Distributional Consequences of the Higher Education Expansion in China and Evolution of the College Wage Premium: 1988-2011*

Yun Feng†

November 13, 2018

Abstract

In 1999, China launched one of the most massive higher education expansions in the world. In this paper, I first investigate the effects of this reform on the evolution of the college wage premium. I then show that the reform interacts with the demographics of workers and creates winners and losers. To do so, I construct and structurally estimate a dynamic labor market general equilibrium model. I innovate in modeling and estimation by incorporating the institutional features of the higher education system of China. Through counterfactual experiments, I find that in the presence of post-reform technological progress, the reform first increases and then decreases the college wage premium. In its absence, however, the reform decreases the college wage premium from the start. I also find that in the latter case, workers induced to go to college by the reform (compliers) gain the most on average, whereas those who go to college with or without it (always-takers) lose the most, because the large increase in the supply of high-skill labor depresses skill prices. Policy experiments are conducted to show, if China were to continue with the expansion, how long it would take for it to reach the average share of high-skill workers in developed countries.

*I'm deeply indebted to my advisors, Moshe Buchinsky, Adriana Lleras-Muney, and Ed Kung for their continuing guidance, encouragement, and support. I would like to thank Maurizio Mazzocco, Till von Wachter, Kathleen McGarry, Dora Costa, and Michela Giorcelli for helpful comments. I also thank Rustin Partow, Nick Carollo, Pecheu Vladimir, Alexandre Fon, and participants at the UCLA Applied Micro Pro-seminar and the All-California Labor Economics Conference for useful feedback. This paper employs parallel computation on the Hoffman2 Shared Cluster provided by the Institute for Digital Research and Education (IDRE) at UCLA.

†Department of Economics, University of California, Los Angeles. Email: yun.feng@ucla.edu

1 Introduction

In 1999, China launched one of the most massive higher education expansion reforms in the history of the world. In anticipation of increasing demand for high-skill workers, the Ministry of Education started to sharply increase the maximum number of students who could be admitted to college in 1999.[1] Compared to 1998, the number of newly admitted college students in 1999 increased from about 1 million to 1.5 million. It then kept increasing in all subsequent years, to 7.2 million in 2014 (Figure 1).[2] As a result, the supply of high-skill workers who have at least some college education has been increasing rapidly since 2001 (Figure 1).[3] In the presence of such a massive expansion of higher education, one might expect a sizable decline in the college wage premium.[4] To the contrary, the college wage premium had been increasing since 1999 and only started to modestly decrease in 2009 (Figure 2). This suggests that strong labor demand-side forces, such as skill-biased technological progress, is shifting relative demand for high-skill versus low-skill workers.

This paper investigates the effects of the higher education expansion reform in China on the college wage premium by disentangling such effects from those resulting from other forces, such as technological progress, changes in cohort size, and changes in capital rental prices. In addition, I examine how the reform interacts with the demographics of workers and affects them differentially. Lastly, I study how long it will take for China to catch up with developed countries in terms of the share of high-skill workers.

To do so, I construct a dynamic general equilibrium model along the lines of Lee and Wolpin (2006). In the model, workers produce skill by acquiring schooling and accumulating work experience, and are paid based on how much skill they supply and their skill types. A worker becomes high-skill if he has at least some college education. There exists a representative firm that decides how much of each type of skill and capital to use. I innovate by modeling the college admissions process. One's admission probability depends on the national admission rate and on proxies for one's ability, such as one's parents' education percentiles within their cohorts.

The model has two novel features. First, low- and high-skill workers have different state spaces.

[1] This maximum is set by the Ministry of Education as an admission quota each year. See Section 2 for details on the background.

[2] Both 2- and 4-year college students are included. Statistics on admissions are calculated based on data collected from China Education Statistical Yearbooks.

[3] Throughout the paper, I define high-skill workers as those who have at least some college education, and low-skill workers as those who have finished high school or less. Most 2-year college students who were admitted in 1999 graduated in 2001.

[4] The college wage premium is defined as the average wage gap between people who have at least some college education (high-skill workers) and those who do not (low-skill workers).

The college admission rate enters the state space of low-skill workers, as they never attended college and the evolution of admission rates affects their future option values. Second, the reform not only directly impacts the admissions process, but also affects workers' decisions and labor market outcomes through their expectations on the evolution of college admission rates. The model is able to separate the effects of the reform on the college wage premium from those coming from technological progress (both skill-biased and skill-neutral), changes in cohort size, and capital rental prices. I structurally estimate the model using survey and aggregate data on China and use the estimated model to conduct counterfactual and policy experiments. I obtain the following main findings.

First, in the presence of post-reform technological progress (both skill-neutral and skill-biased), the reform increases the college wage premium before 2008 with a diminishing effect, from about 0.16 log points in 1999 to virtually zero in 2007. It then decreases the wage premium with an amplifying effect from 0.06 to about 0.14 log points. On average, the reform increases the college wage premium by 21%, with a yearly increment of 0.07 log points. This may seem counterintuitive, as we might expect the reform to decrease the wage premium holding technological progress the same. However, in this paper, wage is defined as a product of the skill price and the amount of skill one supplies.[5] The effect on the college wage premium is therefore determined by two components: changes in the skill-price gap and changes in the average skill-stock gap.[6] Although the reform narrows the skill-price gap between high- and low-skill workers, it widens the average skill-stock gap, since it allows more low-skill workers, who on average have more skill to go to college and become high-skill. The sign of the effect depends on which effect dominates.

In contrast, fixing technological progress at the pre-reform level, I find that the effect of the reform on the college wage premium is negative and increases over time. On average, the reform decreases the college wage premium by 7% per year. Without post-reform technological progress, the reform becomes the main factor that affects skill prices.[7] Both the skill-price gap and the average skill-stock gap between high- and low-skill workers are narrowed as a result. The average skill-stock gap narrows because in the absence of post-reform skill-biased technological progress, the marginal product gap between high- and low-skill workers converges instead of diverging. Hence, fewer low-skill workers will invest in becoming high-skill. The large number of incoming young and relatively inexperienced high-skill workers decreases the average skill stock of high-skill workers and narrows the skill-stock gap. A decomposition of the reform's net effect on the college

[5] Heckman et al. (1998) show the importance of distinguishing wage from skill price in human capital models, as they sometimes move in different directions.

[6] Throughout the paper, the gap is calculated as the high-skill workers' average minus that of the low-skill workers'.

[7] This is not the only factor, since changes in cohort size and capital rental prices are still present.

wage premium shows that most of the effect (99%) comes from this narrowing of the average skill-stock gap. This is because the pre-reform high-skill stock in China was extremely low and the expansion was massive.

In addition, I find that the higher education expansion reform has differential impacts on workers. Cohorts directly affected by the reform gain the most: about 87% compared to the counterfactual without the reform. For cohorts that are not directly exposed to the reform, the effect is positive on average, but is very close to zero for most of them. Cohorts that graduated from high school just a few years before the reform lose modestly, by 0.15%; This is primarily driven by the more able ones who are still young enough to be privately efficient to abandon their jobs and go to college. I also examine the effect on the discounted lifetime wage by treatment group. The group induced to go to college by the reform (compliers) on average gain the most by 97,164 yuan whereas those who go to college even in the absence of the expansion (always-takers) lose by 2.6%. They lose because they suffer from the decrease of high-skill price due to the large increase in the supply of high-skill labor.

Finally, I conduct two policy experiments and show that if China were to continue with the trends in technological progress and admissions process in 2011, by 2052 China would catch up with developed countries in terms of the share of high-skill workers in the working-age population (age 16 - 60). This can be achieved by 2031 if China follows the common practice of college admissions in developed countries by abandoning the explicit constraint on admission quotas beginning in 2012. A back-of-the-envelope calculation shows that the latter is worthwhile if the average cost of adding a seat is not greater than 176,000 yuan.

The paper is structured as follows. Section 2 describes the background and Section 3 sets up the model. Section 4 discusses the intuition on identification and Section 5 describes the data. Section 6 outlines the moments of choice and the estimation strategy. Section 7 discusses the estimation results. Sections 8 and 9 describe the counterfactual experiments that deliver the results on the reform's effects on the college wage premium and its distributional effects, respectively. Section 10 discusses the results of policy experiments. Section 11 concludes.

Related Literature

This paper contributes to three strands of literature. The first examines the evolution of the college wage premium (e.g., Katz and Murphy (1992); Card and Lemieux (2001); Goldin and Katz (2008); Lee and Wolpin (2010); Blundell et al. (2018)), and establishes the importance of various contributing factors, such as skill-biased technological progress, trade, changes in the female labor force participation rate, etc. These papers examine the college wage premium in the setting of developed countries. In contrast, this paper studies a new setting and contributes to the literature by focusing on a developing country in which nationwide government interventions and structural transformation are present. In addition to the factors investigated in prior literature, this new setting allows me to explore how a large-scale education reform affects the college wage premium.

In terms of methodology, this paper contributes to the literature on dynamic general equilibrium models (e.g., Heckman et al. (1998); Lee (2005); Lee and Wolpin (2006); Dix-Carneiro (2014); Llull (2017)). The model employed in this paper is most related to that of Lee and Wolpin (2006), who construct and structurally estimate a labor market general equilibrium model that explains the growth of the U.S. service sector between 1968 and 2000. The model in this paper has two key features that differ from theirs. First, to focus on the role of the higher education reform, this paper incorporates the unique features of China's college admissions process in the model. As a result, the state spaces of low- and high-skill workers are different by an aggregate state variable: the college admission rate. It only enters the state space of low-skill workers because the evolution of admission rates affects their future option values directly. In addition, forward-looking low-skill workers must form expectations on the evolution of the college admission rates, which adds extra complexity to the solution and estimation of the model.

Lastly, this paper adds to the literature on the labor market effects of China's higher education expansion (e.g., Meng et al. (2013); Li et al. (2014); Li et al. (2016)). These papers study different labor market outcomes and how they are related to higher education expansion. Although the primary focus of Meng et al. (2013) is how the increase in the price of unobserved skills could explain the increase in the variance of the earnings of urban male workers, they attribute, to some extent, the slowing down of the rewards to both observed and unobserved skills in the early 2000s to higher education expansion. Li et al. (2014) focus on the effects of the higher education expansion on unemployment, and find that the reform increased the unemployment rate of young college graduates.

Li et al. (2016) construct a general equilibrium model and use simulation to show that a positive demand shock for skill can explain the observation that the college wage premium declines for young workers, but increases for more experienced workers following the reform. Although, I do not concentrate on explaining the different wage premium trends for workers with different experience, I model a worker's human capital as affected by both education and experience, a key feature emphasized by Li et al. (2016). In contrast to their model, in which the changing admissions policy is reflected by a net-return-to-education parameter, I explicitly model the college admissions process and how it changes over time. In addition, I account for both observed and unobserved heterogeneity of workers. The amount of skill a given type of worker supplies is endogenously determined. In addition to studying the evolution of the college wage premium, I also explore how the reform interacts with the demographics of workers and differentially impacts their discounted lifetime wages—a margin that is not well understood in the context of China's higher education expansion. To the best of my knowledge, this is the first paper that structurally estimates a labor market general equilibrium model to study the impact of the higher education expansion reform. The approach I take in this paper allows me to disentangle the effects of the higher education expansion reform on the evolution of the college wage premium from factors such as technological progress (both skill-biased and neutral), changes in cohort size, and capital rental prices. It also provides a way to conduct counterfactual and policy experiments that may inform policy making.

2 Background

The Higher Education System of China

The higher education system in China is very different from that of the US. Every year, those who want to go to college register for the annual College Entrance Exam (CEE) before receiving information on admissions that year. The Ministry of Education releases information on how many registered for the CEE and what the planned national admission quota is months after registration is complete. The quota determines the maximum number that can be admitted to college that year. How large the quota is depends on the number of available seats in universities and the government's anticipation of future demand for high skill. After taking the CEE, students submit a preference list of schools and compete with other students based on their CEE scores.

Tuition

Before 1985, no tuition was charged for college students, and some students from low-income households could receive a monthly subsidy of about 20 yuan. Starting in 1985, China gradually carried out tuition reforms and started to charge tuition. Between 1989 and 1992, tuition was around 200 yuan. As China transitioned to a market-driven economy, tuition started to rapidly increase and reached around 3,000 yuan in 1997 and 4,000 yuan in 2,000. In recent years, tuition is around 5,000 yuan for most public universities. [8]

[8] Tuition varies across majors, universities and provinces. For majors such as Medicine, the tuition is slightly higher (about 5,000-6,000 yuan). Majors in the arts usually charge a much higher tuition; in some cases, about 10,000 yuan.

3 Model

In this section, I construct a dynamic labor market equilibrium model where workers make endogenous choices of education and work and the representative firm decides how much skill and capital to employ. The model is along the lines of Lee and Wolpin (2006). I extend their model by incorporating two key institutional features of the higher education system of China. First, the Ministry of Education usually releases information on the maximum number of college students they plan to admit months after the registration of the College Entrance Exam (CEE). Thus, the people who want to go to college have to make decisions before obtaining accurate information on admissions. Since this capacity constraint is always binding due to the high expected return to college education, determining this admission quota is equivalent to setting a national admission rate. The admission rate enters my model as a key aggregate state variable. It reflects the intensity of the higher education expansion and affects human capital investment decisions through workers' expectations on it. In addition, I model the admissions process, which depends on the national admission rate, proxies of workers' ability, and shocks. In reality, the admissions process starts one or two months after the CEE and determines the actual number of people who are admitted to college that year.

Compared to the literature on dynamic general equilibrium models (Heckman et al. (1998), Lee and Wolpin (2006) and Dix-Carneiro (2014) among others), the model I build has two novel features. First, low- and high-skill workers have different state space. The admission rate only enters the state space of low-skill workers as they never attended college and the evolution of the admission rate affects their future option value. Low-skill workers not only have expectations on the evolution of skill prices but also on college admission rates, through which the policy change affects workers' decisions and labor market outcomes. [9] Second, the model features an admissions process that depends on proxies of ability and captures changes in the composition of workers. The unobserved shock in the admissions process reflects factors that are not in the workers' control during the exam and the admissions process.

There are five forces in the model that affect the college wage premium: changes in the admissions process (both in the admission rate and the parameters), changes in the skill-neutral technological progress, changes in the skill-biased technological progress, changes in capital rental prices and changes in the sizes of entering cohorts. These forces interact and together explain the changes in the college wage premium we observe in the data. I describe the model as follows.

[9] By low-skill workers, I mean they are low-skill at least for some periods when they are in the model economy, not necessarily always like so. Note that the reform also affects existing high-skill workers indirectly through the general equilibrium.

3.1 Workers

Workers make education and labor supply decisions. They choose from three alternatives d_a^k ($k = 1, 2, 3$) at age a. d_a^k is an indicator that equals to 1 if alternative k is chosen at age a. l denotes the unobserved ability types ($l = 1, 2, 3$). I assume there are three ability types. ω_{ik}^l denotes the skill endowment of person i who is of type l if he chooses alternative k. ω is present when the economy starts in this model and is fixed over time. The differences of skill endowment across different alternatives can be interpreted as i's comparative advantage. In terms of preferences, I follow Lee and Wolpin (2006), and assume that the utility of consumption is additively separable from that associated with labor supply decisions. This simplifies the worker's problem tremendously as we can solely focus on labor supply decisions. Ω_{at} is the information set one has given age a and time t.

The flow utility at age a and calendar time t are:

$$U_a = U(d_a^k|\Omega_{at}) + U(c_{at})\ k = 1, 2, 3.\ l = 1, 2, 3$$

Alternative 1 ($d_a^1 = 1$): acquire education

Depending on different stages of education, a worker has to pay tuition t_1 for college and $t_1 + t_2$ for graduate school. Going to school before college is assumed to be free.[10] To capture the fact that most people go to school consecutively, I denote by κ_1 the cost of returning to school if someone didn't go to school in the previous year. η_t^1 is a transitory preference shock one has at t.

$$U(d_{at}^1|\Omega_{at}) = \omega_1^l - t_1 1\left(Educ_a \geq 12\right) - t_2 1\left(Educ_a \geq 16\right) + \kappa_1 \left(1 - d_{a-1,t-1}^1\right) d_{a,t}^1 + \eta_t^1 \qquad (1)$$

Alternative 2 ($d_a^2 = 1$): work

$$U(d_{at}^2|\Omega_{at}) = \gamma + w_{lat}^j\ j = H,\ L\ l = 1, 2, 3 \qquad (2)$$

γ is the "non-pecuniary" benefit associated with working. A worker can be either a high-skill worker ($j = H$) if he has acquired at least some college education ($Educ > 12$) or a low-skill worker ($j = L$) if he never went to college before ($Educ \leq 12$). He rents the amount of skill s he has in the labor market if he chooses to work and gets paid by r^j for every unit of it. $w^j$is the wage bill a worker of skill type j is paid. The wage bill varies by the worker's fundamental type l, age and time.

[10] This is to reduce the number of parameters to be estimated but can be relaxed.

$$w^j_{lat} = r^j_t s^j_{la} \tag{3}$$

A worker produces the amount of human capital or skill using an exponential function. The inputs are his endowment of type l, ω^l_2, his years of education, and work experience. A transitory productivity shock η_2 also affects his skill production.[11]

$$s^j_{la} = exp\left(\omega^l_2 + \beta^j_1 Educ_a + \beta^j_2 Exper_a + \eta_{2a}\right) \tag{4}$$

Alternative 3 ($d^3_a = 1$): stay at home

Similar to the education alternative above, to capture the persistent behavior of staying at home, a fixed benefit κ_3 is introduced if a worker chooses to stay at home for two consecutive years.

$$U(d^3_{at}|\Omega_{at}) = \omega^l_3 + \kappa_3 d^3_{a-1,t-1} d^3_{a,t} + \eta^3_t \tag{5}$$

The worker's problem can be formulated as the following dynamic programming problem.

$$V_a\left(\Omega_{iat}\right) = max_{k\in\{1,2,3\}}\left\{V^k_a\left(\Omega_{iat}\right)\right\} \tag{6}$$

$$V^k_a\left(\Omega_{iat}\right) = \begin{cases} U(d^k_{iat}|\Omega_{iat}) + \delta E_{\epsilon,\eta,admission}\left[V_{a+1}\left(\Omega_{ia+1,t+1}|\Omega_{iat},\ d^k_{iat}=1\right)\right] & a < 60 \\ U(d^k_{iat}|\Omega_{iat}) & a = 60 \end{cases} \tag{7}$$

3.2 Production

I model the firm's side parsimoniously using a CES aggregate production function.[12] The representative firm employs three factors: low skill L, high skill H and capital K. $\alpha's$ are share parameters. I assume them to be time varying to capture within-sector reallocation of factors, which can be interpreted as skill-biased technical change. σ and ν govern elasticities of substitution. Figure 4 shows that 1995 and 2002 are two important turning points for the evolution of wages of different skill groups. Anecdotally, these two years also correspond to the points when China went through major events. In 1995, China started reforms that aimed at downsizing the state-owned enterprises

[11] One concern is that as more and more people are admitted to college, the quality of higher education might decrease and therefore impedes skill production in college. The teacher-student ratio in two- and four-year college in part reflects the quality of higher education. Since this ratio decreased modestly from 0.12 in 1998 to 0.1 in 1999 and was steady around 0.06 starting from 2002, the role of education quality is likely to be limited.

[12] I also tried a different functional form: $Y_t = z_t\left[\alpha_{1t}L^\sigma_t + \alpha_{2t}H^\sigma_t + (1-\alpha_{1t}-\alpha_{2t})K^\sigma_t\right]^{\frac{1}{\sigma}}$. I find that the production function I'm using now explains the data better.

(SOE). In 2002, this process was complete. Moreover, China joined World Trade Organization (WTO).The two turning points I choose capture the timing of structural transformation in China in a parsimonious way. The log difference of the aggregate shocks z between two consecutive years is assumed to follow a AR(1) process.

$$Y_t = z_t \left\{\alpha_{1t} L_t^{\sigma} + (1-\alpha_{1t})[\alpha_{2t} H_t^{\nu} + (1-\alpha_{2t}) K_t^{\nu}]^{\frac{\sigma}{\nu}}\right\}^{\frac{1}{\sigma}} \tag{8}$$

$$\alpha_{kt} = \begin{cases} \alpha_{k0} & if\ t < 1995 \\ \alpha_{k0} + \alpha_{k1}\,(t-1994) & if\ 1995 \le t \le 2001 \\ \alpha_{k0} + 7\alpha_{k1} + \alpha_{k2}\,(t-2001) & if\ 2002 \le t \le 2011\ \ (k=1,2) \end{cases} \tag{9}$$

$$logz_{t+1} - logz_t = \phi_0 + \phi_1\,(logz_t - logz_{t-1}) + \zeta_{t+1} \tag{10}$$

3.3 Labor Market Equilibrium

The competitive equilibrium of this economy is such that all workers maximize their life-time utility, the representative firm maximizes its profit and all markets clear. For computational tractability, I assume the prices in the product market and capital market are determined by the rest of the world.[13]

First-order conditions of the firm's problem

$$\frac{\partial Y_t}{\partial H_t} = r_t^H\ \frac{\partial Y_t}{\partial L_t} = r_t^L \tag{11}$$

$$\frac{\partial Y_t}{\partial K_t} = r_t^K \tag{12}$$

Labor market clears

$$S_t^j = \sum_{a=16}^{60} \sum_{i=1}^{N_{at}} s_{iat}^j 1(skill\ type_i = j)\ j = H,\ L \tag{13}$$

$$S_t^H = H_t\ S_t^L = L_t \tag{14}$$

[13] This assumption is commonly maintained in the literature on dynamic labor market general equilibrium models (e.g. Lee and Wolpin (2006) for the U.S.; Dix-Carneiro (2014) for Brazil) to avoid having to solve for capital rental prices endogenously (if the capital market is closed), given that solving for skill prices in the labor market is already very computationally intensive.

3.4 Admissions Process and State Transitions

3.4.1 Admissions Process

Data on test scores that span across many years are not available, I therefore let one's admission probability depend on the proxies of one's ability. I specify the admissions process as follows.

$$adm_{it} = adm_t + p_{1t}(PctlOfEducAt16_i - E(PctlOfParentsEduc_i))$$

$$+p_{2t}(PctlOfEducAt16_i - E(PctlOfParentsEduc_i)) + \epsilon_{it} \quad (15)$$

One's admission probability adm_{it} depends on the nation level admission rate that year, adm_t, and how high his percentile is with respect to the mean percentile in the distribution of age-16 years of education and parent's education. ϵ_{it} is assumed to be mean zero and uncorrelated with the percentiles. It reflects factors that one cannot control during the College Entrance Exam (CEE) and forecasting errors when submitting one's preference list of schools. p_1 and p_2 change across years since the importance of the percentiles may vary over time. This formulation means the unconditional ex-ante probability of admission is equal to the admission rate, adm_t. Moreover, conditional on the same percentiles, everyone has the same probability of getting admitted up to ϵ_{it}.[14] The deviation from mean percentile variables determine one's (ability) type and are used as proxies for innate ability that determines test scores. There is no available information on how diligent one is in preparing for the exam, but $PctlOfParentsEduc_i$ partly reflects this to the extent that children from a better educated family tends to work harder.

Besides information on percentiles, the national admission rate adm_t plays an important role. It is defined as the ratio of the number of newly admitted students to both 2-year and 4-year college in year t over the number of people who register for the CEE in the same year. As described in the background section, the ministry of education determines a quota after the registration of CEE is complete, which is the maximum number of students who can be admitted.

Although the most salient feature of higher education expansion in this paper is the sharp increase in admission quota starting from 1999, using this quota directly in the model requires a realistic way of ranking all individuals who want to go to college. It will become inevitably arbitrary without detailed information on how observed demographics translate into test scores. Instead, I exploit one important feature of the data and get around this problem. The admissions capacity constraint is always binding in the data because of the high expected return to college education.

[14] $E(adm_{it}|PctlOfEducAt16_i, \quad PctlOfParentsEduc_i) = adm_t + p_{1t}(PctlOfEducAt16_i - E(PctlOfEducAt16_i)) + p_{2t}(PctlOfParentsEduc_i - E(PctlOfParentsEduc_i))$

Hence, as long as the model generates moments that match the number of people who register for the CEE and the national admission rate, the number of admitted students generated by the model must be equal to the observed quota as well.[15] Thus, I use admission rate in the model instead of the quota.

Another reason why admission rate is favorable over quota as modeling choice is that it is consistent with the way people aggregate information on admission in reality. The number of admitted students and that of registration are in terms of millions. It's very difficult for people to keep track of these large numbers and form expectations on future quantities. Instead, at least anecdotally, most people keep track of admission rates and use it as an important source of information when evaluating the chance of getting into college.

In reality, the admissions process occurs after people take the CEE and submit their preference lists of schools. Although one may have some idea on how one is ranked in the ability distribution, this information is far from perfect. Also, as is pointed out above, factors that are not in one's control during the exam and the admissions process make it even more difficult to predict ex-ante what one's admission probability is. To be consistent with these features, in the model, I assume that agents are agnostic as to what p_1 and p_2 are, they decide whether to continue with education once their years of education reach 12.[16] If they do, they will enter the admissions process and continue with education for at least one year if get admitted. If they are not admitted, they will choose their second best option (either working or staying home). The admissions process can be thought of mimicking the reality in the following sense. Based on their skill endowment and innate ability, agents take the exam and submit their preference list of schools. Ex-ante, there's no guarantee one will get admitted. But good students tend to have higher probability of admission.

3.4.2 State Transitions

The transition rule of the state variables are as follows:

Work experience evolves deterministically and increases by one if a worker chooses to work in the previous year. Education evolves stochastically when it's equal to 12 and otherwise the same as work experience.

[15] When doing counterfactual experiments, quota constraint may not be binding and I have to use quota instead of admission rate.

[16] 12 years of education correspond to the completion of high school. According to the Education Statistics Yearbooks of China, among the exam takers, more than 80% are the ones who are about to graduate from high school or professional schools. More than 90% of these students are high school students. I allow for returning exam takers as long as they are not older than 25 and have at least 11 years of education because regulation-wise, it's impossible for adults to go back to high school and acquire formal education. I choose age 25 because the number of people who are older than 25 and take the CEE account for only 0.05% of the total.

$$Exper_{a+1} = Exper_a + d_a^2 \tag{16}$$

$$Educ_{a+1,t+1} = \begin{cases} Educ_{a,t} + d_{a,t}^1 & if\ Educ_{a,t} \neq 12 \\ Educ_{a,t} + d_{a,t}^1 1(admission_t = 1) & if\ Educ_{a,t} = 12 \end{cases} \tag{17}$$

Following the literature (e.g. Lee and Wolpin 2006, Dix-Carneiro 2014), to reduce the dimensionality of the state space, I adopt an adaptive forecasting rule of skill prices to approximate rational expectations.[17] Note that this does not assume away general equilibrium effects. Because the $\rho' s$ are not themselves structural parameters but are functions of other structural parameters of the model. They are estimated such that they are consistent with the equilibrium prices. The$\rho' s$ can be interpreted as the agents' beliefs on the evolution of skill prices. In equilibrium, their beliefs are correct.

$$logr_{t+1}^j - logr_t^j = \rho_0^j + \rho_1^j \left(logr_t^j - logr_{t-1}^j \right) + \xi_t \tag{18}$$

For the admission rate, I adopt a different adaptive forecasting rule: $q_t = q_{t-1}$. This is only relevant to people who have never attended college. They use the admission rate of last year to forecast future admission rates. I choose this forecasting rule for three reasons. First, to calculate the admission rate, one needs information on admission quota and the number of people who register for the College Entrance Exam (CEE). Such information is only available months after the registration of CEE. Therefore, one has to decide whether or not one wants to go to college before knowing the admission rate of that year. Second, to the extent that the number of available seats and the expected return to college education are similar for adjacent years, the admission quota and the number of registered exam takers should also be similar. It is thus reasonable for agents to use q_{t-1} to predict q_t. Third, by modeling the expectations on admission rate in this way, it simplifies a lot the computational burden, which is already quite heavy.

The resulted state space is as follows:

[17] As is noted in Lee and Wolpin (2006), in principle, to solve for a rational expectations equilibrium, the agents should use the whole history of aggregate variables such as skill prices and admission rates (including all the current and past values) as well as the cross-sectional distributions of all individual state variables such as education, work experience etc. However, it's computationally infeasible to take into account all these things when solving for the equilibrium. Instead, to the extent that the one-period lag of the growth rate of the skill prices contains information that's most relevant to determining the current growth rate, we adopt this simplified forecasting rule. Dix-Carneiro (2014) shows that in his setting, the quantitative results are not sensitive to this particular specification by comparing the result to a perfect foresight equilibrium. Such robustness checks require major modifications of the code and the re-estimation of the model, which takes a long time. I will keep working on it and include the results of the sensitivity checks in the future.

$$\Omega_{at} = \{a,\ l,\ j,\ Educ_{at},\ Exper_{at},\ d_{a-1,t-1,}\ q_t,\ r_t,\ r_{t-1},\ \eta_t,\ \xi_t\}$$

The distributional assumptions follow standard practice in the literature. $\eta' s$ are assumed to be correlated across three alternatives and follow a multivariate normal distribution with mean zero. $\eta' s$ are iid across individuals and time. ξ follows a mean zero normal distribution and is iid across time.[18]

4 Identification

This section provides intuition on identification. Intuitively, the data allow us to tell parameters apart as long as they don't move exactly the same set of moments. That is, the model is identified as long as for any parameter θ, there doesn't exist another parameter θ' that moves exactly the same set of moments as θ.

On the production side, the share parameters $\alpha' s$ are directly related to the factor income shares of low-skill workers and the composite of high-skill workers and capital. The main source of variation that is used to identify the $\alpha' s$ come from the time series variations in those inputs. Given the $\alpha' s$, σ is identified off the relative changes over time in low skill supply and the composite of capital and high skill supply, whereas ν is identified off the relative changes over time in high skill supply and capital.

On the worker's side, tuition costs t_1 and t_2 are identified by comparing the proportion of individuals who have a college degree or an advanced degree but of different types. Skill production slope parameters β_1^j and β_2^j are identified off the variations of education and work experience by comparing individuals who are of the same skill types but earn different wages. Given the same skill prices, variation in wages comes from differences in skills. The returning cost to school, κ_1, corresponds to the proportion of individuals who return to school from work or home. The persistence parameter of the alternative of staying home, κ_3, corresponds to the proportion of individuals who stay at home for consecutive periods. Lastly, the skill endowment parameters $\omega' s$ reflect worker's comparative advantage associated with each of the choice alternative. Conditional on the same observables, workers with high skill endowment for alternative k are more likely to choose it. Therefore, the proportion of such workers provides information on the magnitude of ω.

In addition, since the admission quota, cohort size, and capital rental rates are taken as exogenous in the model, variations in these variables are also exogenous from the point of view of

[18] For details on how to solve the model, please see the appendix.

this model. In general, these exogenous variations along with normalizations and functional form assumptions help achieve the identification of this model.

5 Data

This paper combines multiple sources of data together and this section describes each of the dataset.

5.1 Survey Data

Repeated Cross-Section Data

I use three repeated cross-section datasets from the Urban Household Survey (1988-1997), the China Household Income Project (1988-2007) and the China General Social Survey (2003-2012).[19] These data are nationally representative and are the best publicly available data that span across the years I study. [20]

All three datasets provide detailed information on demographics, education, career choice and earnings and the survey design is comparable to the Current Population Surveys (CPS) in the U.S. The Urban Household Survey (UHS) data were collected by the China Statistics Bureau to keep track of the evolution of the socioeconomic conditions of urban Chinese households. The China Household Income Project (CHIP) tracks the dynamics of income and expenditure in China. It was carried out by Chinese and international researchers with the assistance of the China Statistics Bureau.[21] The China General Social Survey (CGSS) is conducted starting from 2003 by researchers from Hong Kong University of Science and Technology and Renmin university. The goal was to construct a nationally representative dataset that can be widely used in empirical social science research. [22]

Panel Data

The panel data used in this paper are from the China Health and Nutrition Survey (CHNS), which were collected since 1989 by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease

[19] Urban Household Survey covers all years from 1988 to 1997. For the periods I study, the China Household Income Project covers 1988, 1995, 1999, 2002, 2007 and the China General Social Survey covers 2003, 2005, 2006, 2008, 2010, 2011, 2012. The CGSS updated its database and included two more years of data (2013, 2015). I plan to extend the period I study to include these years soon.

[20] The Urban Household Survey data are made available by the Chinese University of Hong Kong.

[21] For details on the sampling and survey design, see Eichen and Zhang (1993), Li et al. (2008) and Luo et al. (2013).

[22] For details on sampling and survey design, see Bian and Li (2014).

Control and Prevention (CCDC).[23] Although the survey was designed to track the health and nutritional status of the population, it provides information on basic demographics, work experience and earnings and can be used to construct the work history of the workers.

Main Features of the Data

I restrict the sample to be consistent with the model, where workers enter the economy at age 16 and exit at 60. After dropping observations with missing values for key variables such as age and education, I obtain a pooled repeated cross-section sample containing 317,886 observations. Wages are deflated to their 1988 levels using the GDP deflators of China downloaded from the World Development Indicators database of the World Bank. [24]Among the data I use, the sampling procedure of the CGSS changed in 2006 and 2008 and information on individual weight is not available for most of the years. To make the CGSS data comparable to the other data, I use a weighting procedure that's similar to what is used in the CPS such that the key variables (e.g. the share of people by education category, sex, stratum) are matched closely to the census (see the Appendix for details). [25]

Table 1 shows the evolution of employment shares by education category in urban China. Both the shares of people who are college educated and above and those who have some college education increase. Following the higher education expansion in 1999, we see the sharpest increase in the share of people who have at least some college education.

5.2 Aggregate Data

The aggregate data used in this paper are primarily from the online database of the China Statistics Bureau. The value-added series are readily obtainable from the database. To calculate the capital rental rates, I first use data on labor income and GDP from the database to calculate the share of labor income. Then I calculate the ratio of capital income over capital stock as the capital rental rate.[26]The cohort size data are collected from the China Population Census 1990, 2000, 2010 and

[23]For the period I study, the CHNS covers 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011.

[24]Since GDP is usually deflated using the GDP deflator and wages must be deflated in the same way as GDP, I choose the GDP deflator. Note that the CPI are slightly higher than the GDP deflators for 1989-2004 and are slightly lower for the rest of the years.

[25]CGSS 2005, 2006, 2008 are matched to the 2005 population survey. CGSS 2010-2012 are matched to the 2010 census.

[26]$capital\ rental\ rate = \frac{(1-labor\ share)*real\ GDP}{real\ capital\ stock}$. I follow the argument made in Bai, Qian (2010) and use the GDP calculated by income approach. Such GDP data reflect the factor income distribution of domestic production and is therefore most relevant for the calculation of capital rental rates in this paper. The GDP series are deflated to their 1988 levels using the GDP deflators of China downloaded from the World Development Indicators database of the World Bank.

the 2005 China Population Survey. In addition, the data on the registration of the College Entrance Exam and the number of admitted students are collected from China Education Statistics Yearbooks.

6 Estimation

6.1 Choice of Moments

The model is estimated using the Simulated Method of Moments (SMM). I use three sets of moments from the survey data: moments on career choice, wage distribution and education distribution. I also complement these moments with aggregate data such as the time series of real value-added, capital rental prices, cohort size and the number of newly admitted students. The data moments I choose are as follows.

- Choice distribution (cross-sectional data from the UHS, CGSS and CHIPS)

  1. The proportion of individuals choosing each of the three alternatives by year (1988–2011), age (16–60)
  2. The proportion of individuals choosing each of the three alternatives by year (1988–2011), and schooling level (three categories: <=12, 13 –15, 16+).
  3. The proportion of individuals choosing each of the three alternatives by year (1988–2011)

- Wage distribution (cross-sectional data from the UHS, CGSS and CHIPS)

  1. The mean log real wage by year
  2. The mean log real wage by highest grade completed (<=12, 13–15, 16+), year
  3. The mean log real wage by year, age
  4. The variance in the log real wage by education and year.
  5. The variance in the log real wage and year.

- Wage distribution (panel data from CHNS)

  1. The mean log real wage by work experience (0, 1, 2, 3, 4+ years)

- Schooling distribution (cross-sectional data from the UHS, CGSS and CHIPS)

1. The distribution of highest grade completed (<=12, 13–15, >=16) by year (1988–2011), age (16–60)

- The number of newly admitted students to college (both 2-year and 4-year) by year

6.2 Conditional Type Probabilities

Since the (ability) types are unobserved, we have to estimate the probability of being each type for each person. Individuals are observed for the first time when they enter the economy at age 16. To the extent that variations in observed measures of skills at age 16 reveal information on one's innate ability, we should let one's type probability depend on such observables (or proxies of them). Variables such as years of education, work experience and gender of a worker when he is first observed are typically included in the literature (e.g. Lee and Wolpin 2006, Dix-Carneiro 2014).[27] In China, workers can start to work legally at age 16. Hence, at the beginning of age 16, everyone has zero work experience. For the purpose of this paper, I need observables other than years of education at 16 to better capture the heterogeneity in innate ability.

Since information on common measures of ability such as test scores is not available in any existing surveys for the time span I study, I exploit the intergenerational features of my data sets and use information on parent's years of education. In particular, to control for the fact that the average years of education increase over time, I calculate the percentile of one's parent's education according to where they are in the distribution of years of education of their cohort.[28] Although the compulsory schooling law was introduced in China in the mid-1980s, there was still variation in age-16 years of education. To control for the increase of average age-16 years of education, I also calculate one's percentile in the distribution of age-16 years of education among adjacent cohorts.[29]

I let the type probabilities depend on one's percentile of age-16 education and the average of percentiles of one's parents when one was 16.[30] Holding other demographics constant, being higher in the distribution of years of education at 16 than others suggest a high skill endowment for education. In addition, parents with higher percentiles are more likely to have children with high

[27] In Lee and Wolpin (2006), workers are first observed when they are 16. Therefore, work experience is zero for everyone.

[28] Adjacent cohorts are faced with very similar economic conditions and education resources and therefore comparable. I calculate parent's education percentiles in the pooled three adjacent cohorts. For instance, if one's parent is 40 when he/she is 16, I calculate the parent's education percentile among the age-39, -40 and -41 cohorts.

[29] Although of the same age, depending on what month in which one was born, they may start school in different years. To accommodate more observations and account for whom one's competing with when applying for college, I pool age-15, -16 and -17 cohorts together to calculate the percentile.

[30] It rarely happens that parents gain more years of education after their children reach 16. Even if they do, the effects on Children are very limited according to the literature on early childhood development.

skill endowment. Prior literature (e.g. Keane and Wolpin 1997) shows that the multinomial choice structure is flexible enough for the estimation of conditional type probabilities. I assume there are three types ($l = 1,\ 2,\ 3$), and specify the conditional type probabilities as follows:

$$Prob(type = l|x_{1i}^0,\ x_{2i}^0) = \frac{exp(\pi_{0l} + \pi_{1l}x_{1i}^0 + \pi_{2l}x_{2i}^0)}{1 + \sum_{j=2}^{3} exp(\pi_{0j} + \pi_{1j}x_{1i}^0 + \pi_{2j}x_{2i}^0)},\ l = 1,\ 2,\ 3 \tag{19}$$

where $x_{1i}^0 = PctlOfEducAt16_i$, $x_{2i}^0 = PctlOfParentsEduc_i$

The $\pi' s$ will be estimated together with the rest of the parameters of the model. π_{01}, π_{11} and π_{21} have to be normalized to zero to guarantee the identification of the$\pi' s$.

6.3 Estimation of the Admissions Process

There are two ways to estimate the admissions process. In both ways, the admission rate at year t, adm_t, is calculated as $\frac{No.\ of\ newly\ admitted\ students}{No.\ of\ people\ registered\ for\ the\ CEE}$.[31] One way is to estimate the admissions process with the model, which means there are three more parameters to estimate for each year (p_{1t}, p_{2t} and the variance of ϵ_t). Given that there are 24 years, the parameters to estimate will increase by 72.[32] The other way is to estimate this process outside the model for each year using available information on the two percentile variables and whether or not one is admitted. To reduce the burden of parameter search, I adopt the latter way. Ideally, we need information in the survey on whether and when one registered for the CEE. However, this information is rarely available. Given that almost all high school and professional school graduates take CEE and they account for more than 80% of the registered exam takers, I assume everyone takes CEE when they are just about to graduate from high school or professional school and look for their admission status based on their obtained highest degrees.

7 Estimation Results

7.1 Parameter Estimates

The estimation results are shown in Table 2 through Table 5.[33] Table 2 shows the estimates of preference parameters by the three available options each period. The cost of attending college is

[31] We should use the number of newly admitted students instead of the quota because the admissions process happens after all students take the CEE and submit their preference list of schools. It determines the actual number of admissions, which might be slightly different from the quota.

[32] Although I could restrict the parameters to be the same across years, it doesn't seem to be consistent with data.

[33] To estimate the model, I employ parallel computation using 16 cores on a cluster to improve on speed.

estimated to be about 3200 yuan in 1988 terms, which is about 13000 yuan today. This is comparable to what most public universities charge plus living cost.[34] The extra cost of going to graduate school is about 2000 yuan today. Although some graduate students get paid by working with their professors, others have to pay for the tuition themselves. This estimate reflects an average of the extra cost of attending graduate school, which includes tuition, living cost and the psychic cost due to the pressure of graduating.

Table 2 also shows the skill production parameters. β_1^j and β_2^j measure how efficient type j is in producing skill using years of education and work experience. High-skill workers is better than low-skill ones in producing skill using years of education but not work experience. This does not mean that high-skill workers are not good at on-the-job learning. It's simply because they on average accumulate less years of work experience since they stay in school longer. β_2 reflects the average efficiency of on-the-job learning and it's necessary for low-skill workers to be good at this to generate enough wage growth since it primarily comes from the accumulation of work experience.

Table 3 reports the estimates for the production function. σ and ν govern the elasticities of substitution of the production factors. Since σ is estimated to be bigger than ν, the elasticity of substitution between low-skill labor and the composite is higher than that between high-skill labor and capital. Thus, the estimates suggest high-skill labor is indeed more complementary to capital than low-skill labor. Table 4 and Table 5 show the estimates for the conditional type probability and the admissions process parameters. Type 3 has the highest endowment in education and work whereas Type 2 has the lowest. In general, higher percentiles of years of education at 16 and parents education lead to higher probability of admission.

7.2 Goodness-of-Fit

Figure 5 shows how the choice distribution the model generates fits the data. The model does a good job for years before 1997, but not as good for years between 1999 and 2005. This is partly due to the quality of the data I use for those years. For years before 1998, I use the UHS data collected by the China Bureau of Statistics. The sampling and survey design is maintained in a consistent way across years. For years of 1998 onwards, the data I use are from the CHIP and the CGSS, which were carried out by different research teams and resulted in more noise in the pooled sample. Figure 6 compares how the actual and the simulated average wages of the high- and low-skill workers evolve. Although there is some discrepancy, the trend generated by the model

[34] Most public universities charge around 5000 yuan for tuition and 1000 yuan for dormitory. The average living cost is around 6000 yuan a year.

tracks that of the data. Figure 7 shows the goodness-of-fit for college wage premium measured as the average log wage gap. The log transformation makes the wage premium between 1993 and 1999 look more volatile. For years between 1999 and 2011, where the model will be used for counterfactual exercise, the fit is reasonable.

7.3 Robustness

To address the concern that the parameter estimates may correspond to a local minimum of the objective function, I use the following way to find the initial guess of the parameters. Before implementing the search algorithms such as the Nelder–Mead simplex method, I first specify a grid as fine as possible (100 to 200 points) for each parameter over the possible range. By iterating on the grid of each parameter, I record the best parameter value in each round and use it as the preferred guess for the next round until reaching the end. Although this way doesn't guarantee finding the minimum of the objective, it is a better way to come up with a good initial guess that can be used as a input of the search algorithm. Another upside of this is that the initial guess already reflects information on the shape of the objective function and it partially avoids finding different estimates due to the arbitrary choice of the initial guess. In addition, I use a hybrid search function that combines global minimum solvers such as Simulated Annealing and local minimum solvers such as Pattern Search as a cross-check for the results I got using the standard simplex methods.

8 Effects of the Higher Education Expansion on the College Wage Premium

8.1 Effects on the College Wage Premium

To get the effects of the higher education expansion on the college wage premium, I use the estimated model as a laboratory and conduct several counterfactual experiments. There are five forces in the model that affect the college wage premium: changes in the admissions process (both in quota and/or p_{1t}, p_{2t} and σ_ϵ), changes in aggregate productivity shock z_t, changes in share parameters α_{1t} and α_{2t}, changes in capital rental prices, and changes in the sizes of entering cohorts. These forces interact and together explain the changes in college wage premium we observe in the data. In the following sections, I focus on the effects of the changes in the admissions process while controlling for the other four forces.

Effect of the Reform in the Presence of Post-Reform Technological Progress

To understand what the trend in the college wage premium would have been in the absence of the reform, I simulate a counterfactual in which all the parameters remain the same as their estimated values except for the admissions process. The admission quota and admissions process parameters p_{1t} and p_{2t} are fixed at their pre-reform levels in 1998. To the extent that p_1 and p_2 only measure how the proxies of one's innate ability translate into admission probability, they should only reflect how good one is at taking the College Entrance Exam (CEE) and submitting their preference list of schools. Therefore, as long as the admissions policy remains fixed, these parameters should be relatively stable across time and invariant to changes in technology in the counterfactuals.

One may assume that an alternative way to fix the admissions process is to keep the admission rate constant at its 1998 level, since it also reflects the intensity of the reform. However, the policy instrument the Ministry of Education uses is the quota. In the counterfactual, the number of registered exam-takers won't be the same as in the data, due to changes in skill prices. Keeping the admission rate fixed may result in a number of admitted students that is potentially much different from the planned admission quota. Comparison of Figures 1 and 3 suggests that the admission quota is more likely to be a primary policy instrument, since it is much less volatile. Figure 1 shows that before the expansion, the number of newly admitted students each year increases slowly with very small fluctuations. Figure 3 shows more fluctuations in the admission rate, which is primarily driven by changes in the number of registered exam takers.

In Figure 8, the two lines on top show the simulation results. The solid line is the baseline college wage premium generated by the estimated model, in which all parameters are the same as their estimated values and the admissions process evolves as it does in reality. The dashed line with circles (Counterfactual 1) shows the counterfactual trend without the reform while keeping other factors the same. Since the only difference between the two cases is the reform, their comparison gives us the effect of the reform (see the solid line in Figure 9). The reform increases the college wage premium before 2008 with a diminishing effect from about 0.16 log points in 1999 to virtually zero in 2007. It then decreases the wage premium with an amplifying effect from 0.06 to about 0.14 log points. On average, the reform increases the college wage premium by 21%, with a yearly increment of 0.07 log points.

This may seem counterintuitive as we might expect the reform to decrease the wage premium holding technological progress fixed. However, as will become clear in Section 8.2, the effect on the college wage premium is determined by two components: changes in the skill-price gap and changes in the skill-stock gap. Although the reform narrows the skill-price gap between high- and low-skill workers, it widens the skill-stock gap since it allows more low-skill workers who on average have more skill to go to college and become high-skill. The sign of the effect therefore depends on which component dominates.

Effect of the Reform in the Absence of Post-Reform Technological Progress

In this paper, the technological progress is modeled in a "reduced-form" way in the sense that I do not specify how it depends on other primitives of the model. To the extent that the amount of high-skill workers in the economy may affect the upgrading and diffusion of technology, post-reform technological progress may be different in the absence of the reform. This subsection therefore investigates the effect of the reform on the college wage premium holding technological progress (both skill-biased and skill-neutral) fixed at the pre-reform level.

I simulate two counterfactuals, one with the reform carried out as it is in reality (Counterfactual 2) and another without the reform (Counterfactual 3). In the case without the reform, the admission quota and the admissions process parameters, p_1, p_2 are fixed at the 1998 level for 1999 onwards. In the other case, the admission quota and the $p's$ evolve as they are in reality. In addition, aggregate productivity shock z_t and share parameters α_{1t} and α_{2t} are fixed at their 1998 levels in both cases. The implication is that any technological progress, reforms or other forms of structural transformation that affect productivity or the allocation of factors are shut down.

I allow the entering cohort sizes (age-16) to be exactly the same as in the data, because fertility decisions are already made before 1999.[35] Lastly, capital rental rates remain the same as in the data in both cases, because they are determined by the rest of the world. Comparing these two cases, any differences in the college wage premium must be due to changes in the admissions process.

The bottom two lines in Figure 8 show the evolution of the college wage premium in both cases. The dashed line in Figure 9 shows the net effect of the reform.[36] Without post-reform technological progress, it becomes clear that the reform depresses the college wage premium immediately from the start and the effect increases over time to about 0.08 log points. On average, the reform decreases the college wage premium by 7% per year. One reason why the reform depresses the college wage premium from the start is that the reform allows some workers who would have worked in the labor market as low-skill workers to go to college, which decreases the supply of low-skill workers. The skill price of low-skill workers therefore increases and the skill-price gap shrinks. In Section 8.2, I discuss what drives the increase of the effect over time.

[35]The youngest cohort in the simulated economy is 16 in 2011, which means they were born in 1995.

[36]By net effect, I mean the effect of the reform on the college wage premium in the absence of post-reform technological progress.

Interaction between the Reform and Post-Reform Technological Progress

In the presence of post-reform technological progress, the effect of the reform is a combination of its net effect and its interaction effect with post-reform technological progress. This can be seen in Figure 8. Comparing Counterfactuals 1 and 2 to Counterfactual 3, respectively, the gap shows the net effect of post-reform technological progress and that of the reform on the college wage premium. The gap between the baseline and Counterfactual 3 gives the total effects of both factors, which include their net effects and their interaction effect. The interaction effect on the college wage premium arises from two facts. First, in general equilibrium, the skill prices are affected by both the reform and post-reform technological progress, which in turn affect workers' human capital investment decisions and how much skill they accumulate. Second, the college wage premium is determined by changes in both skill prices and skill stock.

8.2 Understanding the Effects of the Reform

To understand what drives changes in the college wage premium, in this section I decompose such changes into several margins of adjustment. To the extent that wage is a product of skill price and the amount of skill one supplies in the labor market, changes in the wage premium can be decomposed as coming from two main margins: changes in the skill-price gap and changes in the average skill-stock gap. Such decomposition follows naturally from the model. The wage premium in this paper is defined as the average log wage gap of high- and low-skill workers. A change in this gap is a sum of two pieces: the change in the skill-price gap and that in the average skill-stock gap.

$$\triangle Wage\, premium = \triangle \left(\frac{\sum_{i=1}^{N_H} log(r_H S_i)}{N_H} - \frac{\sum_{j=1}^{N_L} log(r_L S_j)}{N_L} \right)$$

$$= \underbrace{\triangle \left(logr_H - logr_L \right)}_{Changes\ in\ the\ Skill-Price\ Gap} + \underbrace{\triangle \left(\frac{1}{N_H} \sum_{i=1}^{N_H} logS_i - \frac{1}{N_L} \sum_{j=1}^{N_L} logS_j \right)}_{Changes\ in\ the\ Average\ Skill-Stock\ Gap} \qquad (20)$$

The change in the average skill-stock gap can be further written as a sum of four components: changes in the average work endowment gap, changes in the average education gap, changes in average work experience gap and the difference between the averages of shocks. The last term should be close to zero if the number of workers is large.

$$\triangle \left(\tfrac{1}{N_H} \sum_{i=1}^{N_H} logS_i - \tfrac{1}{N_L} \sum_{j=1}^{N_L} logS_j \right) = \underbrace{\triangle \left(\frac{1}{N_H} \sum_{i=1}^{N_H} \omega_2^{l_i} - \frac{1}{N_L} \sum_{j=1}^{N_L} \omega_2^{l_j} \right)}_{Changes\ in\ the\ Average\ Work\ Endowment\ Gap}$$

$$+ \underbrace{\triangle \left(\beta_1^H \overline{Educ_i} - \beta_1^L \overline{Educ_j} \right)}_{Changes\ in\ the\ Average\ Education\ Gap}$$

$$+ \underbrace{\triangle \left(\beta_2^H \overline{Exper_i} - \beta_2^L \overline{Exper_j} \right)}_{Changes\ in\ the\ Average\ Experience\ Gap} + \overline{\eta_{2i}} - \overline{\eta_{2j}} \tag{21}$$

Understanding the Effect of the Reform in the Presence of Post-Reform Technological Progress

Figure 10 shows the decomposition of the reform's effect into changes in the skill-price gap and the average skill-stock gap when there is post-reform technological progress. From the decomposition, it's clear that the effect on the college wage premium is determined by two components: changes in the skill-price gap and changes in the skill-stock gap. Holding technological progress the same in the baseline and Counterfactual 1, the reform increases the supply of high skill relative to low skill and therefore narrows the skill-price gap. In the meantime—allowing low-skill workers—who on average have more skill to go to college and become high-skill, the reform widens the skill-stock gap. The widening effect becomes smaller over time, because the reform creates a large and increasing supply of young and inexperienced high-skill workers. As they enter the labor market, the average skill stock of high-skill workers decreases, which creates a countervailing effect on the skill-stock gap.

Understanding the Effect of the Reform in the Absence of Post-Reform Technological Progress

Figures 11 and 12 show the graphical versions of the two decomposition equations, (20) and (21). Figure 11 shows that changes in the average skill-stock gap drive the net effect of the higher education expansion reform on the college wage premium. Changes in the skill-price gap also serve as a force that depresses the college wage premium, but is not the driving force. Figure 12 shows that changes in the education and work experience gaps drive the changes in the average skill-stock gap. Interestingly, changes in the work-endowment gap are positive. This is because as the reform allows more workers who are on average with lower work endowment to go to college, the average work endowment gap widens. Figure 13 shows that the widening of this gap is primarily because the average work endowment of low-skill workers decreases.

Overall, on average changes in the skill-price gap account for 0.93% of the net effect of the higher education expansion on the college wage premium, whereas changes in the average skill-stock gap account for 99.07%. In addition, changes in workers' composition account for -9.71% of the net effect. Changes in the average education gap account for 5.61%, and changes in the average work experience gap account for 103.31%.[37]

Although one may expect that changes in the skill-price gap drive changes in the college wage premium, this result shows that for a developing country like China, where the existing stock of high skill was extremely low (before 1999), massively expanding higher education narrows the skill-stock gap between low- and high-skill workers. This is primarily due to young high-skill workers with zero or little work experience entering the labor market. As the number of such young high-skill workers increases each year, the average skill-stock gap continues to narrow, which leads to the amplifying negative effect on the college wage premium. Note that this narrowing effect is also present when there is post-reform technological progress. However, it does not outweigh the widening effect, because technological progress makes the option of becoming high-skill much more attractive for low-skill workers.

It is important to distinguish the narrowing of the average skill-stock gap from the decrease in the average ability of high-skill workers (measured by work endowment ω_2). As Figure 12 shows, changes in workers' composition actually increase the college wage premium and the effect is small. Overall, changes in the skill-stock gap dominate the reform's net effect on the college wage premium for two reasons. First, the stock of high skill was extremely low in China before the expansion. Second, the reform expands higher education massively and enables a continuous and increasing large supply of young high-skill workers with zero or little work experience to enter the labor market each year.

9 Distributional Effects of the Higher Education Expansion

The rich heterogeneity built into the model allows me to study how the higher education expansion affects people with different demographics. The results of this section are based on the two counterfactuals simulated before on the net effect of the reform.[38] Different cohorts differ in the time when they enter and exit the economy. To assess the consequences of the higher education expansion on the same basis, I have to simulate every cohort until they exit the economy at age 60. Since

[37]The three don't sum up to 1 because in finite sample, the changes in the differences of the average shocks are not exactly zero.

[38]I focus on the case without the presence of post-reform technological progress because I'm primarily interested in the net effect of the reform alone.

the youngest cohort is age 16 in 2011, I simulate the model from 1999 to 2055, when the youngest cohort reaches 60. By comparing the outcomes in the two cases, I can obtain the treatment effect at the individual level.

Although aggregate productivity shock z_t and share parameters,α_{1t} and α_{2t}, are fixed at their 1998 levels, I have to specify how capital rental prices and entering cohort sizes evolve beyond 2011. I first estimate a VAR based on the data I have. I then take the estimated process as the evolution rule and simulate future quantities for years beyond 2011.[39]

$$logN_{c,t+1} - logN_{c,t} = c_0 + c_1 \left(logN_{c,t} - logN_{c,t-1}\right) + \lambda_{t+1} \tag{22}$$

$$logr_{K,t+1} - logr_{K,t} = k_0 + k_1 \left(logr_{K,t} - logr_{K,t-1}\right) + \mu_{t+1} \tag{23}$$

As for the admissions process in the case with expansion, I let the admissions quota and p_1, p_2 be fixed at 2011 level for subsequent years.[40] After simulating the economy to 2055, I calculate the discounted sum of lifetime wages in both cases for all cohorts that are in the economy from 1988 to 2011. By comparing the two cases, I get the net treatment effects of the higher education expansion reform on the discounted sum of lifetime wages for each worker.[41]

9.1 Effects for Different Cohorts

Figure 14 shows the effects on the discounted sum of lifetime wages by cohort. Each cohort is indexed by the year they enter the economy (age 16).

For cohorts 1944 through 1954, the effect is exactly zero, because all of these cohorts exit the economy before the higher education expansion reform starts in 1999.[42] Cohorts 1955 through 1989 overlap with the reform by at least one year. The effect is increasing in the number of years they overlap, and their gains are close to zero but positive. Not surprisingly, the cohorts that gain the most are those that experience the higher education expansion reform when they reach 19 (Cohort 1996 - 2011), which is the first year the majority complete high school and take the College Entrance Exam (CEE). Cohorts 1996 through 2011, on average, gain by about 87% compared to the counterfactual

[39]Cohorts that will enter the economy from 2012 and 2027 are already born in 2011 and I use actual data for these years. Death rates are very low for the newly born cohorts and I assume they won't die. The initial conditions of these new cohorts are unobserved and I assume they are the same as the age-16 cohort in 2011.

[40]Since I only look at cohorts that are already in the economy in 2011, these assumptions I make in this section are only going to affect the results of a limited number of cohorts through general equilibrium.

[41]The discount factor used is 0.95.

[42]For convenience, throughout Section 9, I refer to a cohort by the year when they enter the model economy at age 16.

without the reform.

It's worth noting that cohorts 1990 through 1995 actually lose modestly, by 0.15%. This is primarily driven by returning CEE takers. These cohorts are between 20 and 25 when the reform starts. For those who are relatively better educated and have higher skill endowment in work and education, it's still privately efficient for them to go to college. However, they have to take the CEE as returning exam takers and give up their jobs if admitted. In addition, to transition from work to school, they have to pay an extra cost besides tuition. Although they gain more years of education, they also lose several years of work experience. Overall, being someone who narrowly makes it to become high-skill in the presence of the reform does not guarantee more discounted lifetime wages, compared to the status quo of staying as low-skill in the case without the reform.[43]

9.2 Effects for Different Treatment Groups

Comparing the two counterfactuals, one can also calculate the effects of the higher education expansion reform on the discounted sum of lifetime wages for different treatment groups: those who go to college with or without the reform (always-takers), those who are induced to go to college (compliers), those who are induced to not go to college (defiers) and those who don't go to college regardless of the presence of the reform (never-takers).[44] Figure 15 shows that perhaps not surprisingly, the group that gains the most are the compliers. They gain, on average, by 97,164 yuan—more than five times what they would have earned as low-skill workers in the counterfactual without the reform. The gain is particularly large, because it is calculated off the fraction of the population that is most affected by the reform.[45] A high-skill worker not only accumulates more skill and supplies it at a higher skill price, but is also more efficient in producing skill using years of education. Becoming a high-skill worker early in life leads to the dynamic accumulation of such benefits and the gains could potentially be quite large.

As for the group that goes to college with or without the reform (always-takers), they lose

[43] Note that although they lose on average compared to the baseline, they still choose their first best in the case with the reform.

[44] In the presence of the reform, the skill-price gap between high- and low-skill workers decreases, which lowers the relative attractiveness of attending college versus not. This pattern strengthens as the reform persists, and may change the decisions of those who are at the margin of attending college in the absence of the reform. It's also important to note that although in general, people who go to college in the absence of the reform are of higher ability than those who are admitted in the presence of the reform, due to the randomness in the admissions process, this is only true on average, not necessarily at the individual level. For a discussion of checking the validity of the monotonicity assumption in the LATE literature, please see De Chaisemartin and D'Haultfœuille (2012) and Fiorini and Stevens (2016), among others.

[45] It's important to note that discounted lifetime wages are are calculated as the discounted sum at the first year when one is in the model economy. For cohorts 1988 through 2011, the sum is discounted to age 16. Moreover, the discounted sum is measured in terms of 1988 yuan.

modestly by 2.6% on average. The loss is, in general, increasing in their exposure to the reform. They lose because they suffer from the decrease in high-skill prices due to the large increase in the supply of high-skill labor. The group that doesn't go to college regardless of the presence of the reform (never-takers) gains modestly, by about 8.7%, because as the share of low-skill labor decreases, the demand for low-skill labor increases and the skill price increases.

10 Policy Experiments

I conduct two policy experiments in this section. The main rationale for the higher education expansion is that China is still far lower than the average of developed countries in terms of the share of high-skill workers in the working-age population. There's a debate on whether China should continue to expand higher education. This section aims to address this debate.

In the first policy experiment, I ask when, if China were to continue with the trends in technological progress and admissions process in 2011, it would catch up with the developed countries' average of 30% in terms of the share of high-skill workers in the working-age population. To do so, I let the admission quota, p_{1t}, p_{2t}, z_t, α_{1t} and α_{2t}, be what they should be for years 1988 to 2011. For years after 2011, I keep them fixed at the 2011 levels. I simulate entering cohort sizes and capital rental prices as in the last section.

In the second policy experiment, I ask how much sooner, if China were to follow the common practice of college admissions in developed countries by abandoning the explicit constraint on admission quota from 2012, it could reach the target of 30%?[46] In addition, what is the maximum cost of adding new seats that makes it worthwhile? To do so, I simulate a case in which everything is the same as in the first policy experiment, except that the admission quota is equal to the number of registered CEE takers after 2011. Figure 16 shows the evolution of the share of high-skill workers over time. In policy experiment 1, China will reach 30% of high-skill workers in the working-age population in 2052, whereas in policy experiment 2, China will reach this target in 2031.

Figure 17 shows the evolution of GDP in both experiments. Although technological progress is fixed at the 2011 level, the economy grows as the skill stock increases. However, growth is not going to last forever, because as the share of high-skill workers increases, the marginal product of high skill decreases, which depresses demand. In the meantime, the marginal product of low skill increases will become extremely high as the share of low-skill labor decreases. At some point, the high-skill labor employed in the economy will be low enough such that the GDP decreases.

[46] In this case, although there is no capacity constraint on admissions, not every low-skill worker goes to college because they are subject to tuition cost, the opportunity cost of staying low-skill, and the cost of returning to school (for those who already graduated from school).

Comparison of trends in the GDP in Figure 17 for years between 2011 and 2031 gives the increase in GDP due to eliminating the capacity constraint on admissions. Policy Experiment 2 shows that not everyone wants to go to college in the absence of the admissions capacity constraint. Although the expected return of going to college is high, it's very costly for workers to abandon their jobs and return to school. In order to admit the additional college students, more seats have to be created. Using the fraction of GDP in 2011 that the fiscal budget on higher education accounts for (0.822%), it's possible to do a back-of-envelope calculation on the average cost of adding a seat.[47] Assuming that China continues with this fraction, it is worthwhile to accommodate these extra college students as long as the average cost of adding a seat is not greater than 176,000 yuan.[48]

11 Conclusion

This paper answers what the effects are of the higher education expansion reform on the evolution of college wage premium and investigates its distributional effects on the discounted lifetime wages of workers. To achieve this, I construct a dynamic labor market general equilibrium model in which workers make education and work decisions, and the representative firm decides how much skill and capital to use. I innovate in modeling the college admissions process. The key novel feature of the model is that low-skill workers not only have expectations on the evolution of skill prices but also on college admission rates, through which the policy change affects workers' decisions and labor market outcomes.

I structurally estimate the model and use the estimated model as a laboratory to conduct several counterfactual experiments. The first set of results highlights the importance of post-reform technological progress. It shows that the effects of large-scale higher education reforms on the college wage premium could be different, depending on whether one takes into account the interaction between the reform and post-reform technological progress. In the setting of this paper, in the presence of post-reform technological progress (both skill-neutral and skill-biased), the reform first increases the college wage premium and then decreases it. In contrast, in the absence of post-reform technological progress, I find that the effect of the reform on the college wage premium is negative and increases over time. Post-reform technological progress plays a role because it alters the future option value of attending college, such that the composition of low-skill workers who choose to go to college changes.

[47] The cost of a seat includes all kinds of resources devoted to a student, such as the average cost of hiring teachers and staffs, the average cost of building new institutions and purchasing new equipment etc.

[48] The assumption is that the government behaves exogenously of the model and finances the budget through a lump-sum tax.

The second set of results shows that although the higher education expansion reform has differential impacts on workers, it increases the discounted lifetime wages for the majority. Those who are induced to go to college by the reform (compliers) on average gain the most, whereas those who go to college even in the absence of the expansion (always-takers) lose a small fraction of their lifetime income, because they suffer from the decrease in high-skill prices as a result of the large increase in the supply of high-skill labor.

I also use the model to conduct two policy experiments and show that abandoning the admission capacity constraint would allow China to catch up with developed countries in terms of the share of high-skill workers in the working-age population much sooner and at a reasonable cost.

References

Bai, Chong-En, and Qian Zhenjie. "The Factor Income Distribution in China: 1978–2007." China Economic Review 21.4 (2010): 650-670.

Bian, Yanjie, and Lulu Li. "The Chinese General Social Survey (2003-8) Sample Designs and Data Evaluation." Chinese Sociological Review 45, no. 1 (2012): 70-97.

Blundell, Richard, David Green, and Wenchao Jin. "The UK Education Expansion and Technological Change." (2018).

Card, David, and Thomas Lemieux. "Can Falling Supply Explain The Rising Return to College for Younger Men? A Cohort-Based Analysis." The Quarterly Journal of Economics 116.2 (2001): 705-746.

De Chaisemartin, Clément, and Xavier D'Haultfoeuille. "Late Again with Defiers." (2012).

Dix-Carneiro, Rafael. "Trade Liberalization and Labor Market Dynamics." Econometrica 82, no. 3 (2014): 825-885.

Heckman, James J., Lance Lochner, and Christopher Taber. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." Review of Economic Dynamics 1.1 (1998): 1-58.

Eichen, Marc, and Zhang Ming. "Annex: The 1988 Household Sample Survey - Data Description and Availability." The Distribution of Income in China, pp. 331-336. Palgrave Macmillan, London, 1993.

Fiorini, Mario, and Katrien Stevens. "Assessing the Monotonicity Assumption in IV and Fuzzy RD Designs." Economics Working Paper Series-University of Sidney (2016).

Fleisher, Belton M., and Xiaojun Wang. "Skill Differentials, Return to Schooling, and Market Segmentation in a Transition Economy: the Case of Mainland China." Journal of Development Economics 73, no. 1 (2004): 315-328.

Goldin, Claudia D, and Lawrence F. Katz. "The Race between Education and Technology." Cambridge, Mass: Belknap Press of Harvard University Press, 2008. Print.

Hsieh, Chang-Tai, and Zheng Michael Song. "Grasp the Large, Let Go of the Small: the Transformation of the State Sector in China." No. w21006. National Bureau of Economic Research, 2015.

Katz, Lawrence F., and Kevin M. Murphy. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." The Quarterly Journal of Economics 107.1 (1992): 35-78.

Keane, Michael P., and Kenneth I. Wolpin. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." the Review of Economics and Statistics (1994): 648-672.

Keane, Michael P., and Kenneth I. Wolpin. "The Career Decisions of Young Men." Journal of Political Economy 105.3 (1997): 473-522.

Lee, Donghoon. "An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice." International Economic Review 46.1 (2005): 1-34.

Lee, Donghoon, and Kenneth I. Wolpin. "Intersectoral Labor Mobility and the Growth of the Service Sector." Econometrica 74.1 (2006)

Li, Hongbin, James Liang, and Binzhen Wu. "Labor Market Experience and Returns to Education in Fast Growing Economies." Unpublished paper, Tsinghua University (2016)

Li, Shi, Luo Chuliang, Wei Zhong, and Yue Ximing. "The 1995 and 2002 Household Surveys: Sampling Methods and Data Description." (2008).

Li, Shi, John Whalley, and Chunbing Xing. "China's Higher Education Expansion and Unemployment of College Graduates." China Economic Review 30 (2014): 567-582.

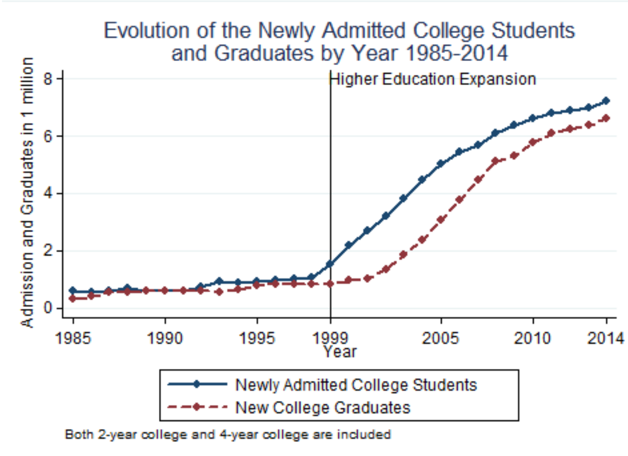
Llull, Joan. "Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model." The Review of Economic Studies 85, No. 3 (2017): 1852-1896.

Luo, Chuliang, Shi Li, Terry Sicular, Quheng Deng, and Ximing Yue. "The 2007 Household Surveys: Sampling Methods and Data Description." Rising Inequality in China: Challenges to a Harmonious Society (2013): 445-464.

Meng, Xin, Kailing Shen, and Sen Xue. "Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988–2009." Journal of Comparative Economics 41.1 (2013)

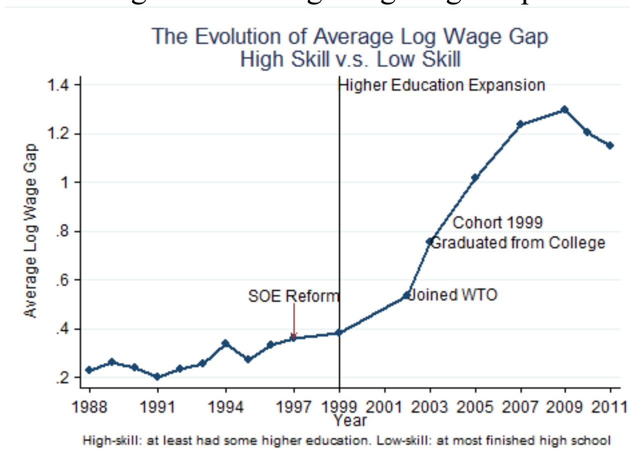
Mincer, Jacob. "Human Capital and Economic Growth." NBER Working Paper (1981): 803.

Figure 1: Annual Admitted College Students and Graduates



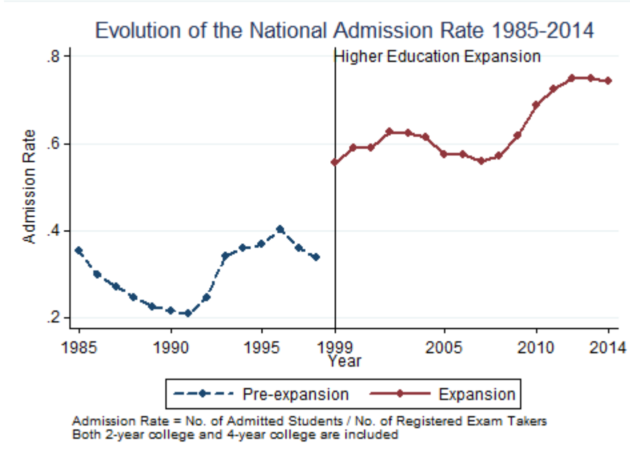
Notes: This figure shows the time series of the annual admitted college students and college graduates. Both two- and four-year college students are included. The data used are collected from the China Education Statistical Yearbooks. In the anticipation of increasing demand for high-skill workers, the Ministry of Education in China started to sharply increase the maximum number of students that could be admitted to college since 1999. Compared to 1998, the number of newly admitted college students in 1999 increased from about 1 million to 1.5 million. It then kept increasing for all subsequent years till 7.2 million in 2014. As a result, the supply of high-skill workers who have at least some college education has been increasing sharply since 2001.

Figure 2: Average Log Wage Gap



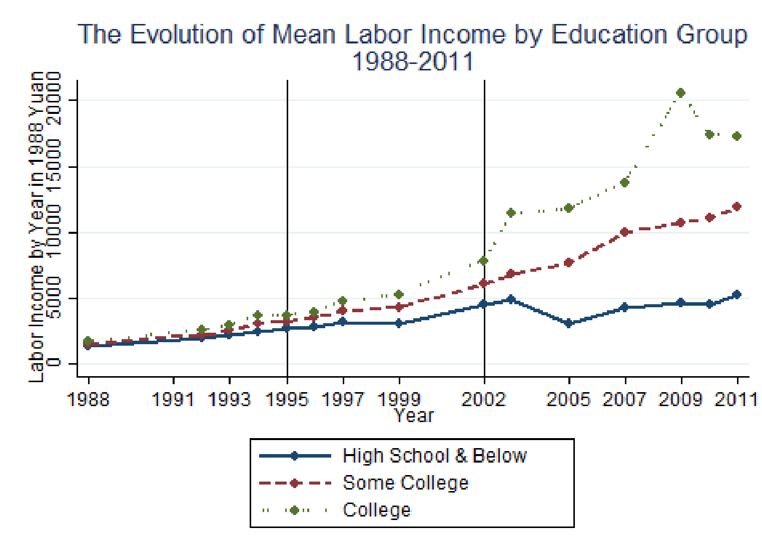
Notes: This figure shows the evolution of the college wage premium, defined as the average log wage gap between high- and low-skill workers, from 1988 to 2011. High-skill workers are those who have at least some college education whereas low-skill workers complete high school or below. The data used are from the Urban Household Survey (1988-1997), the China Household Income Project (1988-2007) and the China General Social Survey (2003-2012). Although the higher education expansion was massive, the college wage premium had been increasing since 1999 and started to modestly decrease in 2009. This suggests that there exist strong labor demand side forces that shift the relative demand for high-skill versus low-skill workers. Indeed, on the demand side, between 1997 and 2002, China aggressively privatized the state-owned enterprises (SOE), which potentially increased the wage gap (e.g. Fleisher and Wang, 2004). In addition, China joined the World Trade Organization (WTO) at the end of 2001, which could increased the wages of low-skill workers and eventually those of high-skill workers as more multinational firms enter China.

Figure 3: Annual College Admission Rate



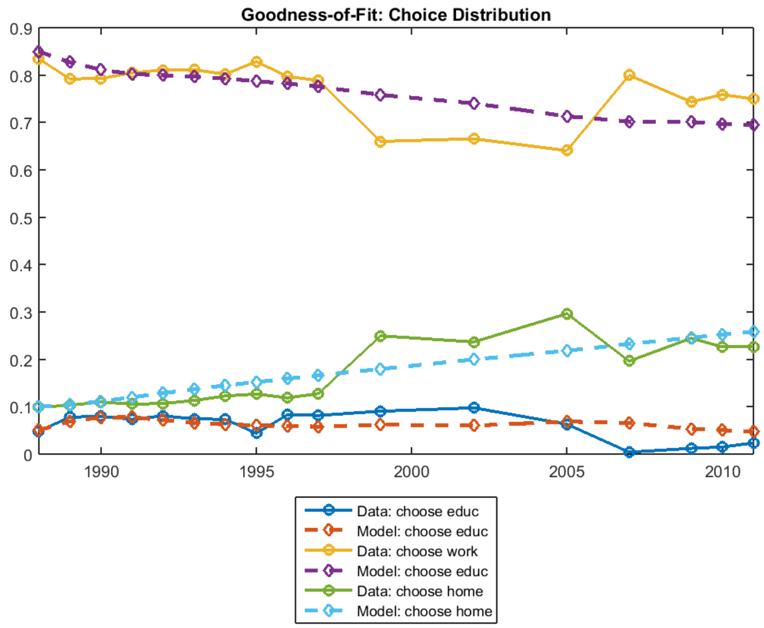
Notes: This figure shows the evolution of the annual college admission rate from 1985 to 2014 (both 2- and 4-year colleges are included). The admission rate is calculated as the ratio of the number of admitted students and the number of people who register for the College Entrance Exam (CEE) each year. The data used are from the China Education Statistical Yearbooks. Before 1999, the admission rate was below 40%. It decreased from 1985 to 1991 mainly because the number of admitted students didn't change much, but the number of registered exam takers increased due to increasing expected returns of college education. It then increased in the 1990s for two reasons. First, the number of admitted students had been slowly increasing. Second, starting in 1990, the value of the outside option of working in the private sector increased a lot, which decreased the number of registered exam takers. After 1999, the admission rate fluctuated around 60% and reached about 75% in recent years.

Figure 4: The Evolution of Mean Wages by Education Group



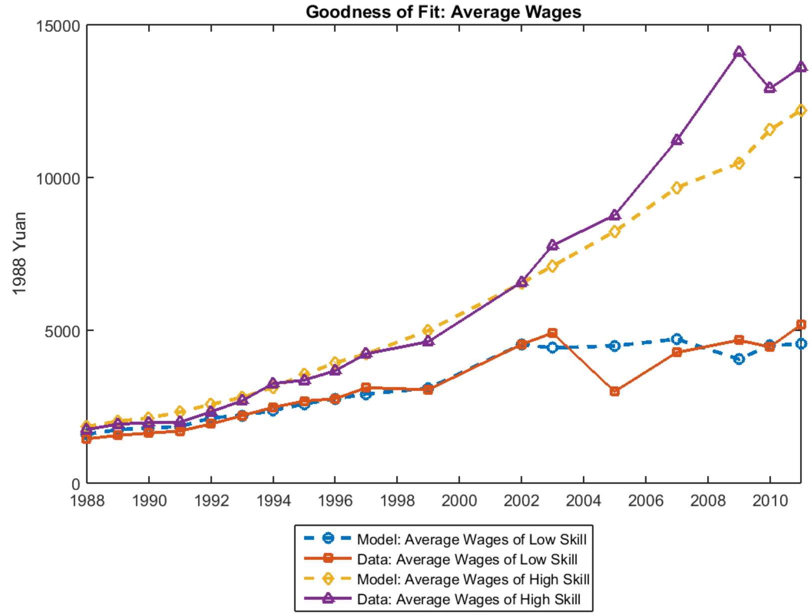
Notes: This figure shows the evolution of average wages for three education groups: those who have high school education and below, those who have some college, and those who are college educated. The data used are from the Urban Household Survey (1988-1997), the China Household Income Project (1988-2007) and the China General Social Survey (2003-2012). In addition, this figure shows that 1995 and 2002 are two important turning points for the evolution of wages of different skill groups.These two years correspond to the points where China went through major events. In 1995, China started reforms that aimed at downsizing the state-owned enterprises (SOE). In 2002, this process was complete. Moreover, China joined World Trade Organization (WTO).

Figure 5: Goodness-of-Fit of Choice Distribution



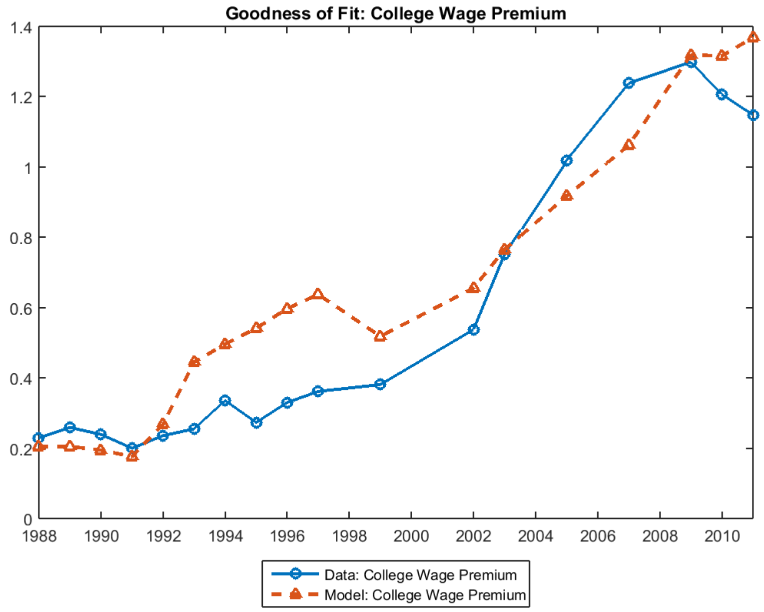
Notes: This figure shows how the choice distribution the model generates fits the data. The model does a good job for years before 1997, but not as good for years between 1999 and 2005. This is partly due to the quality of the data I use for those years. For years before 1998, I use the UHS data collected by the China Bureau of Statistics. The sampling and survey design is maintained in a consistent way across years. For years of 1998 onwards, the data I use are from the CHIP and the CGSS, which were carried out by different research teams and resulted in more noise in the pooled sample.

Figure 6: Goodness-of-Fit of Wage Distribution



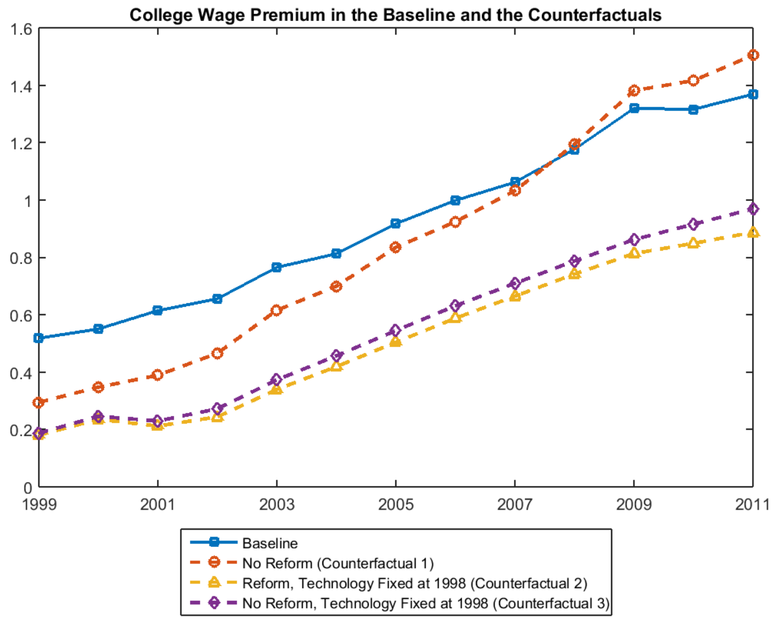
Notes: This figure shows the goodness-of-fit of the model in terms of the key wage moments. It compares how the actual and the simulated average wages of the high- and low-skill workers evolve. Although there is some discrepancy, the trend generated by the model tracks that of the data.

Figure 7: Goodness-of-Fit of the College Wage Premium



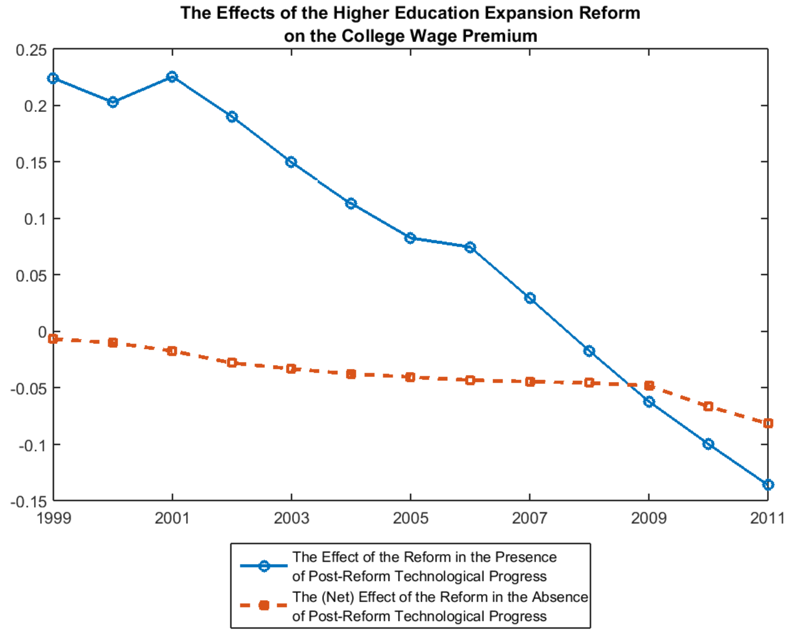
Notes: This figure shows the goodness-of-fit for college wage premium measured as the average log wage gap. The log transformation makes the wage premium between 1993 and 1999 look more volatile. For years between 1999 and 2011, where the model will be used for counterfactual exercise, the fit is reasonable.

Figure 8: The Effects of the Higher Education Expansion on the College Wage Premium



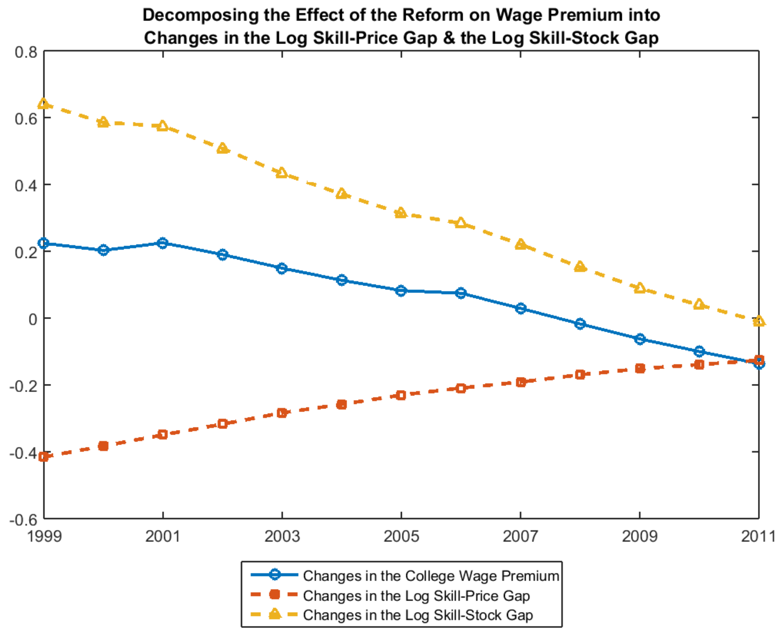
Notes: This figure shows the simulated trends of the college wage premium, defined as the average log wage gap, in different counterfactuals. The solid line is the baseline college wage premium generated from the estimated model, where all the parameters are the same as their estimated values and the admission process evolves as it is in reality. The dashed line with circles (Counterfactual 1) shows the counterfactual trend without the reform while keeping other factors the same. Since the only difference between the two cases is the reform, the comparison of them gives us the effect of the reform. The two lines at the bottom (Counterfactual 2 and 3) show the evolution of college wage premium in the absence of the post-reform technological progress. Without the post-reform technological progress, it the reform depresses the college wage premium immediately from the start. Comparing Counterfactual 1 and 2 to Counterfactual 3 respectively, the gap shows the net effect of the post-reform technological progress and that of the reform on the college wage premium. The gap between the baseline and Counterfactual 3 gives the total effects of both factors, which include the net effects of them and their interaction effect.

Figure 9: The Effects of the Higher Education Expansion Reform on College Wage Premium



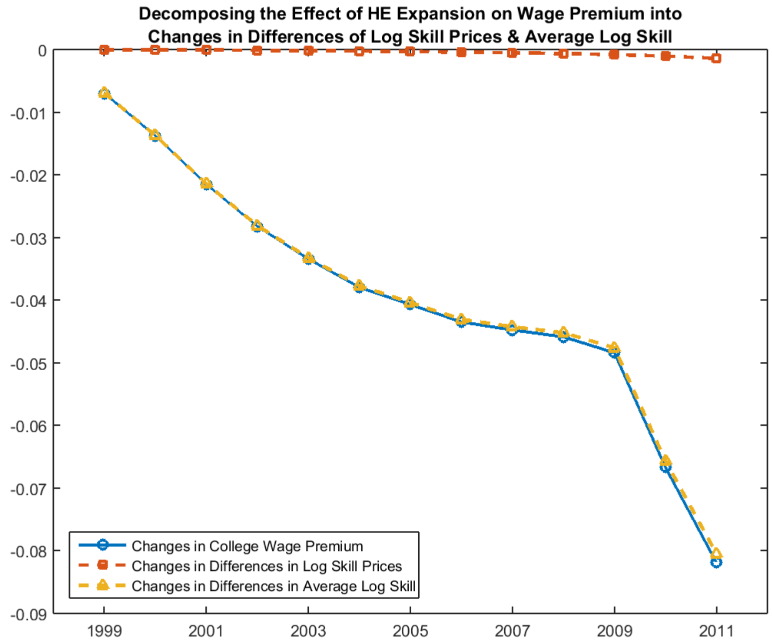
Notes: This figure shows the effects of the reform on the college wage premium with or without the post-reform technological progress. The solid line shows the gap between the baseline and Counterfactual 1. The dashed line shows the gap between Counterfactual 2 and 3. The effect of the reform in the presence of the post-reform technological progress: the reform increases the college wage premium before 2008 with a diminishing effect from about 0.16 log points in 1999 to virtually zero in 2007. It then decreases the wage premium with an amplifying effect from 0.06 to about 0.14 log points. On average, the reform increases the college wage premium by 21%, with a yearly increment of 0.07 log points. The effect of the reform in the absence of the post-reform technological progress: without the post-reform technological progress, it the reform depresses the college wage premium immediately from the start and the effect is increasing over time till about 0.08 log points. On average, the reform decreases the college wage premium by 7% per year.

Figure 10: Decomposing the Effect on College Wage Premium into Changes in the Skill-Price Gap and the Average Skill-Stock Gap in the Presence of Post-Reform Technological Progress



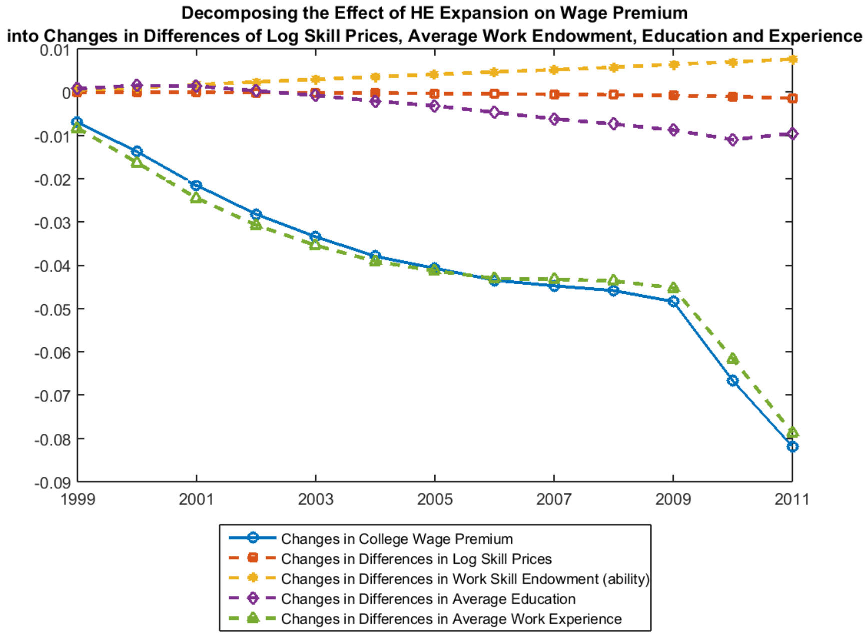
Notes: To understand what drives the effect of the reform on the college wage premium when there is post-reform technological progress, this figure presents the decomposition of the reform's effect. The effect is decomposed into changes in the skill-price gap and changes in the average skill-stock gap. The decomposition shows that the reform's effect is determined by two components: changes in the skill-price gap and changes in the skill-stock gap. Holding the technological progress the same in the baseline and Counterfactual 1, the reform increases the supply of high skill relative to low skill and therefore narrows the skill-price gap. In the meantime, by allowing low-skill workers who on average have more skill to go to college and become high-skill, the reform widens the skill-stock gap. The widening effect becomes smaller over time because the reform creates a large and increasing supply of young and inexperienced high-skill workers. As they enter the labor market, the average skill stock of high-skill workers decreases, which creates a countervailing effect on the skill-stock gap.

Figure 11: Decomposing the Net Effect on College Wage Premium into Changes in the Skill-Price Gap and the Average Skill-Stock Gap (in the Absence of Post-Reform Technological Progress)



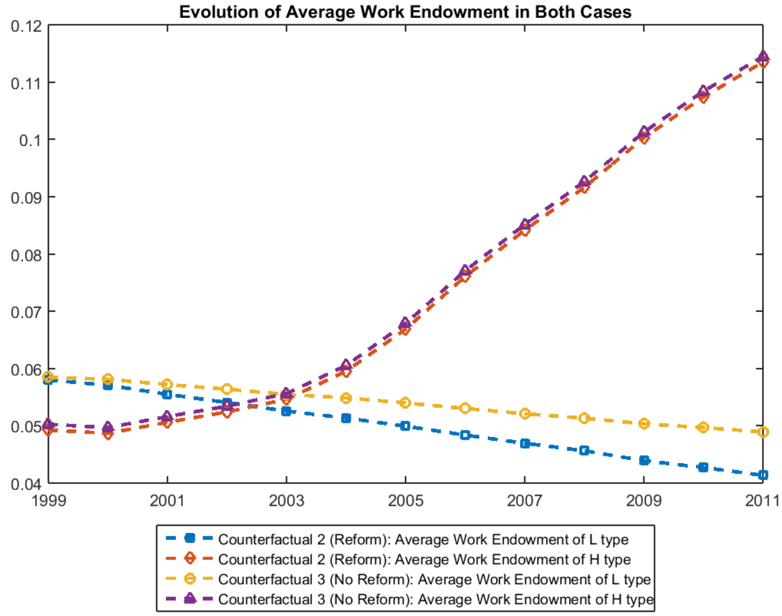
Notes: To understand what drives the effect of the reform on the college wage premium in the absence of the post-reform technological progress, this figure presents the decomposition of the reform's effect. The effect is decomposed into changes in the skill-price gap and changes in the average skill-stock gap. The decomposition shows that changes in the average skill-stock gap drive the net effect of the higher education expansion reform on the college wage premium. Changes in the skill-price gap also serve as a force that depresses the college wage premium but is not the driving force.

Figure 12: Decomposing the Net Effect on College Wage Premium into Changes in the Skill-Price Gap, the Average Work Endowment Gap and the Average Education and Work Experience Gaps



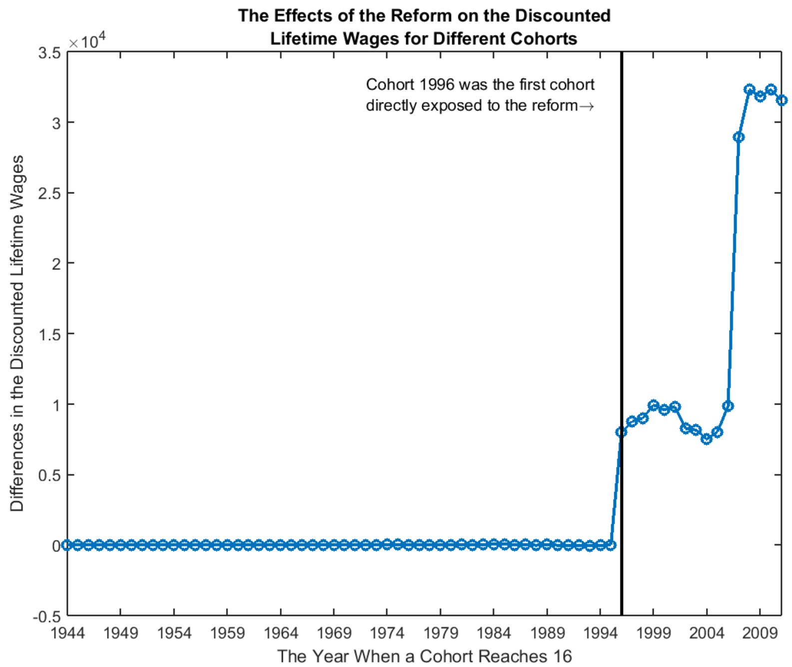
Notes: This figure shows the further decomposition of the changes in the average skill-stock gap into changes in the average work endowment gap, changes in the average education gap and changes in the average work experience gap. The result shows that changes in the average education and work experience gaps drive changes in the average skill-stock gap. Overall, on average, changes in the skill-price gap account for 0.93% of the net effect of the higher education expansion on college wage premium, whereas changes in the average skill-stock gap account for 99.07%. In addition, changes in workers' composition account for -9.71% of the net effect. Changes in the average education gap account for 5.61% and changes in the average work experience gap account for 103.31%.

Figure 13: Comparing the Evolution of Average Work Endowment in Both Cases



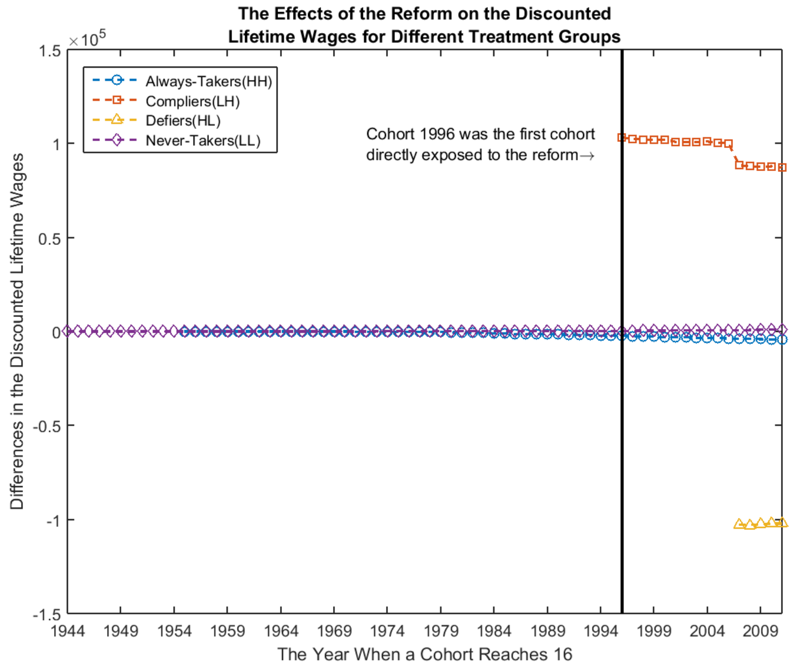
Notes: To understand why the changes in the work-endowment gap in Figure 12 are positive, this figure presents the evolution of average work endowment in Counterfactual 2 and 3. As the reform allows more workers who are on average with lower work endowment to go to college, the average work endowment gap widens. From the figure, we can see that the widening of this gap is primarily because the average work endowment of low-skill workers decreases.

Figure 14: Net Effect of the Higher Education Expansion on the Discounted Sum of Lifetime Wages By Cohort



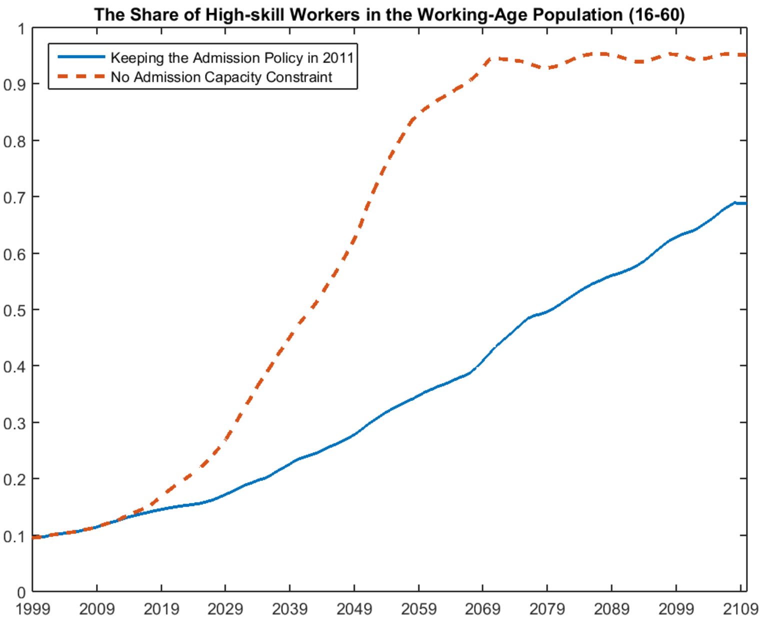
Notes: This figure shows the effects on the discounted sum of lifetime wages by cohort. Each cohort is indexed by the year when they first enter the economy (age 16). For cohorts 1944 through 1954, the effect is exactly zero because all of these cohorts exit the economy before the higher education expansion reform starts in 1999. For cohorts 1955 through 1989, they overlap with the reform by at least one year. The effect is increasing in the number of years they overlap, and their gains are close to zero but positive. Cohorts 1990 through 1995 actually lose modestly by 0.15%. Cohorts 1996 through 2011 gain the most because they are directly exposed to the higher education expansion reform. The expansion was already carried out when they reached 19, the first year when the majority of people complete high school and take the College Entrance Exam (CEE). They on average gain by about 87% compared to the counterfactual without the reform.

Figure 15: Net Effect of the Higher Education Expansion on the Discounted Sum of Lifetime Wages By Cohort and Treatment Group



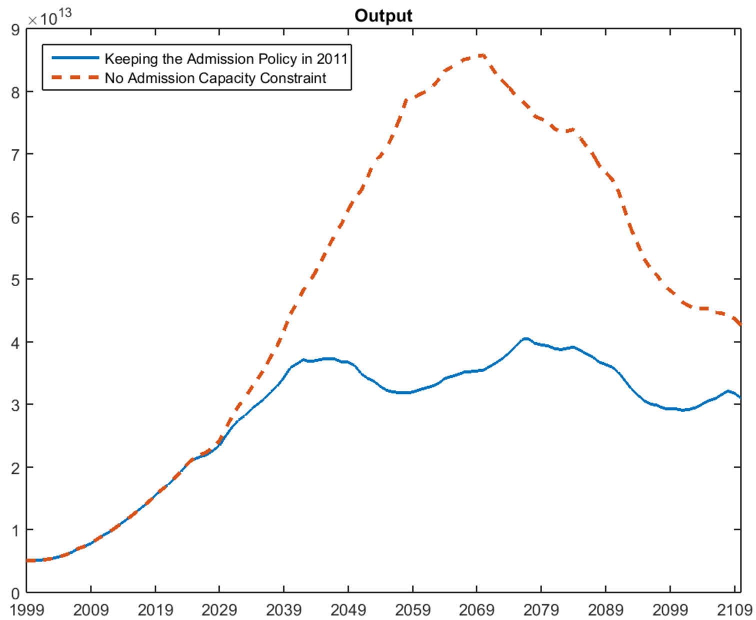
Notes: this figure shows the effects of the reform on the discounted sum of lifetime wages by treatment group. There are four groups in general: people who are induced to go to college by the reform (compliers), those who go to college with or without the reform (always-takers), those who are induced to not go to college as a result of the reform (defiers) and those who never go to college (never-takers). From the figure, we can see that the group that gains the most are compliers. As for the always-takers, they lose modestly by 2.6% on average because they suffer from the decrease of the high-skill prices due to the large increase in the supply of high-skill labor. For never-takers, they gain modestly by about 8.7% because as the share of low-skill labor decreases, the demand for low-skill labor increases and the skill price increases.

Figure 16: Comparing the Evolution of the Share of High-Skill Workers in Policy 1 and 2



Notes: This figure shows the evolution of the share of high-skill workers in the policy experiment 1 and 2. In policy experiment 1, China will reach 30% of high-skill workers in the working-age population in 2052, whereas in policy experiment 2, China will reach this target in 2031.

Figure 17: Comparing the Evolution of the Output in Policy 1 and 2



Notes: this figure shows the evolution of GDP in policy experiment 1 and 2. Although the technological progress is fixed at 2011 level, the economy grows as the skill stock increases. However, growth is not going to last forever because as the share of high-skill workers increases, the marginal product of high skill decreases which depresses the demand. In the meantime, the marginal product of low skill increases will become extremely high as the share of low-skill labor decreases. At some point, the high-skill labor that is employed economy will be low enough such that the GDP decreases.

Table 1: Evolution of Employment Shares by Education Category in Urban China

	88-90	91-95	96-99	00-03	04-08	09-11
College and Above						
Share	-	0.050	0.067	0.081	0.094	0.116
Some College						
Share	-	0.082	0.137	0.179	0.192	0.223
At Least Some College						
Share	0.110	0.166	0.204	0.260	0.286	0.340
High School and Below						
Share	0.890	0.834	0.796	0.740	0.714	0.660
Pooled Sample						
Female	0.507	0.505	0.508	0.509	0.510	0.504
Observations	119159	125893	37013	19561	6947	9313

Notes: This table shows the evolution of employment shares by education category in urban China. The data used are from the Urban Household Survey (1988-1997), the China Household Income Project (1988-2007) and the China General Social Survey (2003-2012). For 1988-1991, the coding of education category in the UHS does not distinguish between some college and college. Both the shares of people who are college educated and above and those who have some college education increase. Following the higher education expansion in 1999, we see the sharpest increase in the share of people who have at least some college education.

Table 2: Preference Parameter Estimates

Panel A: Flow Utility Parameters					
Education / Home	Cost of College	Extra Cost of Grad School	Returning Cost	Persistence	
	t_1	t_2	κ_1	κ_3	
	3178	516	-2319	757	
	(359)	(45)	(211)	(89)	
Work	Non-pecuniary Benefit	Low-Skill Educ	Low-Skill Exper	High-Skill Educ	High-Skill Exper
	γ	β_1^L	β_2^L	β_1^H	β_2^H
	101	0.21	0.15	0.25	0.08
	(11)	(0.026)	(0.014)	(0.028)	(0.027)
	Panel B: Preference Shock Parameters				
	Education	σ_{η_1}	0.01		
			(0.0013)		
	Work	σ_{η_2}	0.07		
			(0.004)		
	Home	σ_{η_3}	0.26		
			(0.022)		

Notes: This table shows the estimates of the flow utility parameters and preference shock parameters. The estimates are shown by the three available options that a worker can choose each period. Standard errors are in parentheses. Please refer to Section 7.1 for the discussion on the parameter estimates.

Table 3: Production Function Parameter Estimates

Production Function Parameters		
Low-Skill Labor Factor Share		
Intercept (before 1995)	α_{10}	0.4
		(0.036)
Slope 1 (1995-2001)	α_{11}	-0.003
		(0.0008)
Slope 2 (2002-2011)	α_{12}	0.007
		(0.0002)
Composite Factor Share		
Intercept (before 1995)	α_{20}	0.4
		(0.065)
Slope 1 (1995-2001)	α_{21}	-0.002
		(0.0004)
Slope 2 (2002-2011)	α_{22}	0.001
		(0.0003)
Elasticity Parameters		
Low-Skill Labor - Composite	σ	0.85
		(0.016)
High-Skill Labor - Capital	ν	0.4
		(0.04)

Notes: This table shows the estimates of the production function parameters. Standard errors are in parentheses. Please refer to Section 7.1 for the discussion on the parameter estimates.

Table 4: Conditional Type Probability Parameters

Endowment	Type 1		Type 2		Type 3	
Education	ω_1^1	0.01	ω_1^2	-0.32	ω_1^3	0.25
		(0.001)		(0.036)		(0.015)
Work	ω_2^1	0.02	ω_2^2	-0.11	ω_2^3	0.24
		(0.004)		(0.037)		(0.017)
Home	ω_3^1	-0.01	ω_3^2	0.04	ω_3^3	-0.46
		(0.002)		(0.01)		(0.051)

Probability Parameters	Type 1		Type 2		Type 3	
Intercept	-	-	π_{02}	0.697	π_{03}	-0.111
	-	-		(0.007)		(0.005)
Percentiles of Education at 16	-	-	π_{12}	0.106	π_{13}	0.029
	-	-		(0.028)		(0.011)
Percentiles of Parents' Education	-	-	π_{22}	-1.19	π_{23}	1.31
	-	-		(0.012)		(0.009)

Notes: This table shows the estimates of the conditional type probability parameters. Standard errors are in parentheses. Please refer to Section 7.1 for the discussion on the parameter estimates.

Table 5: Admission Process Parameters

Year	88	89	90	91	92	93
Percentiles of Education at 16	0.39	0.36	0.40	0.42	0.97	1.08
	(0.023)	(0.024)	(0.026)	(0.027)	(0.03)	(0.028)
Percentiles of Parents' Education	0.26	0.37	0.38	0.51	0.49	0.38
	(0.08)	(0.081)	(0.092)	(0.093)	(0.107)	(0.104)

Year	94	95	96	97	98	99
Percentiles of Education at 16	1.53	1.35	1.01	0.92	2.95	2.28
	(0.028)	(0.029)	(0.027)	(0.03)	(0.031)	(0.03)
Percentiles of Parents' Education	0.37	0.38	0.45	0.50	0.39	0.43
	(0.109)	(0.112)	(0.103)	(0.106)	(0.102)	(0.102)

Year	00	01	02	03	04	05
Percentiles of Education at 16	2.39	2.71	1.60	1.67	1.42	1.44
	(0.035)	(0.035)	(0.037)	(0.036)	(0.04)	(0.44)
Percentiles of Parents' Education	0.32	0.44	0.54	0.45	0.43	0.33
	(0.1)	(0.112)	(0.114)	(0.115)	(0.117)	(0.114)

Year	06	07	08	09	10	11
Percentiles of Education at 16	0.90	2.01	1.03	1.15	0.76	1.88
	(0.432)	(0.509)	(0.705)	(0.39)	(0.08)	(0.056)
Percentiles of Parents' Education	0.39	0.41	0.62	0.45	0.23	0.37
	(0.109)	(0.108)	(0.127)	(0.136)	(0.067)	(0.072)

Notes: This table shows the estimates of the admission process parameters. Standard errors are in parentheses. Please refer to Section 7.1 for the discussion on the parameter estimates.

Appendix

Solving the Model

The rich heterogeneity and the general equilibrium feature of the model makes it computationally-intensive to solve the model. This section describes how the model is solved.

1. Given a set of reasonable model parameters, solve the dynamic programming problem for each cohort, each age and every possible point in the support of the state space. The outcome of the first step is a matrix of coefficients that are cohort and age specific. The coefficients come from a second order polynomial regression that uses contemporary state variables to predict the Emax function evaluated at the states. The regressors include a constant, education, experience, ability type dummies, skill prices and admission rate at t, and all second-order terms such as squares and interactions. The regression is used to approximate the Emax function, which is the expectation of the value function at age a with respect to the distributions of unobservables. The model economy starts in 1988 and ends in 2011 and there are 45 overlapping generations each year aged from 16 to 60.

2. Solve the model by computing a sequence of equilibrium prices that is consistent with the model parameters and the beliefs imposed by the forecasting rule.

    (a) The following inputs are required for this step: Initial conditions on the state space distribution, time series of output and capital rental prices from 1988 to 2011. I simulate 1500 people for each cohort starting from 1988. This means for each year, we have a cross section of 67500 observations. To account for cohort size variations, I weight each cohort by its cohort size.

    (b) Guess a vector of skill prices for 1988. Use this guess, the combinations of states from (a), simulated shocks to alternatives, guessed forecasting rule parameters and coefficients obtained from step 1, solve the DPP for everyone alive at 1988 and obtain the choice distribution. Individual choices are then aggregated to construct the total skill supply of skills for high- and low-skill workers. Given such skill supply, capital rental prices, output and the model parameters, we can solve for z_t and K_t from the first order condition of capital use (i.e. demand for capital) and the resource constraint. Applying these solved values to the first order conditions of high skill and low skill, we can solve for the skill prices. These prices are in general different from the initial guess, which means we should update the guess of skill prices using them. This is done repeatedly for 1988 until the skill prices solved from the first-order conditions are very close to those used as the guess.

(c) Repeat (b) for 1989 through 2011 and we can get a time series of skill prices. These are not yet the equilibrium prices since they are obtained under a set of guessed belief parameters. To update the beliefs, we need to estimate forecasting rule as a vector autoregression (VAR) using the time series of skill prices. The estimates of the VAR then become the updated belief parameters. Repeat (a) and (b) and keep updating such parameters until they are stable. The stable belief parameters are then consistent with the equilibrium of the model.

Weighting the CGSS Data

For the weighting to be reasonable, I use the whole sample of the CGSS containing information on both rural and urban areas.

1. First weight CGSS 2005-2008 to the 2005 Population Survey, and CGSS 2010-2012 to the 2010 census by the sex-age-rural/urban-stratum cell. It's worth noting that the sampling frame used in 2005-2008 is different from that used in 2010-2012, which affects the choice of strata. For 2005-2008, the largest strata are Eastern China, Central China and Western China. For 2010-2012, the largest strata used are the three municipalities in Eastern China (including Beijing, Shanghai and Tianjin) and the rest of China. I match the CGSS to the information in the 2005 Population Survey and the 2010 Census according to the proper divisions of the strata.[49]

2. On top of Step 1, which matches information by sex-age-rural/urban-stratum cell, this step matches information by sex-educ-region cell.

   (a) I calculate a scale factor such that after multiplying it by the weight constructed in the first step, the population shares by education category in the CGSS sample match those in the aggregate data. I can first inflate the aggregate data by the sampling proportion

[49] For the weighting procedure to be reasonable, I must make sure that, within each stratum, the smaller sampling units are similar in terms of population. This assumption seems less valid for 2010-2012 than for 2005-2008 given the heterogeneity across the rest of China after excluding the three big municipalities in Eastern China. Although the coding of counties is confidential and a direct check is not feasible, the CGSS sampling design team conducted factor analysis based on population density, proportion of non-agriculture population and county GDP per capita for all counties besides the three municipalities. and ranked them. Moreover, the number of counties were made as similar as possible for all strata (50 in total) before sampling. This procedure resembles that done for the CPS and was used to make sure the strata as homogeneous as possible.

because the aggregate data is already adjusted by weight so as to be representative. Thus, everyone represents the same number of people in the population.

(b) Similar to the first-stage adjustment in the CPS, I count the sample in each sex-educ-region cell using the weight that I constructed before to inflate it so as to match the population counts. If the sample cell is under-representative, the adjustment factor will inflate it, and deflate it otherwise.

$$adjustment\ factor\ of\ cell\ j = \frac{population\ counts\ in\ cell\ j}{weighted\ sample\ counts\ in\ cell\ j}$$

(c) After applying the adjustment factor, urban, age, female are still generally consistent across years, which is reassuring. These variables are chosen because they should be stable within a short period of time (5 years).

Table A.1: Descriptive Statistics after Weighting (2005-2008)

Year	2005		2006		2008	
Variable	Mean	Population Survey	Mean	Aggregate Stats	Mean	Aggregate Stats
Age	42.2314	-	41.9951	-	41.5525	-
College	0.0183	0.0227	0.0193	-	0.0276	-
Some College	0.0493	0.0444	0.0568	-	0.0541	-
HS Equivalent	0.1321	0.1321	0.1379	-	0.1452	-
Grad Degree	0.0012	0.0017	0.0011	-	0.0010	-
At Least Some College	0.0688	0.0688	0.0772	0.0769	0.0828	0.0829
Female	0.5053	0.4847	0.5007	0.4848	0.4998	0.4853
Work Exper	25.7510	-	25.4966	-	24.9802	-
Urban Area	0.4219	0.4299	0.4255	0.4434	0.4418	0.4699

Notes: This table shows the weighted averages of key variables in the CGSS (2005-2008) and their corresponding values in the representative samples. CGSS 2005 is matched to the 2005 Population Survey. Samples of the other years are matched to the China Statistics Yearbooks. The mean for each education category in the table is the population share of that category. The population shares by detailed education category are only available in the 2005 Population Survey.

Table A.2: Descriptive Statistics after Weighting (2010-2012)

Year	2010		2011		2012	
Variable	Mean	Census	Mean	Aggregate Stats	Mean	Aggregate Stats
Age	41.1218	-	41.5234	-	41.6160	-
College	0.0309	0.0448	0.0326	-	0.0421	-
Some College	0.0767	0.0670	0.0802	-	0.0786	-
HS Equivalent	0.1529	0.1592	0.1570	-	0.1634	-
Grad Degree	0.0037	0.0041	0.0043	-	0.0025	-
At Least Some College	0.1113	0.1159	0.1172	0.1176	0.1232	0.1238
Female	0.4948	0.4873	0.4942	0.4874	0.4939	0.4875
Work Exper	24.4411	-	24.8085	-	24.8514	-
Urban Area	0.5270	0.4995	0.5361	0.5127	0.5248	0.5257

Notes: This table shows the weighted averages of key variables in the CGSS (2010-2012) and their corresponding values in the representative samples. CGSS 2010 is matched to the 2010 Census. Samples of the other years are matched to the China Statistics Yearbooks. The mean for each education category in the table is the population share of that category. The population shares by detailed education category are only available in the 2010 Census.