based on the "Blueprint for the Ninth Five-Year Plan for Education" published in 1996, the MOE originally planned to raise the gross college enrollment rate (college enrollment among those in the same age group) from 6.5 to 8% by 2000. However, in December of 1998, the State Council released a special issue of the policy book "Plan to Revitalize Education for the 21st Century," which sought to raise the target gross college admission rate further to 11%, tripling the original growth goal year-over-year from 1.5 to 4.5%.

announced a 38% increase in the admissions quota. Since then, college enrollment has increased dramatically in China.

The MOE sets a binding enrollment quota for each college every year. If a college plans to make any adjustment, it needs to seek approval from its provincial government and the MOE. As each college follows the pre-assigned quota closely, the MOE plays the main role in determining the total enrollment of a college and therefore the total college enrollment of the country. To formulate the enrollment quota for each college, the MOE issues a plan every five years that outlines the growth of college enrollment for an upcoming five-year time frame. To do this, the MOE audits the on-campus infrastructure at each of the country's colleges and then constructs benchmark enrollment quotas accordingly. For example, the MOE used the auditing results from 1996 to compute the benchmark college enrollment quotas for each college from 1997 to 2004. The MOE sets annual enrollments based on the benchmark quotas and the country's overall growth plan. After the State Council unexpectedly raised the target gross college admission rate in 1999, the MOE did not revise the benchmark enrollment quota but only increased the growth rate of annual enrollment at the national level.

Fig. 1 shows that the annual college admission growth rate was 2% on average between 1993 and 1998 based on the data published by the National Bureau of Statistics (NBS) of China. However, from 1999 to 2001 it suddenly increased by 30% per year. It slowed to a 15% annual growth rate in the following 5 years. As a result, the average proportion of population with any college education between 20 and 40 years of age in China increased from 6.9% in 2000 to 18.9% in 2010, according to the Table 1.

How does the enlarged college-educated workforce affect the equilibrium wage in the labor market? To document the change in the college wage premium at the national level, we run an ordinary least squares (OLS) regression using the UHS data from 1990 to 2009. This allows us to estimate the college wage premium each year and the results of this can be seen in Fig. 2.[9] We restrict the sample to workers between 20 and 40 years of age at the time of the survey because workers in this age group have similar work experiences and they were directly affected by the college expansion program.[10] In Fig. 2, the squares represent the point estimates for the estimated college wage premium for workers aged 20–40. During this period, the increase in the supply of college-educated workers accelerated. However, the average college wage premium remained steady. Also of note is that unemployment rate of college-educated workers decreased compared to workers without any college education (see Feng, Yingyao, & Moffitt, 2017 and Fig. A1).

This college expansion program also increased variation in the skill mix across cities. Although the MOE did not publish details of how the college expansion program was implemented, the data suggests that the MOE likely increased each college's enrollment quota using a national growth rate. Fig. 3 shows that changes in a city's college enrollment between 1998 and 2009 is highly correlated to the city's initial college enrollment in 1998. Therefore, cities that had higher college enrollments before 1999 were likely to have faced larger influxes of college-educated workers compared to others after the college expansion program.

While the relative supply of college-educated workers quadrupled from 1990 to 2009 for 20 to 40-year-olds based on UHS data, the college wage premium during this period did not decrease. The national acceptance rate of the college entrance exam nearly doubled from 34% in 1998 to 62% in 2009. This change in selection cutoff decreases the average unobserved abilities of college graduates (e.g., Juhn, Kim, & Vella, 2005). As colleges had insufficient infrastructure in the first few years to accommodate so many new students, the quality of college education could decline during the beginning of the college expansion program (see Meng, Shen, & Xue, 2013). Overall, changes in students' unobserved abilities or quality of college education during this period were likely to decrease rather than increase the college wage premium holding the relative labor demand constant.

Our explanation for the observed college wage premium trend is a shift in the labor demand that was likely caused by firms' responses to local labor force skill mix changes. In the following sections, we develop a model with two production technologies to illustrate how firms adjust labor demand when faced with an exogenous change in the skill mix of a local labor market.

3. Model

There are four possible elements in a production function: low-skill labor (L), high-skill labor (H), skill-complementing capital (K_h), and labor-complementing capital (K_l). We define workers who received any college education as high-skill labor, and workers who have received high-school education or below as low-skill labor. We define capital in our model as a generalized machinery input in the production function, which can be measured by the total value of capital in firm-balance sheet or other related variables such as R&D expenditure. Motivated by observed heterogeneity in production technology among firms in a large developing country like China, we assume that there are two types of firms: firms that use skill-biased capital and firms that use labor-biased capital. Both types of firms employ a mix of high- and low-skill labor.

Consider an environment in which two types of firms both produce one final good denoted by Y. Each type of production technology follows a nested constant elasticity of substitution (CES) production function (see Eqs. (1)–(2)), which has been widely adopted in economics literature since Sato (1967). In particular, Y_1 equals the amount of output produced by firms using technology 1 and Y_2 represents the total output of firms using technology 2.

[9] The dependent variable is the log of a worker's annual earnings in each year, which consists of basic wages, bonuses, subsidies, and other labor-related income. The explanatory variables include a dummy variable that indicates whether a worker is college educated and other control variables that include demographics, provincial-fixed effects, and years of experience by education levels.

[10] Li et al. (2017) suggest that the college expansion program has different effects on labor market performance across workers in different age groups. We also verify that the pattern holds for college wage premium of workers aged 15–64.

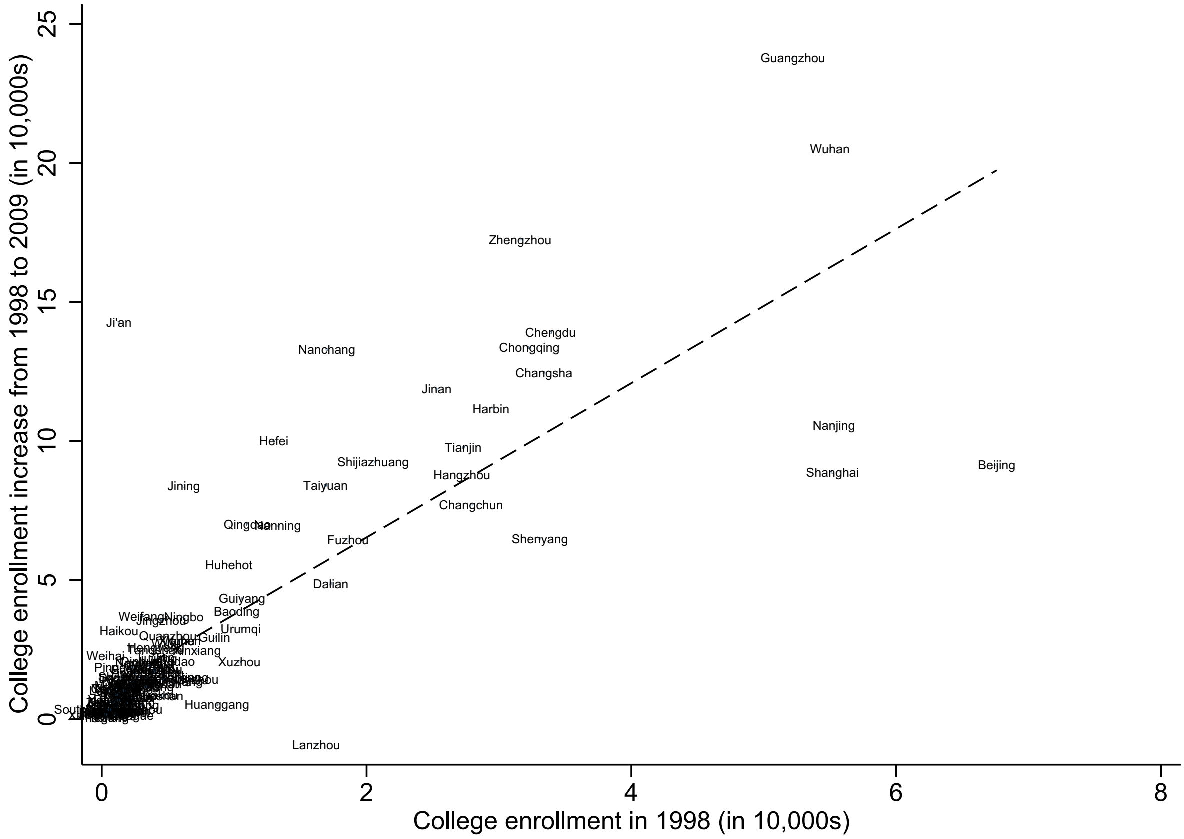

Fig. 3. The increases of college enrollments in each city from 1998 to 2009 strongly relates to cities' college enrollments in 1998. The data is from the Statistical Yearbook of Regional Economy.

$$Y_1 = \left[\alpha_1 (K_h^{\rho_1} + L_1^{\rho_1})^{\frac{\mu}{\rho_1}} + (1-\alpha_1) H_1^{\mu}\right]^{\frac{1}{\mu}}, \quad 0 < \alpha_1 < 1, \quad \rho_1, \mu \le 1 \tag{1}$$

$$Y_2 = \left[\alpha_2 (K_l^{\rho_2} + H_2^{\rho_2})^{\frac{\mu}{\rho_2}} + (1-\alpha_2) L_2^{\mu}\right]^{\frac{1}{\mu}}, \quad 0 < \alpha_2 < 1, \quad \rho_2, \mu \le 1 \tag{2}$$

In these production functions, the elasticity of substitution between labor and capital is fixed. For example, the elasticity of substitution between L_1 and K_h is $\frac{\mu}{1-\rho_1}$ under the first production technology. When ρ_1 is equal to 1, L_1 and K_h are perfect substitutes, see example in Autor, Levy, and Murnane (2003). When ρ_1 is equal to μ, capital is skill-neutral, i.e. the elasticity of substitution between L_1 and K_h is equal to that between H_1 and K_h. If $\rho_1 > \mu$, capital is generally a substitute to low-skill labor and a complement to high-skill labor.

Capital is not always skill complementing. Acemoglu (2002) summarizes studies such as Goldin and Katz (1998) on how machines invented in the 19th Century were non-skill-biased. An example from today's economy is that sewing robots and sewing machines are simultaneously used in the textile industry. Sewing robots require high-skill workers to program designs while sewing machines simply assist low-skill workers in making clothes. In this example, the second type of production technology, sewing machines, is an illustration of capital (K_l) as a complement to low-skill labor (thus $\rho_2 > \mu$), as opposed to the more general assumption that capital only complements to high-skill labor.

In this economy, we further assume that the price of capital is exogenous (denoted as r_h and r_l); and the sum of each type of labor is exogenous ($H_1 + H_2 = H$ and $L_1 + L_2 = L$). Both types of firms maximize their profits by choosing the optimal amount of capital and labor. As the production functions in our model satisfy constant returns to scale, we can normalize the number of firms to be 1 under each type of production technology. In this case, a price-taking firm aims to maximize profits by solving the problems in Eqs. (3)–(4) after we normalize the product price to be 1 as well.

$$max\,\Pi_1 = Y_1 - w_H H_1 - w_L L_1 - r_h K_h \tag{3}$$

$$max\,\Pi_2 = Y_2 - w_H H_2 - w_L L_2 - r_l K_l \tag{4}$$

We define a competitive equilibrium as a situation when firms maximize profit, labor markets clear, and each type of firm earns the same profit for a given output price and input price. The constant returns to scale of the production function imply that each type of firm earns zero profit. In equilibrium, we can solve how a firm allocates inputs by taking the first-order derivative of a firm's profit with respect to each input. In particular, we propose the following:

Proposition 1. When firms can adjust capital freely and the range of parameters allows an interior solution for the equilibrium, the relative wage of high-skill workers compared to that of low-skill workers, $ln\left(\frac{w_H}{w_L}\right)$, only depends on exogenous production function parameters (μ, α_1, α_2, ρ^h and ρ^l) and capital rental price (r_h, r_l). See proof in Appendix A1.

To see the adjustment of relative wages of high- and low-skill workers in each type of firm, we can focus on the first-order conditions summarized in Eq. (5).

$$\begin{aligned} ln\left(\frac{w_H}{w_L}\right) &= ln\frac{(1-\alpha_1)}{\alpha_1} + \left(1-\frac{\mu}{\rho_1}\right)ln\left(1+\frac{K_h^{\rho_1}}{L_1^{\rho_1}}\right) + (\mu-1)ln\frac{H_1}{L_1} \\ &= ln\frac{\alpha_2}{(1-\alpha_2)} + \left(\frac{\mu}{\rho_2}-1\right)ln\left(1+\frac{K_l^{\rho_2}}{H_2^{\rho_2}}\right) + (\mu-1)ln\frac{H_2}{L_2} \end{aligned} \tag{5}$$

Suppose firms do not adjust capital (i.e. K_h and K_l are fixed). According to Eq. (5), the relative wage of high-skill workers compared to that of low-skill workers, $ln\left(\frac{w_H}{w_L}\right)$, decreases as the relative supply of college-educated workers ($ln\frac{H_1}{L_1}$ or $ln\frac{H_2}{L_2}$) increases when $\mu < 1$. However, if the price of the capital is constant and firms can adjust the quantity of capital freely, Proposition 1 shows that the relative wage keeps constant regardless of skill-mix changes under both technologies.

While our model is an extension of Beaudry et al. (2010) and Lewis (2011), it is different in three aspects: 1) Lewis (2011) points out skill-capital complementarity can explain why low-skill immigration has little effect on wages. Our model, however, exhibits that adjustment in both skill- or labor-biased capital can explain the observed constant wage premium after relative labor supply shocks; 2) our model allows for the co-existence of different types of production technologies and predicts a zero response in college wage premium instead of a diminished response as that in Beaudry et al. (2010)[11]; 3) in this model, we assume that firms are endowed with different production technologies, whereas firms in Beaudry et al. (2010) and Lewis (2011) are homogeneous.

In addition, Eq. (5) indicates that a sufficient condition for $\frac{H_1}{L_1} > \frac{H_2}{L_2}$ is $\alpha_1 + \alpha_2 < 1$. This is aligned with previous estimates. For example, Stokey (1996) finds $\alpha_1 = 0.38$ and Duffy and Papageorgiou (2000) use $\alpha_1 = \alpha_2 = 0.4$ in the estimation of nested CES production functions. This observation provides a method for how to identify firms that are using skill- or labor-biased capital and motivates our next proposition.

Proposition 2. When more high-skill workers enter the labor market, firms that use skill-biased capital hire more workers and invest more in capital. However, the capital intensity measured by the value of capital per worker remains constant among both types of firms. For proof, see Appendix A2.

4. Data

The main data sets used in this study are micro-level surveys conducted by the National Bureau of Statistics (NBS) of China. The first data set is the Annual Survey of Industrial Firms (ASIF), which consists of over 2 million observations of large domestic- and foreign-invested manufacturing firms in China from 1998 to 2008.[12] The second data set combines the micro sample of Census 2000 and 2010 and the microcensus in 2005 (an intercensal survey to measure the population). The micro samples include 0.95% of total population for 2000, 0.2% of total population for 2005 and 0.1% of total population for 2010, which provide relatively accurate measures of city-level skill-mix. The third data set comes from the Urban Household Survey (UHS), which is a repeated cross-sectional survey similar to the March Current Population Survey Income Supplement in the United States. We have access to the UHS data from 1990 to 2009, which provides the additional information on the urban labor market in China. By combining all these data, we can connect large manufacturing firms with local labor-force compositions before and after the college expansion program.

The micro sample of Census record detailed information about school completion levels and residential locations. The education levels of workers are grouped into the categories of "high school and below" and "college or above."[13] We restrict our microcensus sample to workers who are between 20 and 40 years old to calculate each city's share of college-educated workers at the prefecture-level. We list the summary statistics in Table 1. The average proportion of population with any college education increased from 6.9% to 18.9% during the period between 2000 and 2010, and that share among population between 15 and 64 years of age increased from 5.2% to 11.9%.[14]

[11] As a robustness check, we use the production functions in Beaudry et al. (2010) (both firms use skill-biased capital but with different substitution elasticities) to replace our Eqs. (1–2) and we obtain the same set of results.

[12] Examples of recent works that use the ASIF data include Brandt et al. (2017) and Che and Zhang (2018).

[13] The category of "college or above" includes workers who have received any post-secondary education. Note that previous literature may define the college wage premium as the wage of college graduates relative to wages of high school graduates, we choose our definition of college wage premium due to the sparse information on education attainment in the Census.

[14] Part of the observed change may be explained by the lower levels of education achieved by people who were college-aged during the Cultural Revolution. During that time, college enrollment was nearly halted.

The ASIF contains information on the income statements and balance sheets of each large manufacturing firm in China from 1998 to 2008. The ASIF includes all firms that are either state-owned or are non-state firms with current-year sales over CNY 5 million in the manufacturing sector. Our unit of analysis in this data set is at a plant level, as we observe one firm having multiple plants in different cities. Among all the firms, more than 90% of all observations in our sample are single-plant firms. For firms with multiple plants, we treat each plant as an independent unit and assign them separate IDs. The definition of unit ID is by a firm's tax filer number and location code (or 6-digit administration code). For simplicity of notation, we refer the unit of analysis as a "firm" regardless whether it is from single- or multiple-plant firms. Empirically, most of our observations in the data are at a firm-level. To measure firms' real capital stock, we follow Brandt, Van Biesebroeck, and Zhang (2012) by using the perpetual inventory method. We exclude firms with missing, zero, and negative values for real capital stock and employment, as well as firms with fewer than eight employees (such firms are considered individually-owned businesses in China). We list the summary statistics in Table 2.

We then drop firms that did not enter the ASIF in 2004. We impose this restriction as firm-level skill mix is crucial for our analysis and we can only track the skill-mix information of any firm in the ASIF that existed in 2004 using the 2004 Economy Census. Table A1 lists the summary statistics for all firms in our analysis sample. Ideally, we hope to find a skill intensity measure that was collected before 1999, but the only available data on skill intensity at firm level in China is from 2004 Economy Census, at least to the knowledge of the authors. The college enrolment expansion started in 1999 in China and it often takes 3–4 years for enrolled students to graduate, we therefore expect that firm-skill mix has not changed a lot in 2004. All of the firms surveyed by the ASIF are quite large in scale of employment due to the design of the data collection. In our sample of interest, the annual real value added is CNY 33 million (as measured in 1998) and the average workforce size is 289.

As supplementary data, we use city-level outcomes that from the statistical yearbooks published by the NBS. The city-level college enrollment and graduate data are from the "Statistical Yearbook of Regional Economy". One control we have is based on the city macroeconomic indicator data from the "City Statistical Yearbooks". In addition, we use micro-level data from the 1995 and 2004 Economy Censuses. We also use export and import records from the General Administration of Customs (GAC) from 1998 to 2008 to provide additional controls and tests.

To capture the likelihood of a firm's use of skill- or labor-biased capital (K_1 or K_2), we use the share of college-educated workers in a firm based on the 2004 Economy Census. The overall share of high-skill workers among Chinese manufacturing firms was approximately 12% in 2004 and there is substantial variation in the firm-level college worker ratio within an industry according to the Economy Census. We use the ratio of high-skill over low-skill workers in 2004 Economy Census as the measure of skill intensity (*SI*). Table A3 lists the raw and the standardized *SI* at 2-digit industry level. Fig. A2 gives examples of skill-intensity heterogeneity within an industry.

5. Empirical strategy

Suppose each city is a separate labor market. The market for high- and low-skill labor is purely local, with exogenously fixed local supplies that vary across cities and time.[15] To test whether firms invest in capital differently after a skill-mix shock, we introduce the regression in Eq. (6). Consider a basic firm-level equation for the logged real value of capital K_{ict} in firm i of industry j in city c at time t as:

$$K_{ict} = \beta_i R_{ct} + \eta_i + \gamma_1 O_{it} + \gamma_2 C_{ct} + \epsilon_{ict} \tag{6}$$

where $R_{ct} = log\frac{H_{ct}}{L_{ct}}$ represents the city-level skill mix, and the capital, K_{ict} is interpreted broadly as production technology which can be measured by a firm's value of capital or R&D expenditure. We measure city level skill-mix, R_{ct}, using the logged ratio of workers with college or above education over workers with high school and below education. The η_i represents a firm's fixed but unobservable characteristics, which would be absorbed after taking the difference in a firm's outcome between two periods. To control for the effects of SOE reform in China, we add a time-specific ownership dummy (O_{it}).[16] The city economic indicators (C_{ct}) help absorb macroeconomic shocks at the city level. Finally, to cancel out firm fixed effects (η_i), we use a long-difference regression to test our model's prediction, where Δ represents the long difference operator.

$$\Delta K_{ict} = \beta_i \Delta R_{ct} + \gamma_1 \Delta O_i + \gamma_2 \Delta C_c + \Delta\epsilon_{ict} \tag{7}$$

Our model predicts that the sign of β_i depends on the production technology a firm uses. From a firm's high level of SI_i, we can infer that this firm is more likely to use more skill-complementary capital. Therefore, we use a parametric approximation to estimate firm specific responses to changes in a city's skill mix, β_i:

$$\beta_i = \beta_0 + \beta_1 SI_i \tag{8}$$

We re-write Eq. (7) as Eq. (9) to test our model's prediction in the data.

[15] The Chinese residency-permit system (hukou) largely restricts the inter-city labor mobility, which helps to justify our identification assumption.

[16] We include five ownership categories: SOEs; domestically private-owned enterprises or enterprises owned by shareholders; foreign invested enterprises; enterprises owned by investors from Hong Kong, Macau, or Taiwan; and collectively owned enterprises and enterprises with hybrid ownership.

Table 2
Summary statistics of the ASIF firm sample – panel data.

Analysis sample 1998–2008	(1) Obs	(2) Mean	(3) S.D.
State-owned enterprise ratio	1,050,881	0.145	0.352
Real value-added (millions)	934,459	33.02	315.75
Real capital (millions)	1,050,881	34.89	423.20
R&D expenditure (millions)	855,937	1.73	51.59
Average number of employees	1,050,881	288.64	1157.79
Annual salary per employee	1,049,382	16,233	74,929

Note: All value variables are measured in 1998 CNY. The sample includes firms that with 2-digit industry code range from 13 to 42 according to the NBS industry classification from 1998 to 2008 ASIF data. We drop firms that did not enter the ASIF in 2004, as well as firms with missing, zero, and negative values for real capital stock and employment,and firms with fewer than eight employees.

$$\Delta K_{ict} = \beta_0 \Delta R_{ct} + \beta_1 \Delta R_{ct} \times SI_i + \gamma_1 \Delta O_i + \gamma_2 \Delta C_c + \Delta \epsilon_{ict} \quad (9)$$

The coefficient of the interaction term between SI_i and change in R_{ct} shows if firms that use skill-biased capital invest more in capital after an influx of college-educated workers in a local labor market. The model's prediction is reflected in the positive coefficient of the interaction term, β_1.

The main challenge for estimating Eq. (9) is that the change in a city's skill mix (ΔR_{ct}) can be related to unobserved shocks $\Delta \epsilon_{ict}$. For example, if college workers are more likely to migrate to cities with higher labor demand, the OLS estimates would be upward biased as cities with higher share of skilled workers are just booming economically. Another example is that if more firms enter or expand in a city with high skilled worker ratio but the Hukou system—acting as worker's mobility constraint—discourages entry of workers, then OLS estimates would be downward biased. We thus use an instrumental variable that is based on college enrollment expansion to isolate the exogenous shock in a city's skill mix. The instrumental variable, ΔS_{ct}, uses the changes in predicted college graduates divided by base-year population (P_{t0}). One year's predicted college enrollment, $E_{ct} = N_t \eta_c$, equals the national college enrollment in year t, N_t, multiplied by a city's average share of the national college enrollment from 1995 to 1997 (η_c).[17] Eq. (11) describes how we construct the instrument variable, ΔS_{ct}, for the change in city-level skill-mix, ΔR_{ct}.

$$E_{ct} = \eta_c \cdot N_{t-4} \quad (10)$$

$$\sigma_{ct} = \frac{\Sigma_{t_0}^{t} E_{ct}}{P_{t_0}}$$

$$\Delta S_{ct} = \Delta log \frac{\sigma_{ct}}{1 - \sigma_{ct}} \quad (11)$$

Fig. A3 shows that a city's college enrollment share stays steady over time, which explains why our predicted college enrollment matches well with a city's actual college enrollment. Our instrument is different from a conventional shift-share instrument because we do not assign future college graduates to cities where college-educated workers were located before the college expansion program. Nevertheless, we impose the assumption that a city's initial college enrollment share of the national college enrollment is randomly assigned after controlling for city observable characteristics, i.e. $Cov(\Delta \epsilon_{ict}, \eta_c | C_{ct}) = 0$. We further discuss the plausibility of this assumption in below, which also relates to a recent discussion on shift-share instrument by Goldsmith-Pinkham, Sorkin, and Swift (2020).

First, to find out which city level controls (C_{ct}) we need to control for, we examine which city characteristics are likely to correlate with a city's pre-policy college enrollment share. According to Table 3, a city's type — whether it is a municipality that is directly overseen by the federal government, a provincial capital, or another type of prefecture-level city – significantly influences the a city's college enrollment numbers.[18] A city's GDP positively correlate with a city's enrollment share, but the city's migrant population and average wage does not strongly correlate with a city's college enrollment share.

Notice that while the number of each city's newly added college graduates is proportional to this city's base year college graduates before the college expansion program, the changes in local skill mix brought by the college enrollment expansion do not follow the same proportion across cities for several reasons. First, each city's initial level skill mix are different, thus the accumulated effect of college expansion varies across cities. Second, as college graduates were assigned to the work place by the government before 1997, a city's skill mix in 1998 does not strongly correlate with its historical enrollment share of that city once controlling for city types and regions (shown in Table 3).

We then conduct the following validity tests to examine our instrument. In our firm-level long difference regressions, we add city type, region dummies and the base year level GDP as controls to address the concern that local governments may impose industrial

[17] The city-level college enrollment and graduation data in our study comes from the Statistical Yearbook of Regional Economy from 1999 to 2009. We use a 4-year lag to recover a city's college enrollment in 1998 and we verify that the college graduation rate in China is approximately 100%. We also use the number of colleges established as of 1998 to predict the enrollment share of each city in 1998, which generates similar results.

[18] To follow the higher education model of the Soviet Union, there was a major integration and reorganization through building and moving of all universities in 1952 which largely determined the location of the majority of current universities to be at capital cities.

Table 3
A city's initial share of national college enrollment in 1995–1997.

Dependent variable	Initial share of national college enrollment			
	(1)	(2)	(3)	(4)
Relative supply of skilled workers in 1998	0.002***	0.001	0.001	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
Municipality	0.032***	0.028***	0.028***	0.027***
	[0.007]	[0.007]	[0.007]	[0.007]
Provincial capital	0.014***	0.012***	0.012***	0.011***
	[0.002]	[0.002]	[0.002]	[0.002]
Middle region	0.001**	0.001*	0.001*	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
Eastern region	0.002***	0.001	0.001	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
−1997 average GDP		0.002***	0.002***	0.002***
		[0.001]	[0.001]	[0.001]
migrant population			0.000	0.000
			[0.000]	[0.000]
−1997 city average wage				−0.002*
				[0.001]
Observations	274	241	241	237
R-squared	0.739	0.765	0.765	0.768

Note: The data source is China City Statistical Yearbook. The independent variables include GDP, migrant population, and city average wage, which are measured in a logarithm scale. Robust standard errors are shown in brackets. ***p<0.01, **p<0.05, *p<0.1.

policies or provide support that targets provincial capitals, coastal cities, or cities with higher GDP. Another concern is that cities with higher baseline enrollment shares have been more exposed to trade since China entered the WTO in 2001. Therefore, we add city-level export volume by destination country's WTO membership as a control.

We also test our instrument's validity by conducting a pre-trend regression. We add data from the 1995 Economy Census to supplement the firm-level observations in pre-policy periods, and then carry our main specification by using $S_{ct'}$ at future periods to replace the R_{ct} in Eq. (9). For example, we can regress a firm's capital change from 1995 to 1999 on the change in S_{ct} from 2004 to 2008 using the same set of other control variables as that in Eq. (9). We will discuss the empirical result of the pre-trend test in the next section.

Our policy-induced instrumental variable is similar to that in Fortin (2006). To the extent that past college enrollment is exogenous to current demand, two-stage least squares (2SLS) estimates should consistently estimate the supply shock. In Table 4, we run the following regression (Eq. (12)) to examine whether the instrumental variable (ΔS_{ct}) has a strong correlation with the city level skill mix.

$$\Delta R_{ct} = \delta \Delta S_{ct} + \theta X_{ct}^{'} + \epsilon_{ct}. \tag{12}$$

The control variable list X_{ct} includes municipality, provincial capital, and region dummies, and a city's change in GDP.

According to Equation Eq. (9), the change in a city's skill mix (ΔR_{ct}) should be measured in the same time frame as the change in firm-level outcomes (ΔK_{ict}), which span from 1998 to 2008. However, as the most accurate city-level skill mix information is from the 2000 and 2010 censuses, we use a city's change in skill mix from 2000 through 2010 as a proxy for the change that city experienced from 1998 to 2008. In Table 4, we show that the change in a city's actual skill mix as measured in the Census 2000 and 2010, ΔR_{ct}, strongly correlates with the predicted skill mix from 1998 to 2008, ΔS_{ct}. We use the time window from 1998 to 2008 for ΔS_{ct} to match with firm-level outcome variables. The correlation does not weaken significantly after we control for the city characteristics. The coefficient of S_{ct} shows that a 10% increase in the predicted relative supply of college graduates predicts a 5% increase in a city's relative supply of college-educated workers whose age are between 20 and 40.[19] The attrition of students migrating from the city where they attended college only attenuates our first-stage estimate towards 0.

6. Results

In this section, we first verify our Proposition 1 by estimating the relative labor demand for college-educated workers using wage data from the UHS. We then present evidence to support Proposition 2, showing that after an influx in college-educated works, firms using skill-biased capital increase capital stock, R&D expenditure and employment compared to other firms. By using other sources of firm-level data, we find that confounding factors, such as China's joining WTO, SOE reform or rural-to-urban migration, are not the

[19] We collect the placement reports of 116 universities (top-tier universities in China) in "Project 211" of the MOE in 2014 to estimate how likely students are to work in the city where they attended college relative to other cities. Of the 48 universities with city-level placement information, 46.5% of students work in the same city after graduation. For the other 28 universities with provincial-level placement information, we know that 62.6% of students choose to work in the same province after graduation. It appears that the ratio of students who choose to stay local after graduation in the placement reports is consistent with the observed significant first stage regression result.

Table 4
Correlation between actual and predicted skill-mix change based on predicted college graduates.

Dependent variable	Changes in relative supply of skilled workers: ΔR_{ct}					
	Age 20–40			Age 15–64		
	(1)	(2)	(3)	(4)	(5)	(6)
Changes in predicted relative supply: ΔS_{ct}	0.423***	0.533***	0.578***	0.328***	0.355***	0.401***
	[0.099]	[0.116]	[0.110]	[0.125]	[0.135]	[0.121]
Municipality, Capital & Region FE		Yes	Yes		Yes	Yes
Changes in city GDP (in log)			Yes			Yes
Observations	197	197	197	197	197	197
R-squared	0.061	0.110	0.158	0.051	0.095	0.165

Note: ΔR_{ct} is a proxy for changes in a city's logged ratio of college- to non-college-educated workers 1998–2008 based on Census 2000 & 2010. ΔS_{ct} is change in the ratio of predicted college graduates share to the share of those without college education from 1998 to 2008. Robust standard errors are in brackets. ***p<0.01, **p<0.05, *p<0.1.

cause of the change.

We use worker-level data to test how changes in skill mix affect the equilibrium wage in the labor market. Based on the UHS wage records, we estimate the relative demand curve at a city level to verify our wage prediction following how Katz and Murphy (1992) estimate the inverse elasticity of substitution between college-educated and high-school and below workers. More specifically, we test if the outward shift in the relative demand for college-educated workers offsets the reduction in the supply of low-skilled workers using UHS data as predicted by our Proposition 1. The regression uses the dependent variable as the change in logged relative wage of college-educated workers among workers aged 20 to 40 between 1998 and 2008 measured by the data from the UHS. The key independent variable is the change in logged relative supply of college-educated workers in the same age group of that city, $log\left(\frac{H_{ct}}{L_{ct}}\right)$, which is measured by the data from Census 2000 and 2010. The additional control variables, noted by X_{ct}, include city types and change in GDP from 1998 to 2008. We use a city's logged predicted relative supply of college-educated workers from 1998 to 2008, $\log S_{ct}$, as the instrument variable.

$$\Delta log\frac{w_{ct}^{h}}{w_{ct}^{l}} = \theta_0 + \theta_1 \Delta log\frac{H_{ct}}{L_{ct}} + \theta_2 X_c + \zeta_{ct} \tag{13}$$

According to Table 5, the point estimate of the slope of relative labor demand (the inverse elasticity of substitution between college-educated and high school-educated workers) is not statistically different from 0 in both OLS ans 2SLS specifications. Overall, our estimates for the inverse elasticity of substitution between college-educated and high school-educated workers are slightly smaller in magnitude than previous estimates in Katz and Murphy (1992), Angrist (1995), Card and Lemieux (2001), and Fortin (2006) (which range from −0.2 to −0.7). However, due to the limited precision in our 2SLS estimation, we cannot rule out typical estimates of the inverse elasticity of substitution between college and non-college workers. Among recent studies, Carneiro, Liu, and Salvanes (2018) study the change of college wage premiums after the construction of new colleges in several cities in Norway in the 1970s. Although their setting is very different from ours, the results regarding changes in college wage premium are similar, and the same pattern in UK has been pointed out by Blundell, Green, and Jin (2018). It appears that all recent estimates on wage premium change is consistent with our claim in Proposition 1.

To test Proposition 2 (firms' have heterogeneous responses to changes in local labor market skill-mix), we define skill intensity (SI_i) using a firm's number of college-educated workers in the 2004 Economy Census divided by the number of high-school-educated workers. We use this proxy because our model predicts that firms using skill-biased capital hire a higher share of college-educated workers, verified by Table A1. In that table, we observe that firms with higher skill intensity are capital intensive and spend more in R&D.

With this proxy of production technology endowment, we use a firm-level long difference regression according to Eq. (9) to examine how city-level skill mix shocks affect firms differently in Table 6. We measure a firm's logged real capital change in the period of 1998 to 2008 as the outcome variable. The coefficient of the interaction between city level skill mix shock ΔR_{ct} and firm skill intensity SI_i measures how much more capital a firm invests if its college over non-college educated workers ratio is one standard deviation higher. Suppose there is a 100% increase in the relative supply of a city's college-educated workers population (which is close to the average change across cities in our sample from 1998 to 2008), the point estimator for β_0 suggests that all firms on average invest 40–70% more capital. The point estimator for β_1 suggests that firms with one standard deviation higher skill intensity will accumulate additional 2.8% more in total capital. Yet, among firms in the high-tech industry (such as information technology equipment), the standard deviation of SI can be as large as 3, so the heterogeneity effect of local labor market skill-mix change is much larger among firms in the high-tech industry. Both OLS and IV estimates in Table 6 are significant and similar in magnitude, and we find that our first-stage firm-level regression has a strong Montiel-Pflueger Effective F statistics.[20]

[20] The first-stage firm-level regression with the interaction term between firm level skill intensity and ΔS_{ct} also has a strong Montiel-Pflueger Effective F statistics.

Table 5
The effect of changes in relative labor supply on college wage premium 1998–2008.

Dependent variable	$\Delta log\frac{w_{ct}^h}{w_{ct}^l}$ in the UHS		ΔR_{ct}
	(1) OLS	(2) IV	(3) F-S
Changes in relative supply: ΔR_{ct}	−0.038	−0.301	
	[0.124]	[0.217]	
Changes in predicted relative supply: ΔS_{ct}			0.647***
			[0.078]
Municipality, Capital & Region FE	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes
Observations	112	112	112
Montiel-Pflueger Effective F stat			68.02

Note: The wage data is based on workers aged 20–40 in the UHS 1998 & 2008. The change in relative labor supply (ΔR_{ct}) is calculated by population age 20–40 in Census 2000 & 2010, which is instrumented by ΔS_{ct} between 1998 and 2008 in column 2. Robust standard errors are shown in brackets. F—S stands for First-Stage. ***p<0.01, **p<0.05, *p<0.1.

Table 6
The effect of a city's skill mix on firms' capital investment 1998–2008.

Dependent variable	Changes in logged value of real capital				ΔR_{ct}
	OLS		2SLS		F-S
	(1)	(2)	(3)	(4)	(5)
Change in relative supply: ΔR_{ct}	0.422***	0.427***	0.739***	0.749***	
	[0.114]	[0.115]	[0.210]	[0.214]	
Firm skill intensity X Change in R_{ct}		0.028***		0.028***	
		[0.004]		[0.004]	
Changes in predicted relative supply: ΔS_{ct}					0.298***
					[0.118]
Municipality, Capital & Region FE	Yes	Yes	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes	Yes	Yes
Changes in ownership category	Yes	Yes	Yes	Yes	Yes
Observations	17,814	17,814	17,814	17,814	17,814
Montiel-Pflueger Effective F stat					35.753

Note: The firm-level real capital is calculated from the ASIF 1998–2008. The change in relative labor supply (ΔR_{ct}) is the change in the skill mix among workers age 20–40 in Census 2000 & 2010, which is instrumented by ΔS_{ct} between 1998 and 2008 in columns 3–4. Firm skill intensity is the standardized ratio of college-educated to non-college-educated workers at firm level in the 2004 Economy Census. Clustered standard errors at cities are shown in brackets. F—S stands for first-stage regression. ***p<0.01, **p<0.05, *p<0.1.

Note that magnitude of $\beta_0$and β_1 cannot be compared directly, because while β_0 only indicates the average effect on capital investment while β_1 shows the marginal effect of one percentage point increase in skill mix ratio across firms with different level of skill intensity. The main message of our result is about the interaction term rather than the scale effect. To rule out that the differential response on skill-intensity is derived from the firm's skill-intensity rather than the skill-intensity of the sector the firm is in, we add 3-digit industry fixed effect as additional control and find that estimates are nearly identical, so the differential response we observe indeed occurred within industry.

We turn to R&D as another measure of skill-biased capital. Table 7 uses the change in firms' R&D expenditures from 2001 to 2007 as a dependent variable to document firms' responses to changes in the skill mix of local labor markets. The ASIF has R&D records from 2001 to 2007. As most manufacturing firms in China are not associated with high-tech, it is not surprising that 59.3% of firms in the ASIF have zero R&D expenditure in any given year. To include all firms with zero expenditure in the regression, we use a firm's R&D expenditure divided by its revenue to calculate the R&D share as the outcome variable in Table 7. We find consistent point estimators for the skill-mix change coefficient for both OLS and 2SLS specifications: the point estimator for β_1 suggests that firms with one standard deviation higher skill intensity will spend additional 0.1% of total revenue in R&D when there is a 100% increase in the relative supply of a city's college-educated workers population. As R&D expenditure takes 5% of total revenue on average, the magnitude of firms' response in R&D is close to that of capital. As R&D expenditure are directly observable indicators of technology adoption (see Machin & Van Reenen, 1998), a city's increase in R&D in response to an increasing share of college-educated workers suggests that firm-level adjustment may include endogenous production technology adoption.[21]

To test our instrumental variable's validity, we use additional records on firms' real capital to run a pre-trend regression. We use the change in the city skill mix in future periods as the key independent variable. The results are presented in Table 8 and we find

[21] Existing theories and evidence on endogenous technological adoptions or innovations include Goldin and Sokoloff (1984), Krusell, Ohanian, Ros-Rull, and Violante (2000), Beaudry et al. (2010), Lewis (2011) and Acemoglu (1998, 2007).

Table 7
The effect of a city's skill mix on firms' R&D expenditure divided by revenue 2001–2007.

Dependent variable	Changes in R&D share				ΔR_{ct}
	OLS		2SLS		F-S
	(1)	(2)	(3)	(4)	(5)
Change in relative supply: ΔR_{ct}	0.002*	0.002**	−0.002	−0.002	
	[0.001]	[0.001]	[0.002]	[0.002]	
Firm skill intensity X Change in R_{ct}		0.001***		0.001***	
		[0.000]		[0.000]	
Changes in predicted relative supply: ΔS_{ct}					0.706***
					[0.114]
Municipality, Capital & Region FE	Yes	Yes	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes	Yes	Yes
Changes in ownership category	Yes	Yes	Yes	Yes	Yes
Observations	17,649	17,649	17,155	17,155	17,155
Montiel-Pflueger Effective F stat					67.459

Note: R&D expenditure share is calculation based on from the ASIF 2001–2007. The change in relative labor supply (ΔR_{ct}) is the change in the skill mix among workers age 20–40 in Census 2000 & 2010, instrumented by ΔS_{ct} between 2001 and 2007 in columns 3–4. Firm skill intensity is the standardized ratio of college-educated to non-college-educated workers at firm level in the 2004 Economy Census. F—S stands for first-stage regression. Clustered standard errors at cities are shown in brackets. ***p<0.01, **p<0.05, *p<0.1.

Table 8
Pre-trend test: the correlation between city's future skill mix and firms' total capital.

Dependent variable	Change in total capital in 1995–1999				$\Delta R_{ct'}$
	OLS		2SLS		F-S
	(1)	(2)	(3)	(4)	(5)
2004–2008 change in relative supply: $\Delta R_{ct'}$	0.027	0.029	−0.005	−0.003	
	[0.035]	[0.035]	[0.084]	[0.085]	
Firm skill intensity X Change in $\Delta R_{ct'}$		0.027		0.026	
		[0.023]		[0.027]	
Changes in predicted relative supply: ΔS_{ct}					0.745***
					[0.229]
Municipality, Capital & Region FE	Yes	Yes	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes	Yes	Yes
Changes in ownership category	Yes	Yes	Yes	Yes	Yes
Observations	11,010	11,010	10,678	10,678	10,678
Montiel-Pflueger Effective F stat					28.501

Note: The firm-level real capital is from the Economy Census 1995 and the ASIF 1999. The change in relative labor supply (ΔR_{ct}) is the skill mix among workers age 20–40 in Census 2005 & 2010, instrumented by ΔS_{ct} between 2004 and 2008 in columns 3–4. Firm skill intensity is the standardized college-educated to non-college -educated workers ratio in the 2004 Economy Census at firm level. Clustered standard errors at cities are shown in brackets. F—S stands for first-stage regression. ***p<0.01, **p<0.05, *p<0.1.

insignificant correlation between a firm's capital investment between 1995 and 1999 and the future influx of college-educated workers between 2004 and 2008. Unfortunately, as the R&D expenditure data is only available in the ASIF from 2001 to 2007, we cannot run a similar regression for firms' R&D expenditure for that time period.

We then continue to examine whether firms that use more skill-biased capital increase their total employment as the share of college-educated workers in the local labor market increases. According to Table 9, the point estimates suggest that when a city increases its relative supply of college-educated workers by 100%, firms that have one standard deviation higher skill intensity hire 3% more workers compared to other firms in the same city. This result is consistent with Proposition 2 as well.

To compare our results to that of Lewis (2011), we test whether firms that use more skill-biased capital adjust their capital intensity (measured by the value of capital per worker) after a change in local labor market skill mix. Lewis (2011) finds that firms in areas with larger increases in unskilled labor invest less in capital per worker because capital and skill are always complements. In Table 10, we find that firms that use more skill-biased capital do not have different capital intensities as the coefficient of the interaction between city level skill mix shock ΔR_{ct} and firm skill intensity SI_i is not statistically different from zero.

To rule out confounding factors, we conduct robustness checks in Table 11 based on firm-level regressions (Eq. (9)). Cities with different college worker ratios might have faced different trade barriers before China joined the WTO in 2001, therefore firms' investments in capital and R&D could be responses to changes in international trade.[22] By using the GAC data from 1998 to 2008, we aggregate each city's export volume by destination country (collapsed into two groups based on whether a destination country belongs

[22] Several trade models (e.g., Acemoglu, 2003; Bloom et al., 2016) assert that exports from less developed countries to the U.S. or Europe lead to a decrease in skill premium in less-developed countries.

Table 9
The effect of a city's skill mix on firms' employment 1998–2008.

Dependent variable	Changes in logged employment				ΔR_{ct}
	OLS		2SLS		F-S
	(1)	(2)	(3)	(4)	(5)
Change in relative supply: $R_{ct} = log\frac{H_{ct}}{L_{ct}}$	0.057	0.062	−0.596***	−0.586***	
	[0.088]	[0.087]	[0.140]	[0.138]	
Firm skill intensity X Change in R_{ct}		0.032***		0.029***	
		[0.006]		[0.005]	
Changes in predicted relative supply: ΔS_{ct}					0.717***
					[0.120]
Municipality, Capital & Region FE	Yes	Yes	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes	Yes	Yes
Changes in ownership category	Yes	Yes	Yes	Yes	Yes
Observations	17,814	17,814	17,814	17,814	17,814
Montiel-Pflueger Effective F stat					35.753

Note: The firm-level employment data is from the ASIF 1998–2008. The change in relative labor supply (ΔR_{ct}) is the change in the skill mix among workers age 20–40 in Census 2000 & 2010, instrumented by ΔS_{ct} between 1998 and 2008 in columns 3–4. Firm skill intensity is the standardized ratio of college-educated to non-college-educated workers at firm level in the 2004 Economy Census. Clustered standard errors at cities are shown in brackets. F—S stands for First-Stage. ***p<0.01, **p<0.05, *p<0.1.

Table 10
The effect of a city's skill mix on firms' capital per worker 1998–2008.

Dependent variable	Changes in logged capital per worker				ΔR_{ct}
	OLS		2SLS		F-S
	(1)	(2)	(3)	(4)	(5)
Change in relative supply: $R_{ct} = log\frac{H_{ct}}{L_{ct}}$	0.365***	0.364***	1.335***	1.335***	
	[0.134]	[0.134]	[0.305]	[0.305]	
Firm skill intensity X Change in R_{ct}		−0.004		−0.001	
		[0.004]		[0.003]	
Changes in predicted relative supply: ΔS_{ct}					0.0717***
					[0.120]
Municipality, Capital & Region FE	Yes	Yes	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes	Yes	Yes
Changes in ownership category	Yes	Yes	Yes	Yes	Yes
Observations	17,814	17,814	17,814	17,814	17,814
Montiel-Pflueger Effective F stat					35.753

Note: The firm-level capital per worker is calculated from the ASIF 1998–2008. The change in relative labor supply (ΔR_{ct}) is the change in the skill mix among workers age 20–40 in Census 2000 & 2010, which is instrumented by ΔS_{ct} between 1998 and 2008 in columns 3–4. Firm skill intensity is the standardized ratio of college-educated to non-college-educated workers at irm level in the 2004 Economy Census. Clustered standard errors at cities are shown in brackets. F—S stands for first-stage regression. ***p<0.01, **p<0.05, *p<0.1.

to the WTO or not) and use this controlling variable to tease out the possible influence of change in trade access. We find that adding a city's export volume by destination country does not change the key coefficient of β_1 in all specifications, suggesting that city-specific trade barriers are unlikely to be a factor in explaining firms' responses in capital investment and R&D expenditure to changes in labor market skill mix.

Another variable we considered as a possible confounding factor was rural-to-urban migrant workers, which may influence how firms adopt technology. We compare a city's change in college enrollment share to changes in the migrant population and find that there is no correlation between the two (see Fig. A4). As cross-validation, we also run a firm-level regression by including a city's extrapolated migrant population based on the 1990, 2000, and 2010 Census.[23] Table 11 shows that the point estimates for the key coefficients are similar to those in the baseline specification after adding controls for a city's migrant population.

One additional concern is that SOEs are different from non-SOE firms. To demonstrate that our results are not driven by the SOEs alone, we run regressions on firms' capital, R&D and employment again and exclude SOE firms. As shown in Table 11, we mostly find consistent and similar estimates for β_1 across all three outcome variables.

One alternative explanation for our Proposition 1 is the Rybczynski theorem of the Heckscher-Ohlin trade model. It argues that changes in the relative supply of college graduates have no effect on city-specific relative wages as long as there are within city output adjustments which can fully absorb any endowment shock. To evaluate the possibility that a firm changes its output mix rather than

[23] As the only accurate counts of migrants are from Census years, we use a quadratic time trend to fill in the gap years.

Table 11
The effect of a city's skill-mix on firms' outcomes: robustness check.

Dependent variable	Capital		R&D		Employment		Capital per worker	
(changes in logged value)	1998–2008		2001–2007		1998–2008		1998–2008	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
	Control for city's export volume by destination countries							
Firm skill intensity X Change in R_{ct}	0.028***	0.028***	0.001***	0.001***	0.032***	0.030***	−0.004	−0.001
	[0.004]	[0.004]	[0.000]	[0.000]	[0.006]	[0.005]	[0.004]	[0.003]
Observations	17,733	17,733	17,515	17,077	17,733	17,733	17,733	17,733
	Control for city's migrant population							
Firm skill intensity X Change in R_{ct}	0.028***	0.028***	0.001***	0.001***	0.032***	0.029***	−0.004	−0.001
	[0.004]	[0.004]	[0.000]	[0.000]	[0.006]	[0.005]	[0.004]	[0.003]
Observations	17,814	17,814	17,649	17,155	17,814	17,814	17,814	17,814
	Excluding State-owned Enterprises							
Firm skill intensity X Change in R_{ct}	0.025***	0.026***	0.001***	0.001***	0.031***	0.028***	−0.006	−0.002
	[0.003]	[0.003]	[0.000]	[0.000]	[0.006]	[0.005]	[0.004]	[0.004]
Observations	15,102	15,102	14,694	14,293	15,102	15,102	15,102	15,102
	Excluding firms that changed product category							
Firm skill intensity X Change in R_{ct}	0.055***	0.057***	0.001***	0.001***	0.074***	0.064***	−0.019	−0.007
	[0.019]	[0.020]	[0.000]	[0.000]	[0.018]	[0.018]	[0.017]	[0.019]
Observations	10,322	10,322	10,247	9935	10,322	10,322	10,322	10,322
Municipality, Capital & Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Changes in city GDP (in log)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Changes in ownership category	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: ΔR_{ct} and its interaction terms are instrumented by ΔS_{ct} in columns (2, 4, 6, 8). Firm skill intensity is the standardized ratio of college-educated to non-college-educated workers at firm level in the 2004 Economy Census. Clustered standard errors at the cities level are shown in brackets. ***p<0.01, **p<0.05, *p<0.1.

production technology, we categorize a firm's main product using the 3-digit classification table published by the NBS in 2002 and track if firms in our ASIF sample have ever changed their main product over time.[24] By restricting our sample to firms with the same main product, we find larger estimates of β_1 in capital, R&D and employment.

7. Conclusion

The returns to college education have steadily increased in many countries of the world, including China, despite an increased supply of college graduates. We hypothesize that this might be due to firms' endogenous responses to labor market skill mix changes in terms of their technological use. Using Chinese data that covers the period with a massive higher education expansion program, we provide direct empirical evidence to show that how firms actively respond to changes in the local skill ratio depends on technologies they use. Specifically, firms endowed with skill-biased production technology would accumulate more capital and spend more on R&D when faced with an influx of college-educated workers. In conclusion, an education policy that alters the skill mix of a country can have profound effects in determining how firms adjust their production technology.

Appendix A. Appendix

A.1. Proof for Proposition 1

Proposition 1. When firms can adjust capital freely and the range of parameters allows an interior solution for the equilibrium, the relative wage of high-skill workers compared to that of low-skill workers, $ln\left(\frac{w_H}{w_L}\right)$, only depends on exogenous production function parameters (μ, α_1, α_2, ρ^h and ρ^l) and capital rental price (r_h, r_l).

Proof. At the equilibrium with interior solution, each firm produces according to the first order conditions (Eqs. (14)–(19)) of profit maximization problem in Eqs. (3)–(4):

$$\frac{\partial Y_1}{\partial K_h} = \alpha_1 y_1^{1-\mu} K_h^{\rho_1 - 1}(K_h^{\rho_1} + L_1^{\rho_1})^{\frac{\mu}{\rho_1}-1} = r_h \quad (14)$$

$$\frac{\partial Y_1}{\partial H_1} = (1-\alpha_1) y_1^{1-\mu} H_1^{\mu-1} = w_H \quad (15)$$

[24] We use the established natural language processing method to map each firm's self-reported main product to the classification table published by the NBS.

$$\frac{\partial Y_1}{\partial L_1} = \alpha_1 y_1^{1-\mu} L_1^{\rho_1 - 1} (K_h^{\rho_1} + L_1^{\rho_1})^{\frac{\mu}{\rho_1} - 1} = w_L \tag{16}$$

$$\frac{\partial Y_2}{\partial K_l} = \alpha_2 y_2^{1-\mu} K_l^{\rho_2 - 1} (K_l^{\rho_2} + H_2^{\rho_2})^{\frac{\mu}{\rho_2} - 1} = r_l \tag{17}$$

$$\frac{\partial Y_2}{\partial H_2} = \alpha_2 y_2^{1-\mu} H_2^{\rho_1 - 1} (K_l^{\rho_2} + H_2^{\rho_2})^{\frac{\mu}{\rho_2} - 1} = w_H \tag{18}$$

$$\frac{\partial Y_2}{\partial L_2} = (1-\alpha_2) y_2^{1-\mu} L_2^{\mu - 1} = w_L \tag{19}$$

As the production function has constant return to scale, we have the zero profit conditions in Eqs. (20)–(21).

$$Y_1 - w_L L_1 - w_H H_1 - r_h K_h = 0 \tag{20}$$

$$Y_2 - w_L L_2 - w_H H_2 - r_l K_l = 0 \tag{21}$$

We introduce the simplifications in Eqs. (22)–(27) to help solve the equilibrium wage of each type of worker.

$$(14)\ \ and\ \ (16) : L_1 = \left(\frac{w_L}{r_h}\right)^{\frac{1}{\rho_1 - 1}} K_h \tag{22}$$

$$(14), (15)\ \ and\ \ (22) : H_1 = \left[\frac{\alpha_1 w_H}{(1-\alpha_1) r_h}\right]^{\frac{1}{\mu - 1}} \left[1 + \left(\frac{w_L}{r_h}\right)^{\frac{\rho_1}{\rho_1 - 1}}\right]^{\frac{\mu - \rho_1}{(\mu - 1)\rho_1}} K_h \tag{23}$$

$$(17)\ \ and\ \ (18) : H_2 = \left(\frac{w_H}{r_l}\right)^{\frac{1}{\rho_2 - 1}} K_l \tag{24}$$

$$(17), (19)\ \ and\ \ (24) : L_2 = \left[\frac{\alpha_2 w_L}{(1-\alpha_2) r_l}\right]^{\frac{1}{\mu - 1}} \left[1 + \left(\frac{w_H}{r_l}\right)^{\frac{\rho_2}{\rho_2 - 1}}\right]^{\frac{\mu - \rho_2}{(\mu - 1)\rho_2}} K_l \tag{25}$$

$$(15)\ \ and\ \ (23) : Y_1 = \left(\frac{\alpha_1}{r_h}\right)^{\frac{1}{\mu - 1}} \left[1 + \left(\frac{w_L}{r_h}\right)^{\frac{\rho_1}{\rho_1 - 1}}\right]^{\frac{\mu - \rho_1}{(\mu - 1)\rho_1}} K_h \tag{26}$$

$$(19)\ \ and\ \ (25) : Y_2 = \left(\frac{\alpha_2}{r_l}\right)^{\frac{1}{\mu - 1}} \left[1 + \left(\frac{w_H}{r_l}\right)^{\frac{\rho_2}{\rho_2 - 1}}\right]^{\frac{\mu - \rho_2}{(\mu - 1)\rho_2}} K_l \tag{27}$$

In summary, the first-order conditions for the equilibrium can be simplified as Eqs. (28)–(29)

$$(1-\alpha_1)^{\frac{1}{1-\mu}} w_H^{\frac{\mu}{1-\mu}} + r_1^{\frac{\mu}{\mu - 1}} \alpha_1^{\frac{1}{1-\mu}} \left[1 + \left(\frac{w_L}{r_h}\right)^{\frac{\rho_1}{\rho_1 - 1}}\right]^{\frac{\mu(1-\rho_1)}{(1-\mu)\rho_1}} - 1 = 0 \tag{28}$$

$$(1-\alpha_2)^{\frac{1}{1-\mu}} w_L^{\frac{\mu}{1-\mu}} + r_2^{\frac{\mu}{\mu - 1}} \alpha_2^{\frac{1}{1-\mu}} \left[1 + \left(\frac{w_H}{r_l}\right)^{\frac{\rho_2}{\rho_2 - 1}}\right]^{\frac{\mu(1-\rho_2)}{(1-\mu)\rho_2}} - 1 = 0 \tag{29}$$

We further introduce two sufficient conditions (condition 1 and 2) so that there exists w_H and w_L satisfying both the first-order and the second-order condition for profit maximization.

Condition 1: A sufficient condition for there exist two w_H and w_L subject to Eqs. (28)–(29) for every μ is that

$$g < \frac{(1-\alpha_2)e(h-g) + g}{1 - (1-\alpha_1)(1-\alpha_2)(h-g)(f-e)} < min\left\{h, \frac{1}{1-\alpha_1}\right\} \tag{30}$$

$$e < \frac{e + g}{1 - (1-\alpha_1)(1-\alpha_2)(h-g)(f-e)}\Bigg) < min\left\{f, \frac{1}{1-\alpha_2}\right\} \tag{31}$$

where

$$e = \left(\alpha_1^{\frac{\rho_1}{\rho_1 - 1}} - r_1^{\frac{\rho_1}{\rho_1 - 1}}\right)^{\frac{\rho_1 - 1}{\rho_1}} \tag{32}$$

$$f = \left(\alpha_1^{\frac{\rho_1}{\rho_1-1}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}} \tag{33}$$

$$g = \left(\alpha_2^{\frac{1}{\rho_2-1}} - r_2^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}} \tag{34}$$

$$h = \left(\alpha_2^{\frac{\rho_2}{\rho_2-1}} - r_1^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}}. \tag{35}$$

We derive this condition by finding two fixed points in Eq. (28), $\left(0, \left(\alpha_1^{\frac{\rho_1}{\mu(\rho_1-1)}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}}\right)$ and $\left(\frac{1}{1-\alpha_1}, \left(\alpha_1^{\frac{\rho_1}{\rho_1-1}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}}\right)$, that independent of the value of μ. Similarly, Eq. (29) has two fixed points, $\left(\left(\alpha_2^{\frac{\rho_2}{\mu(\rho_2-1)}} - r_2^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_1-1}{\rho_1}}, 0\right)$ and $\left(\left(\alpha_2^{\frac{\rho_2}{\rho_2-1}} - r_1^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}}, \frac{1}{1-\alpha_2}\right)$. Given that $\mu < \rho_1$ and $\mu < \rho_2$, we have

$$0 < \left(\alpha_1^{\frac{\rho_1}{\mu(\rho_1-1)}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}} < \left(\alpha_1^{\frac{1}{\rho_1-1}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}} \tag{36}$$

$$0 < \left(\alpha_2^{\frac{\rho_2}{\mu(\rho_2-1)}} - r_2^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}} < \left(\alpha_2^{\frac{1}{\rho_2-1}} - r_2^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}} \tag{37}$$

If we link these fixed points and define them as:

$$A = \left(0, \left(\alpha_1^{\frac{\rho_1}{\rho_1-1}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}}\right) = (0, e) \tag{38}$$

$$B = \left(\frac{1}{1-\alpha_1}, \left(\alpha_1^{\frac{\rho_1}{\rho_1-1}} - r_1^{\frac{\rho_1}{\rho_1-1}}\right)^{\frac{\rho_1-1}{\rho_1}}\right) = \left(\frac{1}{1-\alpha_1}, f\right) \tag{39}$$

$$C = \left(\left(\alpha_2^{\frac{1}{\rho_2-1}} - r_2^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}}, 0\right) = (g, 0)\Big) \tag{40}$$

$$D = \left(\left(\alpha_2^{\frac{\rho_2}{\rho_2-1}} - r_1^{\frac{\rho_2}{\rho_2-1}}\right)^{\frac{\rho_2-1}{\rho_2}}, \frac{1}{1-\alpha_2}\right) = \left(h, \frac{1}{1-\alpha_2}\right) \tag{41}$$

We further derive the point of intersection E as

$$E = \left(\frac{g-(1-\alpha_2)e(h-g)}{1+(1-\alpha_1)(1-\alpha_2)(h-g)(f-e)}, \frac{e+g(f-e)(1-\alpha_1)}{1-(1-\alpha_1)(1-\alpha_2)(h-g)(f-e)}\right) \tag{42}$$

Notice that w_H is continuous concave function of the w_L in Eq. (28), but a continuous convex function of the w_L in Eq. (29). If the intersection point E is between fixed points A, and B, C and D, we can always find two sets of w_H and w_L that satisfy Eqs. (28)–(29) for every μ. The following condition guarantee the position of the intersection point E.

$$g < \frac{(1-\alpha_2)e(h-g)+g}{1-(1-\alpha_1)(1-\alpha_2)(h-g)(f-e)} < min\left\{h, \frac{1}{1-\alpha_1}\right\} \tag{43}$$

$$e < \frac{e+g}{1-(1-\alpha_1)(1-\alpha_2)(h-g)(f-e)}\Big) < min\left\{f, \frac{1}{1-\alpha_2}\right\} \tag{44}$$

Fig. A5 illustrates how we derive the sufficient condition using an example when $\alpha_1 = \alpha_2 = 0.4, r_1 = r_2 = 1$. If $\rho_1 = 0.3, \rho_2 = 0.5, \mu = 0.5$, the sufficient condition is satisfied and we can find two solutions for the Eqs. (28)–(29). However, if $\rho_1 = 0.3, \rho_2 = 0.3, \mu = 0.9$, the condition is not satisfied and there is no solution for the Eqs. (28)–(29).

Condition 2: The range of parameter generate interior solutions guarantees the existence of positive level of capital for each type of firms, K_h and K_l.

Under above conditions, we claim the relative wage of high- and low-skill workers $ln\left(\frac{w_H}{w_L}\right)$ does not change after the influx of high-skill workers in local labor markets in equilibrium. As we cannot analytically verify that at least one set of solution satisfy the second-order condition, we numerically verified that the Hessian matrix (based on second derivative of the profit function at the equilibrium) is always negative definite given above two conditions. The numerical computation codes are attached in the online appendix.

A.2. Proof for Proposition 2

Proposition 2. When more high-skill workers enter the labor market, firms that use skill-biased capital hire more workers and invest more in capital. Correspondingly, firms use labor-biased capital employ fewer workers and divest capital. However, the capital intensity measured by value of capital per worker remains constant among both types of firms.

Proof. In equilibrium with interior solution, labor markets for both types of workers clear.

$$H_1 + H_2 = H \tag{45}$$

$$L_1 + L_2 = L \tag{46}$$

By combining Eqs. (45)–(46) with Eqs. (22)–(25), we rewrite the labor market clearing conditions as

$$aK_h + bK_l = H$$

$$cK_h + dK_l = L$$

where

$$a = \left[\frac{\alpha_1 w_H}{(1-\alpha_1)r_1}\right]^{\frac{1}{\mu-1}} \left[1^{+} > 0\right.$$

$$b = \left(\frac{w_H}{r_2}\right)^{\frac{1}{\rho_2-1}} > 0$$

$$c = \left(\frac{w_L}{r_1}\right)^{\frac{1}{\rho_1-1}} > 0$$

$$d = \left[\frac{\alpha_2 w_L}{(1-\alpha_2)r_2}\right]^{\frac{1}{\mu-1}} \left[1^{+} > 0\right.$$

thus we have

$$K_h = \frac{dH - bL}{ad - bc} \tag{47}$$

$$K_l = \frac{aL - cH}{ad - bc} \tag{48}$$

$$\frac{K_h}{H_1 + L_1} = \frac{1}{a + c} \tag{49}$$

$$\frac{K_l}{H_2 + L_2} = \frac{1}{b + d} \tag{50}$$

Notice that

$$ad - bc \propto ad \cdot K_h \cdot K_l - bc \cdot K_h \cdot K_l = H_1 \cdot L_2 - H_2 \cdot L_1 > 0 \tag{51}$$

As $\frac{H_1}{L_1} > \frac{H_2}{L_2}$, if $\alpha_1 + \alpha_2 < 1$ (see discussion on Eq. (5) in the main text). Therefore, we have following comparative statics

$$\frac{\partial K_h}{\partial H} = \frac{d}{ad - bc} > 0 \tag{52}$$

$$\frac{\partial K_l}{\partial H} = -\frac{c}{ad - bc} < 0 \tag{53}$$

$$\frac{\partial (H_1 + L_1)}{\partial H} = (a + c)\frac{\partial K_h}{\partial H} > 0 \tag{54}$$

Eqs. (52)–(53) indicate that firms that use skill-biased capital further invest in capital if there is an influx of high-skill workers in the local labor market, and firms that use labor-biased capital reduce capital stock after more college-educated workers enter the related labor markets. These adjustments in K_1 and K_2 also help explain the constant college wage premium after the influx of college-educated workers observed in Fig. 2. According to Eqs. (49)–(50), the capital per worker does not change after the skill-mix shock in this model, which we can empirically test as well. In addition, Eqs. (54) shows that the total employment of firms that use skill-biased capital will increase as well.

Table A1
Summary statistics of the ASIF firm sample – cross-sectional data.

Analysis sample firms in 1998	A: presence in 2007			A: presence in 2004 and 2007		
	(1) Obs	(2) Mean	(3) S.D.	(4) Obs	(5) Mean	(6) S.D.
State-owned enterprise ratio	32,730	0.432	0.495	30,320	0.425	0.494
Real value-added (millions)	32,730	19.40	14.91	30,320	19.24	14.97
Real capital (millions)	32,730	50.23	45.56	30,320	49.31	45.30
Average number of employees	32,730	502.65	2008.86	30,320	496.35	1965.91

Note: All value variables are measured in 1998 CNY. The sample includes firms that with 2-digit industry code range from 13 to 42 according to the NBS industry classification in ASIF 1998 data. We drop firms that did not enter the ASIF in 2007 in column A and firms that did not enter in either 2007 or 2004 in column B.

Table A2
Correlation between a firm's skill intensity and other outcomes in 2004.

Dependent variable (in logged value)	Capital	Capital per worker	Employment	R&D expenditure
	(1)	(2)	(3)	(4)
Firm skill intensity in 2004	0.019**	0.057***	−0.038***	0.082***
	[0.009]	[0.013]	[0.006]	[0.014]
Ownership dummies	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Observations	197,164	197,164	197,164	197,164
R-squared	0.082	0.052	0.088	0.055

Note: Firm skill intensity is the standardized ratio of college- to non-college-educated workers at firm level in the 2004 Economy Census. Clustered standard at cities are shown in brackets. ***p<0.01, **p<0.05, *p<0.1.

Table A3
Industry average skill-intensity (*SI*) in the 2004 Economy Census of China.

Industry name	2-digit CIC	Raw $\frac{H}{L}$	Standardized $\frac{H}{Ls}$	S.D. $\frac{H}{Ls}$
Agriculture products	13	0.26	−0.012	0.661
Foods	14	0.329	0.028	0.888
Beverages	15	0.34	0.035	0.870
Tobacco	16	0.324	0.025	0.217
Textiles	17	0.081	−0.117	0.211
Apparel, footwear, and caps	18	0.097	−0.107	0.332
Leather, fur, feather products	19	0.081	−0.117	0.344
Wood, bamboo, grass products	20	0.103	−0.104	0.243
Furniture	21	0.149	−0.077	0.339
Paper and paper products	22	0.134	−0.086	0.366
Printing and recorded media	23	0.227	−0.031	0.366
Articles for education and sport	24	0.108	−0.101	0.253
Petroleum, coking, nuclear fuel	25	0.349	0.04	0.616
Chemical raw materials and products	26	0.373	0.054	1.052
Medicine manufacturing	27	0.79	0.298	1.436
Chemical fibers	28	0.148	−0.077	0.417
Rubber	29	0.213	−0.039	1.987
Plastics	30	0.16	−0.071	0.341
Non-metallic mineral products	31	0.148	−0.078	0.335
Processing of ferrous metals	32	0.144	−0.08	0.644
Processing of non-ferrous metals	33	0.195	−0.05	0.411
Metal products	34	0.198	−0.048	0.604
General purpose machinery	35	0.284	0.002	0.792
special purpose machinery	36	0.54	0.152	1.324
Transport equipment	37	0.299	0.011	1.013
electrical machinery and equipment	39	0.375	0.055	1.193
Information technology equipment	40	1.077	0.466	3.130
Measuring instruments and machinery	41	1.214	0.546	2.392
Manufacture of artwork and others	42	0.126	−0.091	0.316
Recycling and disposal of waste	43	0.252	−0.017	0.527

Note: CIC stands for China Industry Classification system, which identifies 95 different industry categories.
We use the version of "GB/T 4754" that is updated in 2002.

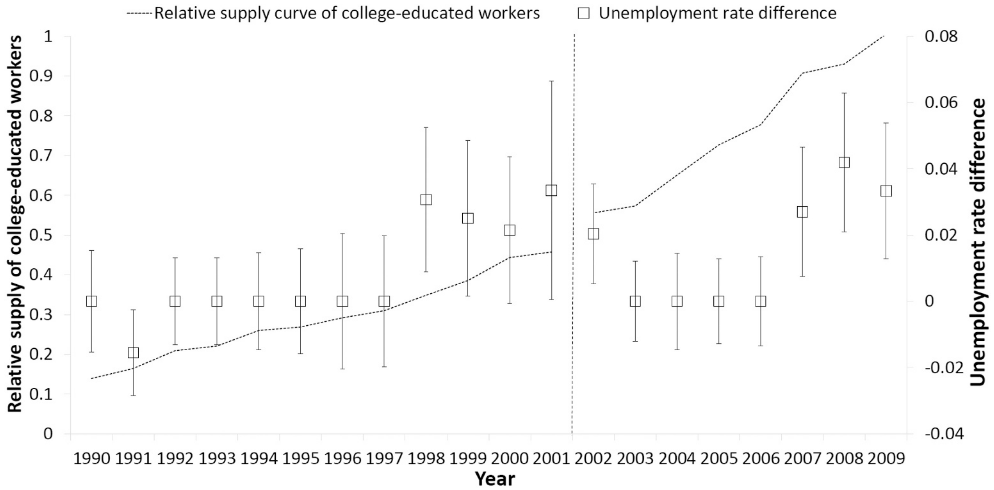

Fig. A1. The trend of unemployment rate premium of workers age between 20 and 40 after the increase in relative labor supply. We calculate the college wage premium using individuals' wage and education achievement data in the UHS.

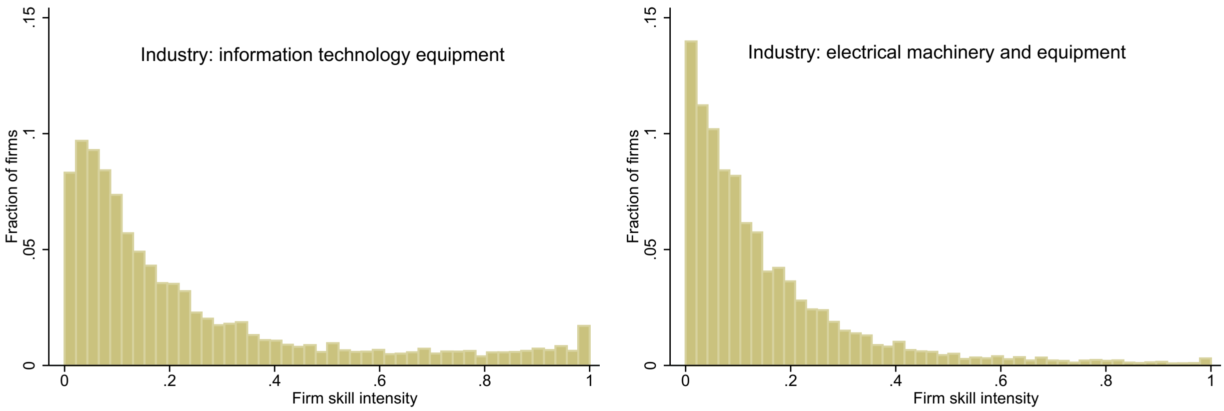

Fig. A2. Examples of skill-intensity heterogeneity within an industry. The data is from 2004 Economy Census.

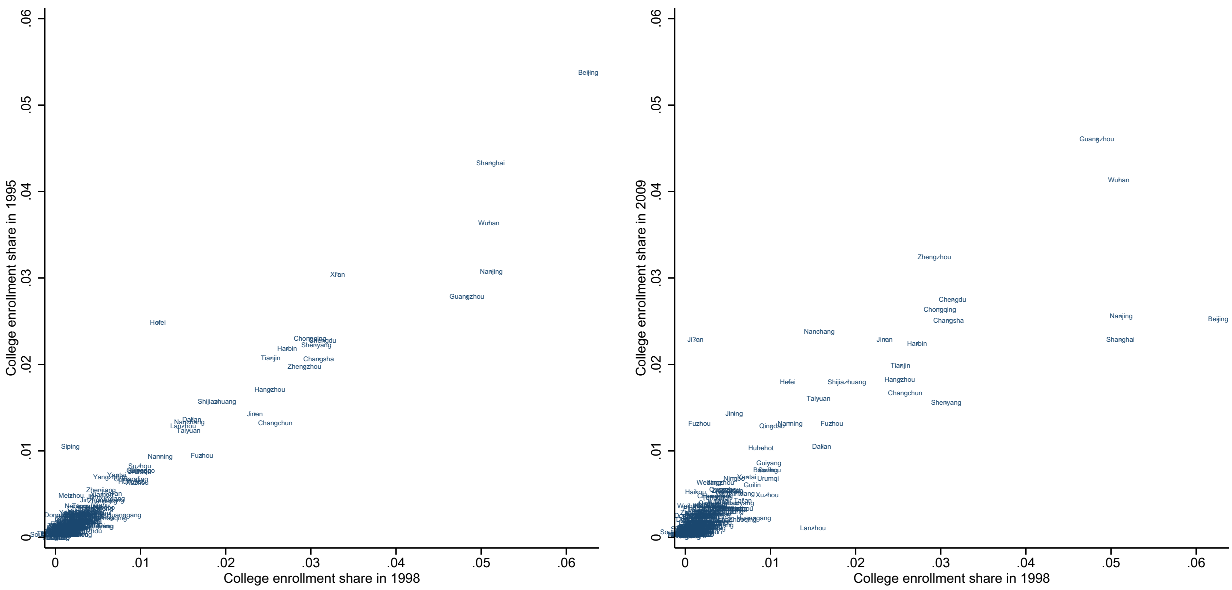

Fig. A3. There is a strong correlation in a city's college enrollment share over time (1995 is the earliest year with city-level enrollment). The data is from the Statistical Yearbook of Regional Economy.

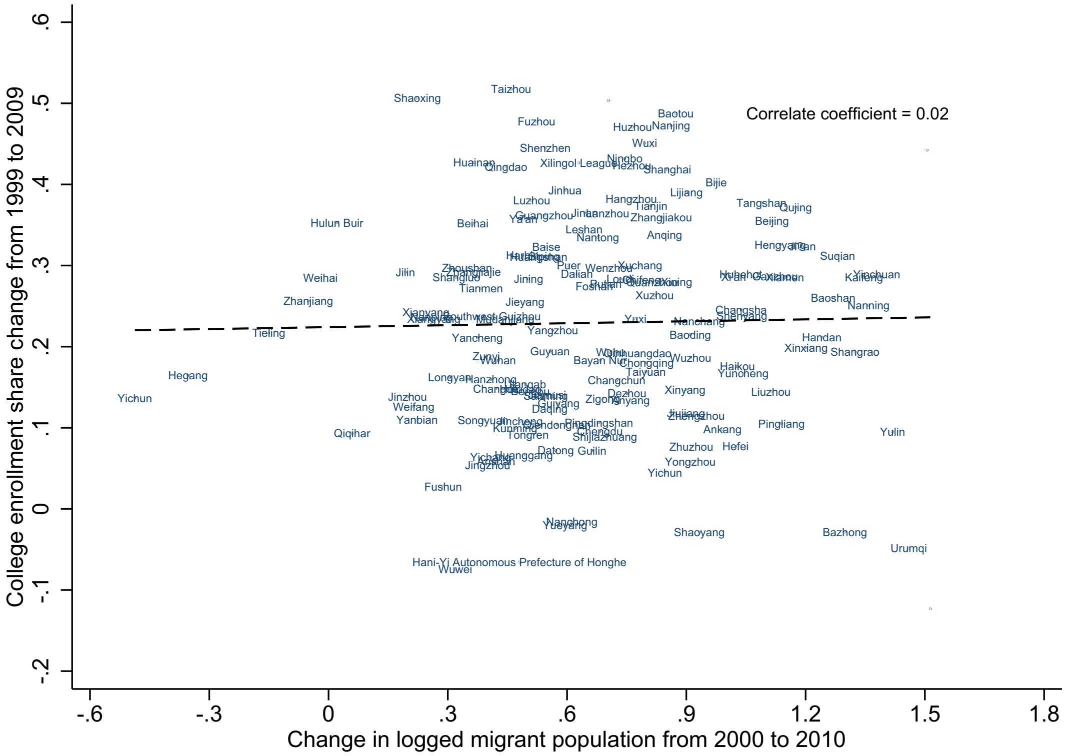

Fig. A4. Correlation between the change in a city's college enrollment share and the proportional change in the city's migrant population. The data is from Census.

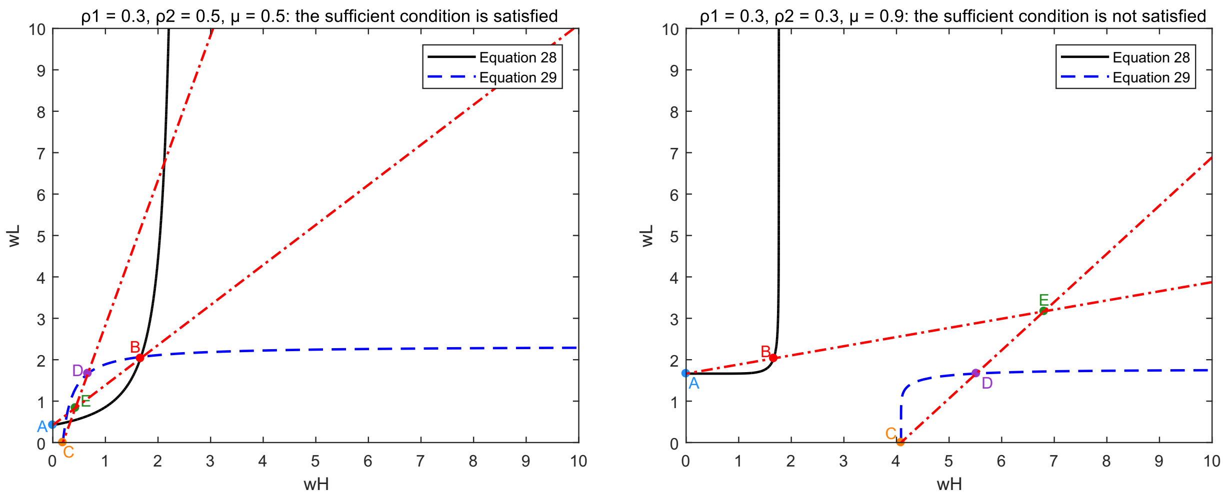

Fig. A5. Illustrating the sufficient condition for model solutions.

References

Acemoglu, D. (1998). Why do new technologies complement skills? Directed technical change and wage inequality. *The Quarterly Journal of Economics, 113*(4), 1055–1089.

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature, 40*(1), 7–72.

Acemoglu, D. (2003). Patterns of skill premia. *Review of Economic Studies, 70*(2), 199–230.

Acemoglu, D. (2007). Equilibrium bias of technology. *Econometrica, 75*(5), 1371–1409.

Angrist, J. D. (1995). The economic returns to schooling in the West Bank and Gaza Strip. *American Economic Review*, 1065–1087.
Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *The Review of Economics and Statistics, 90*(2), 300–323.
Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics, 118*(4), 1279–1333.
Barro, R., Caselli, F., & Lee, J.-W. (2013). Symposium on human capital and economic development: An introduction. *Journal of Development Economics, 104*, 181–183.
Barro, R., & Lee, J.-W. (1994). Sources of economic growth. *Carnegie-Rochester Conference Series on Public Policy, 40*(1), 1–46.
Beaudry, P., Doms, M., & Lewis, E. (2010). Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas. *Journal of Political Economy, 118*(5), 988–1036.
Bekman, E., Bound, J., & Machin, S. (1998). Implications of skill-biased technological change: International evidence. *The Quarterly Journal of Economics, 113*(4), 1245–1279.
Benhabib, J., & Spiegel, M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics, 34*(2), 143–173.
Bils, M., & Klenow, P. (2000). Does schooling cause growth? *American Economic Review, 90*(2), 1160–1183.
Bloom, N., Draca, M., & Van Reenen, J. (2016). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *Review of Economic Studies, 83*(1), 87–117.
Blundell, R., Green, D., & Jin, W. (2018). *The UK education expansion and technological change*.
Brandt, L., Van Biesebroeck, J., Wang, L., & Zhang, Y. (2017). WTO accession and performance of Chinese manufacturing firms. *American Economic Review, 107*(9), 2784–2820.
Brandt, L., Van Biesebroeck, J., & Zhang, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics, 97*(2), 339–351.
Card, D. (1999). The causal effect of education on earnings. In *, Vol. 3. Handbook of labor economics* (pp. 1801–1863).
Card, D., & Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics, 116*(2), 705–746.
Carneiro, P. M., Liu, K., & Salvanes, K. G. (2018). *The supply of skill and endogenous technical change: Evidence from a college expansion reform*.
Che, Y., & Zhang, L. (2018). Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s. *The Economic Journal, 128*(614), 2282–2320.
Chevalier, A., Harmon, C., Walker, I., & Zhu, Y. (2004). Does education raise productivity, or just reflect it? *The Economic Journal, 114*(499), 499–517.
Chow, G. C. (1993). Capital formation and economic growth in China. *The Quarterly Journal of Economics, 108*(3), 809–842.
Ciccone, A., & Peri, G. (2006). Identifying human-capital externalities: Theory with applications. *Review of Economic Studies, 73*(2), 381–412.
Duffy, J., & Papageorgiou, C. (2000). A cross-country empirical investigation of the aggregate production function specification. *Journal of Economic Growth, 5*(1), 87–120.
Dustmann, C., & Glitz, A. (2015). How do industries and firms respond to changes in local labor supply? *Journal of Labor Economics, 33*(3), 711–750.
Falvey, R., Greenaway, D., & Silva, J. (2010). Trade liberalisation and human capital adjustment. *Journal of International Economics, 81*(2), 230–239.
Feng, S., Yingyao, H., & Moffitt, R. (2017). Long run trends in unemployment and labor force participation in urban China. *Journal of Comparative Economics, 45*(2), 304–324.
Fortin, N. M. (2006). Higher-education policies and the college wage premium: Cross-state evidence from the 1990s. *American Economic Review, 96*(4), 959–987.
de la Fuente, A., & Doménech, R. (2006). Human capital in growth regressions: How much differences data quality make? *Journal of the European Economic Association, 4*(1), 1–36.
Gennaioli, N., Porta, R. L., Lopez-de-Silanes, F., & Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics, 128*(1), 105–164.
Goldin, C., & Katz, L. F. (1998). The origins of technology-skill complementarity. *The Quarterly Journal of Economics, 113*(3), 693–732.
Goldin, C., & Sokoloff, K. (1984). The relative productivity hypothesis of industrialization: The American case, 1820 to 1850. *The Quarterly Journal of Economics, 99*(3), 461–487.
Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review, 110*(8), 2586–2624.
Gregorio, J., & Lee, J.-W. (2002). Education and income inequality: New evidence from cross-country data. *Review of Income and Wealth, 48*(3), 395–416.
Iranzo, S., & Peri, G. (2009). Schooling externalities, technology, and productivity: Theory and evidence from US states. *The Review of Economics and Statistics, 91*(2), 420–431.
Juhn, C., Kim, D. I., & Vella, F. (2005). The expansion of college education in the United States: Is there evidence of declining cohort quality? *Economic Inquiry, 43*(2), 303–315.
Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics, 107*(1), 35–78.
Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984–1989. *The Quarterly Journal of Economics, 108*(1), 33–60.
Krugman, P. R. (2000). Technology, trade and factor prices. *Journal of International Economics, 50*(1), 51–71.
Krusell, P., Ohanian, L. E., Ros-Rull, J.-V., & Violante, G. L. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica, 68*(5), 1029–1053.
Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. *The Quarterly Journal of Economics, 126*(2), 1029–1069.
Lewis, E. (2013). Immigration and production technology. *Annual Review of Economics, 5*(1), 165–191.
Li, H., Ma, Y., Meng, L., Qiao, X., & Shi, X. (2017). Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. *China Economic Review, 46*, 10–26.
Li, S., Whalley, J., & Xing, C. (2014). China's higher education expansion and unemployment of college graduates. *China Economic Review, 30*, 567–582.
Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3–42.
Machin, S., & Van Reenen, J. (1998). Technology and changes in skill structure: Evidence from seven OECD countries. *The Quarterly Journal of Economics, 113*(4), 1215–1244.
Meng, X., Shen, K., & Xue, S. (2013). Economic reform, education expansion, and earnings inequality for urban males in China, 1988–2009. *Journal of Comparative Economics, 41*(1), 227–244.
Moretti, E. (2004a). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics, 121*(1), 175–212.
Moretti, E. (2004b). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review, 94*(3), 656–690.
Sato, K. (1967). A two-level constant-elasticity-of-substitution production function. *The Review of Economic Studies, 34*(2), 201–218.
Song, Z., Storesletten, K., & Zilibotti, F. (2011). Growing like China. *American Economic Review, 101*(1), 196–233.
Stokey, N. L. (1996). Free trade, factor returns, and factor accumulation. *Journal of Economic Growth, 1*(4), 421–447.
Walker, I., & Zhu, Y. (2008). The college wage premium and the expansion of higher education in the UK. *Scandinavian Journal of Economics, 110*(4), 695–709.
Xu, C. (2011). The fundamental institutions of China's reforms and development. *Journal of Economic Literature, 49*(4), 1076–1151.