China’s College Expansion Policy: 
Ability Selection and Labor Market Effects*

Shihui Ma
University of Minnesota

JOB MARKET PAPER
November 29, 2014

Abstract

In 1999, the Chinese government embarked on a program to increase the entry class to tertiary education by 42% from the previous year; the college admission rate stayed at the higher level since then. The expansion of college education represents a large and exogenous increase in supply of the college graduates to the labor market. This paper identifies the key role of the relative college labor supply in driving the changes of college wage premium after the expansion program. Assuming imperfect substitutability of workers in different education and age groups, I propose an overlapping-generation model with endogenous educational choice to account for college premium trends in distinct demographic groups. The estimation results provide the basis for evaluating the welfare effects of the college expansion in different subgroups.

Keywords: Return to Education, Higher Education, Policy Experiments
JEL classifications: I2, J24, J31

*I am very grateful to my advisor, Fatih Guvenen for his continuous guidance and support on the project. I thank Jonathan Heathcote, Ellen McGrattan, Paul Glewwe, and Alessandra Fogli for inspiration, encouragement, and support. I also thank the participants of the Labor Workshop at the University of Minnesota for helpful comments. Please check http://shihuima.weebly.com/research.html for the latest version.
1 Introduction

In June 1999, the State Council of China decided to increase the entry class to tertiary education by 42% from previous year, and the admission rate stayed at this higher level since then. This paper assesses the role of supply, driven by the exogenous college expansion policy, in accounting for the large change in the returns to college education over different subgroups. The various trends of changing returns to college education in different age and residential groups motivates my assumption of imperfect substitutability between different age and education groups in production, and my estimation of the elasticities are comparable with the estimates by Katz and Murphy (1992) and Card and Lemieux (2001). To further account for the differential dynamics of the college premium across age and residential groups, and to evaluate the welfare gain or loss of the college expansion policy, I propose an overlapping-generation model with endogenous college education choice. And in the end, I conduct counterfactual analysis and solve for the optimal college expansion size.

As documented by many literatures, the wage of college graduates compared to high school graduates increased dramatically after the opening up in 1978. However, a remarkable trend in the China’s labor market in the recent years is the overall declining college premium, and this is especially phenomenal among the young college graduates. Using a longitudinal data set China Health and Nutrition Survey (CHNS) that include 26,000 individuals in eleven provinces of China, I documented the fact that the college premium for the young graduates (21-25 years old) begin to decrease immediately after the first cohorts of college students after the 1999 college expansion policy graduated in 2003, while the college premium for the older cohorts kept on growing until late 2000s. The panel data drawn from the China Health and Nutrition Survey (CHNS) also allow me to separately identify the extent to which the life-cycle earnings profile is determined by the curvature of the working-experience accumulation versus the changing efficiency of the labor input.

To account for the divergent college premium trends for different age groups, I build a heterogeneous-agent model with imperfect substitutability of workers in different education and age groups. The production technology in my model is based on the seminal analysis of Card and Lemieux (2001), which explains the diverging trends of college premiums for old and young workers by a model of imperfect substitution between different age workers with same education. In particular, they segregate the labor inputs by five-year age groups and education groups. For each age group, they separately estimate the college premiums, and
calculate the corresponding relative college labor supply by a ratio of weighted sum of hours worked by equivalent college and high school workers. Adopting a similar strategy for measuring the effective supply of college labor and college premium, I could estimate the demand elasticity of substitution between workers with the same education in different age groups, and the elasticity of substitution between college and high school labors. The estimated elasticity of substitution between college and high school labor is in the range of 1.1 to 1.9, comparable to the estimations by Katz and Murphy (1992) and Card and Lemieux (2001). And my estimation of elasticity of substitution between different age groups are in the range of 4.4 to 7.6, generally higher than the estimates from Card and Lemieux (2001).

In order to analyse the effect of the college expansion policy on residential groups with different financial resources and ability distributions, I further include the mechanism of college education choice with heterogeneity in abilities and initial assets. At high school graduation, individuals are endowed with abilities that will affect their labor earnings through effective labor inputs and the chance of access to higher education. They are also endowed with different initial assets. The ability threshold to college education is exogenously determined by the government policy, and the expansion of college entry class is equivalent to decreasing the college entry ability threshold if we assume a constant ability distribution over time.

The huge exogenous college expansion will induce two channels that affect the college wage premiums. Firstly, the increase in the relative college labor supply will decrease the relative college labor price. Secondly, the decrease in the ability threshold for college entry will affect both the average ability of college and high-school labor, and that will change the effective labor supply in both education groups. The endogenous college education choice could correspond to both the above channels, where after the college expansion policy was announced, individuals perfectly expect the potential decrease in college premium out of the first supply force, and make their college education choice on their initial wealth and ability levels accordingly.

The dynamic model and numerical solutions could correspond to the evolution of the labor

---

1. They divided workers into five educational groups: high school dropouts, high school graduates, workers with some college, college graduates, and workers with a postgraduate degree, then calculate the total supply of high-school and college equivalent labor by a weighted sum of different education groups’ total annual hours, the weight being the regression coefficients of each group’s average wage on the average wages of high school and college graduates.

2. By comparison, Katz and Murphy (1992) report an estimate of elasticity at 1.4 and the estimates from Card and Lemieux (2001) is in the range of 1.1-1.6.

3. See documented evidence in Section 2.2
market over the recent years. The overall college premiums constantly increased initially; the college expansion exogenously increased the supply of young college graduates and drove down the wage for inexperienced graduates while the wage for prime age graduates continues to rise.

From a policy perspective, the model can be used to analyse the welfare effects of the college expansion policy on different subgroups. The counter-factual analysis on an alternative policy of limiting the growth of the higher education sector lay the foundation for a policy analysis of the development of higher education in China as well as other rapidly developing countries.

**Related Literature** My paper is closely related to the literature of the labor supply effects on the wage structure, such as the cohort effects in Welch (1979), changes in college education by age groups and cohorts in Katz and Murphy (1992) and Card and Lemieux (2001), and changes in return to experiences in Jeong, Kim, and Manovskii (2014). However, instead of a clear exogenous supply change, all the above studies are involved with some endogenous response, such as the schooling and birth choice, to the underlying social-economic factors. This paper features a dramatic exogenous policy change, as a sharp identification strategy, to estimate the production demand elasticities on different categories of labor input. As a result, the estimated elasticities are generally comparable with the previous literature, with a slightly higher elasticities for different age groups.

The second strand of research is on the structural model on higher education choice with heterogeneous ability, such as Heckman, Lochner, and Taber (1998) and Epple, Romano, and Sieg (2006), where students differ in ability and family income, and school enrolment decisions are made with perfect information. My model builds on the framework and combines it with heterogeneous income profiles (HIP) assumptions, as in ?, where workers with different abilities have various income growth rates to study how ability selection affects measured college wage premiums.

Finally, the work is motivated by various empirical studies on the evolutionary policies on education sectors, such as Duflo (2001) and Lui and Suen (2005), and Wang, Fleisher, Li, and Li (2007) as well as empirical estimates of the education returns in China for recent years, as in Zhang, Zhao, Park, and Song (2005), Lee and Malin (2009), and de Brauw and Rozelle (2008).

The rest of the paper is organized as follows. Section 2 describes the data and discusses the
background of China’s tertiary education evolution system as well as the college expansion policy, we also explain why the policy might affect different residential subgroups differently. Section 3 set up the two-sector economy and establish the key variables of elasticities we need to estimate from microlevel data. Section 4 illustrates the strong relationship between the relative supply of college labor with the college premium and reports the structural estimation results for the elasticities. Section 5 provides the calibration and quantitative results. In Section 6, I apply the calibrated results to study the counter-factual effects of policy measures of limiting the growth of the higher-education sector.

2 Data Sources and Institutional Background

This section provides a brief introduction of the two survey data sets and a general description of the college education system in China. My main data sources are drawn from a longitudinal data set of China Health and Nutrition Survey (CHNS) with 9 survey waves from 1989 to 2011, I use another cross-sectional data from China General Social Survey (CGSS) for further information after 2003. I also present a brief history of evolution of college education as well as its private financing in China since the 1980s.

2.1 Data Sources

To quantify the effects of college expansion policy on both college enrolment for the rural and urban students and increased human capital within subgroups requires micro-level data with detailed information about the income, residential status, education and employment of workers. To satisfy the requirements, two micro-level survey data sets are used in this paper, a longitudinal data set China Health and Nutrition Survey (CHNS), and a cross-sectional data set (CGSS).

The CHNS data is an unbalanced panel household survey with refreshment that includes 26,000 individuals in nine provinces of China, including Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong, as shown in the Figure 1.

---

4The CHNS is jointly conducted by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. More detailed information about the CHNS can be found at the CPC UNC website: http://www.cpc.unc.edu/projects/china

5The sample in 2011 is expanded by adding Shanghai, Beijing and Chongqing, the three mega municipalities in China.
provinces sampled are broadly representative of China’s regional variation. Of these provinces, two are dynamic high-growth ones in China’s east coastal region (Jiangsu and Shandong); two are located in the northeast region (Liaoning and Heilongjiang), with one heavily industrialized (Liaoning); three are located in the middle region (Henan, Hubei, and Hunan); and two are in the southwest with populations concentrated of ethnic minorities (Guangxi and Guizhou). These surveys provide wage, education, employment and demographic information for the survey years 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011.

Figure 1: Map of Survey Regions

The main advantages of the CHNS relative to other publicly-available data sets are that it includes data on both rural and urban areas, and its survey waves cover all the relevant years of the college expansion policy, as well as the recent trends. One potential concern is whether the unweighed panel is representative of the whole China on the key dimensions, including the college enrolment, labor supply and wage differentials between rural and urban residents. To allay the concern, in Table 1 we compare summary statistics from the nine waves of the CHNS (first panel) to the four census and China Statistical Yearbooks (CSYs) (second panel).

---

6 Other publicly available data sets like the Chinese Household Income Project survey (CHIP) and China General Social Survey (CGSS) only include years before or after the college expansion, or only urban households like the Urban Household Income and Expenditure Survey (UHIES).

Comparison of the two panels of Table 1 shows that the CHNS oversampled high-school graduates relative to college graduates, and in general oversampled males. Therefore, in each survey year, we separate the CHNS sample into cells by four education types and eight age groups, so that in each education-age cell, the proportion is comparable with the national survey. The share of workers with education levels of primary-school and below, middle-school graduates, high-school graduates and college/university graduates in the weighted sample remain roughly in line with census data in each of the 5-year age groups.

Table 1: Summary Statistics of CHNS Sample

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Provinces</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Full-time workers</td>
<td>3,670</td>
<td>3,128</td>
<td>2,792</td>
<td>2,790</td>
<td>2,842</td>
<td>3,613</td>
<td>3,806</td>
<td>4,197</td>
<td>5,556</td>
</tr>
<tr>
<td>Percentage of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Male workers</td>
<td>59.62</td>
<td>58.95</td>
<td>60.06</td>
<td>59.17</td>
<td>56.32</td>
<td>56.57</td>
<td>56.85</td>
<td>56.37</td>
<td></td>
</tr>
<tr>
<td>-Urban hukou</td>
<td>69.06</td>
<td>71.14</td>
<td>68.02</td>
<td>64.60</td>
<td>57.30</td>
<td>37.75</td>
<td>37.28</td>
<td>35.62</td>
<td>46.87</td>
</tr>
<tr>
<td>-High-school graduates</td>
<td>24.55</td>
<td>26.69</td>
<td>29.48</td>
<td>35.25</td>
<td>37.16</td>
<td>27.48</td>
<td>26.12</td>
<td>23.73</td>
<td>27.27</td>
</tr>
<tr>
<td>-College graduates</td>
<td>3.11</td>
<td>4.99</td>
<td>4.19</td>
<td>4.76</td>
<td>7.11</td>
<td>6.64</td>
<td>9.22</td>
<td>8.46</td>
<td>15.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Provinces</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>Percentage of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Male workers</td>
<td>-</td>
<td>55.04</td>
<td>-</td>
<td>54.26</td>
<td>54.66</td>
<td>-</td>
<td>54.58</td>
<td>-</td>
<td>55.46</td>
</tr>
<tr>
<td>-High-school graduates</td>
<td>-</td>
<td>15.12</td>
<td>-</td>
<td>12.01</td>
<td>14.60</td>
<td>-</td>
<td>14.02</td>
<td>-</td>
<td>16.24</td>
</tr>
<tr>
<td>-College graduates</td>
<td>-</td>
<td>2.92</td>
<td>-</td>
<td>3.23</td>
<td>5.11</td>
<td>-</td>
<td>7.68</td>
<td>-</td>
<td>12.52</td>
</tr>
</tbody>
</table>

The CGSS survey data sets focus on the post-policy change years. It only includes urban households in 2003 survey, but both urban and rural households are included in the survey waves afterwards. Around 6,000-10,000 households are included each survey year. These surveys contain detailed information on individual education history, annual income and working experience as well as family background.

From these data sets, I create two samples: (1) a wage sample including weekly wages of
workers by demographic groups and (2) a working hours count sample that we use to measure the amount of labor supplied by each group. We divide the sample into different labor groups, distinguished by sex, education (less than high-school, high-school diploma, college drop-outs, college/university diploma, and graduate school), and 8 age-groups (21-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56-60). And we construct a separate sequence of college application and enrolment ratios of rural versus all high-school graduates in each of the college entrance examination years from the precise information on individual education history from the CGSS survey data sets.

Both samples include all individuals aged 21-60 who worked at least 20 hours per week in the preceding year. The wage measure that we use throughout the paper is the average weekly wage, computed as total annual or monthly labor earnings divided by total weeks worked. To get the real wage, we adjust the nominal wage using last year’s CPI (equal to 100 in 1988). We compute total working hours for each demographic group in each year by calculating the product of total annual hours (weeks worked times usual weekly hours) and the individual sample weight. Since the CHNS doesn’t have the corresponding individual weight, we construct the sample weight by population census data in the year 1990, 1995, 2000, 2005 and 2010.

2.2 Institutional Background

China’s Higher Education  China has the largest higher education system in the world. According to the Ministry of Education, there are 2,790 public higher education institutions enrolling 29.7 million students, including 1,867 regular higher education institutions with a total enrolment of 23.9 million, and 348 institutions for adults with a total enrolment of 5.8 million; and 707 other private institutions with a total enrolment of 5.3 million. With a centralized educational system, all the higher education institutions are under the control of the Ministry of Education through enrolment planning, funding, and evaluation.

Since 1977, The National College Entrance Examination (NCEE) is a prerequisite for entrance into almost all higher education institutions at the undergraduate level, and the unique criteria for college admission. The examination was uniformly designed by the Ministry of Education

---

8Most recent publicly available data in 2012 at http://www.moe.edu.cn

9Starting 2003, some top pilot schools were given autonomy of admission quota of no more than 5% of their total enrolments.
until the early 2000s\textsuperscript{10}. Exams are held annually, and generally taken by high school students in the end of their third year, although there has been no age restriction since 2001.

The initial college admission rate\textsuperscript{11} was astonishingly low, at only 4.7% in 1977. Since the open up in 1978, China’s economy has been growing at a stellar rate, and so does its physical capital accumulation. The high gross saving rate, together with the complementarity between the physical capital and the high-skill workers increase the return to college education\textsuperscript{12}. Up to early 1990s, the admission rate to college education maintained at a relatively constant growth of around 8.5% annually.

However, during the Asian financial crisis in early 1999, in order to alleviate the unemployment, stimulate the home consumption and to meet the tremendous demand for higher education, the State Council decided to embark on a radical enrolment expansion plan\textsuperscript{13}. It only took 4 months for the policy to be announced and taken into effect, within just one year, the college admission rate was increased by more than 40%, from 33.8% in 1998 to 55.6% in 1999, as shown in the Figure 2. The admission rate was kept around 60% since then. The initial target of increasing the gross enrolment ratio in tertiary education to 15% by 2010 was already surpassed in 2005, the actual gross enrolment rate increasing from less than 10% in 1998 to over 20% in 2005. After 2006, the government began to control the rapid growth of tertiary education, and the growth of college admission rate transformed to a steady and moderate trend.

The substantial increase in the college entry class would require a wider base of support for higher education. However, Between 1992 and 2003, the proportion of government expenditures in total education expenditures decreased from 84% to 62\textsuperscript{14}. As a result, the burden of financing for tertiary education shifted from entirely government to the point where household has to finance a substantially increasing amount through tuition and fees. Indeed, the expenditure on education ranked the first in total household expenditure in the 10th 5-year-plan\textsuperscript{15}.

\textsuperscript{10}The Ministry of Education allowed the College Enrolment Office of Shanghai and Guangdong to employ an independent exam in 1985. Starting from 2003, Beijing, Tianjin, Jiangsu, and Zhejiang were allowed to adopt independent propositions. Till now, there have been 16 provinces and municipalities adopting customized exams.

\textsuperscript{11}Calculated as the ratio of college admitted population by the population taken the NCEE in the same year according to the Ministry of Education

\textsuperscript{12}Wang (2011) and Heckman and Yi (2012)

\textsuperscript{13}Thirty years of evolution and development in education: http://theory.people.com.cn/GB/49157/49166/8143089.html

\textsuperscript{14}China Statistical Yearbook 2005

\textsuperscript{15}According to the China Youth and Child Research Center (CYCRC) 2007, Report on the Development
Private Financing and Tuition Fees  Higher education was free of charge and heavily subsidized by the government until 1989, when the dual-track system of tuition was set up, wherein students who scored below a cut-off line on the NCEE could attend colleges at a higher tuition level. The tuition fee was only 200 RMB per student in 1989, it grew to 610 RMB in 1993, when the State Council stated that higher education is non-compulsory, and students should pay tuition in principle. The tuition system was unified in 1997, and tuition and fees vary by institutions, locations and programs. As shown in the Table 2, the average per capita tuition and fees grows at at an annual rate of 11.8% from 1996 to 2011, peaking at 7850 RMB in 2011.
annual growth around 30% from 1996 to 1999.

Reliance on household wealth to finance the students’ tertiary education, the lack of well-established credit markets, and substantial increases in tuition and fees might restrict the opportunities for students from poor families to attain college education. Studies on China’s household wealth universally report the huge and increasing income and wealth gap ever since the economic reform in 1978, between rural and urban households. The Figure 3 pictures the tuition burden by household residential status, calculated as the ratio of average per capita tuition and fees over the annual per capita disposable household income\textsuperscript{17}. The tuition and fees are well over the per capita disposable income of rural household since 1999, while for urban household, it is around half of the per capita income. Thus China’s geographic unbalance and segregation would potentially affect the enrolment decisions through financial constraints faced by different residential groups.

3 Baseline Model

In this section, I consider a heterogeneous-agent OLG model. The economy is populated with a continuum of individuals, each live for finite periods, with age indexed by $j = 1, \ldots, T$. Of the China Youth During the 10th and 11th 5-Year-Plan

\textsuperscript{16} According to the Guidelines for China’s Education Reform and Development in 1993

\textsuperscript{17} Based on the tuition and income statistics from NBS, various years
Individuals start with high school graduation \((j = 1)\), survive from age \(j\) to \(j + 1\) with probability \(\zeta^j\), and die after period \(T\) for sure. Time is discrete, at each period a new cohort of measure one enters the economy. Since cohort size and survival probabilities are time-invariant, the model age distribution is stationary.

We assume a risk-free asset market for simplicity, consumers could borrow at the risk-free interest rate up to a borrowing constraint \(a\). Workers in different education and age groups are imperfect substitutes in production.

### 3.1 Timing

At birth, individuals observe their initial wealth and ability level, and choose to attend college or not. If they choose to go to college, they will spend the first 4 periods in college, living on their initial wealth and a very small amount of stipends endowed by the government. If instead they choose non-college, they work from period 1 and will earn the high-school wage.

Life time is composed of working and retirement periods, \(T = T_W + T_R\). The working periods are 1 through \(T_W\) for high school workers, and 5 through \(T_W\) for college labor. In period \(T_W + 1\) through \(T\), individuals earn retirement pension that is a function of their effective working wage.

### 3.2 Consumer’s Problem

**Value Function** Instead of working, college students get an annual stipend \(y_{col}\) in the first 4 periods: \(j = 1, \ldots, 4\), after graduation, college workers work in periods \(j = 5, \ldots, T_w\). High school graduates work through the whole working periods.

A worker’s labor income is composed by a wage level on efficient units of work by the corresponding age and education group, and her efficient units of labor supplied, as an exponential function of her ability, age and education level. Consider a worker at age \(j\), ability level \(b\) and education \(e \in \{C, H\}\), denoting high school and college education respectively. Age efficiency schedule \(N(j, b, e)\) is assumed to follow an exponential function of a second-degree polynomial
with coefficients depending on worker’s ability and education.

\[ V(j, e, b, a) = \max_{c, a'} \{ U(c) + \beta \zeta^j V(j + 1, e, b, a') \} \]

\[ c + \zeta^j a' = (1 + r)a + y \]

\[ y = \begin{cases} 
  w_j^e N(j, b, e)(1 - \tau) & \text{if working} \\
  y_{col} & \text{if } j=1, \ldots, 4 \text{ and } e=C 
\end{cases} \]

\[ \zeta^j a' \geq a \]

Consumers are subject to borrowing constraint \( a \) each period, and making their consumption as well as saving choice for next period \( a' \). One unit of savings delivers \( 1/\zeta^j \) units of assets next period, reflecting the annuity-market survivors’ premium.

After retire at the end of period \( T_w \), individuals earn retirement pension \( \phi(e, b) \) as a function of their effective working wage determined by education and ability levels, for periods \( j = T_w + 1, \ldots, T = T_w + T_R \)

\[ V(j, e, b, a) = \max_{c, a'} \{ U(c) + \beta \zeta^j V(j + 1, e, b, a') \} \]

\[ c + \zeta^j a' = (1 + r)a + \phi(e, b) \]

\[ a' \geq a \]

**Education Choice** In the beginning of their life-time, individuals make college enrolment decisions with full knowledge of their initial ability and asset endowments. The government control the total college admission by choosing the ability threshold \( B \), where in equilibrium the desired college admission \( E^C_t \) satisfies:

\[ E^C_t = \int \int_{B^t} \eta(b, a) dF(b) dF(a) \]  

(1)
With both types of workers’ value function defined above, individual’s education choice is belowing:

\[
V(b,a) = \begin{cases} 
\max\{V(1, C, b, a), V(1, H, s, a)\} & \text{if } b \geq B_t \\
V(1, H, b, a) & \text{if } b < B_t
\end{cases}
\]

(2)

\[
\eta(b, a) = \begin{cases} 
1 & \text{if } V(b, a) = V(1, C, b, a) \\
0 & \text{if } V(b, a) = V(1, H, b, a)
\end{cases}
\]

(3)

The first order condition can be written as, if \(a\) is not binding:

\[
u'(c) = \beta(1 + r)u'(c')
\]

(4)

3.3 Firm’s Problem

There is a large number of competitive firms, with the following production technology, where \(H_t\) and \(C_t\) denote the aggregate high-school and college educated labor.

\[
F(H_t, C_t, K_t) = AK_t^{1-\alpha}(\theta_cC_t^\rho + \theta_hH_t^\rho)^{\frac{\alpha}{\rho}}
\]

I assume imperfect substitution between different age and education groups, following the assumption in Card and Lemieux (2001). \(-\infty < \rho \leq 1\) is a function of elasticity of substitution \(\sigma_E\) between the two education groups, where \(\rho = 1 - 1/\sigma_E\). Similarly, \(-\infty < \eta \leq 1\) is a function of the partial elasticity of substitution \(\sigma_A\) between different age groups and same education level, \(\eta = 1 - 1/\sigma_A\). \(\alpha_j\), \(\beta_j\) and \(\{\theta_c, \theta_h\}\) sequences are the relative effective parameters between age and education groups. When \(\alpha_j = 1\), \(\beta_j = 1\) and \(\eta = 1\), labor inputs are perfect substitutable with same education and different age.

\[
H_t = \left(\sum_{j=1}^{T_w} \alpha_j H_{jt}^\eta\right)^{\frac{1}{\eta}}, \quad C_t = \left(\sum_{j=1}^{T_w} \beta_j C_{jt}^\eta\right)^{\frac{1}{\eta}}
\]

In each period, the representative firm takes both wages and interest rate as given, and choose its optimal college and high-school labor \(c_j, h_j\) and investment demand \(i\):

\[
\Omega(k) = \max_{c_j, h_j, i} F(H, C, k) - \sum_j w_j^c c_j - \sum_j w_j^h h_j - i + \frac{1}{1+r}\Omega(k')
\]

(5)
subject to: \( k' = (1 - \delta)k + i - \frac{\varepsilon}{2} (\frac{i}{k} - \delta)^2 k \), \( \delta \) denotes the steady-state depreciation rate, and \( \varepsilon \) is the investment cost parameter.

The first order condition is:

\[
\frac{1 + r_t}{1 - \varepsilon (}\frac{i}{k_t} - \delta) = \alpha A k_t^{\alpha - 1} \left( \theta_{ct+1} C_{t+1} + \theta_{ht+1} H_{t+1}^\rho \right)^{\frac{1-\alpha}{\rho}} + \frac{1 - \delta + \frac{(i_{t+1}}{k_{t+1}} - \delta^2)}{1 - \varepsilon (}\frac{i}{k_{t+1}} - \delta)
\]

\[
w^C_{jt} = \alpha \beta_j \theta_{ct} Ak^{1-\alpha} C_t^{\rho - \eta} (\theta_{ct} C_t + \theta_{ht} H_t^{\rho})^{\frac{\alpha - 1}{\rho}} C_{jt}^{\eta - 1}
\]

\[
w^H_{jt} = \alpha \alpha_j \theta_{ht} Ak^{1-\alpha} H_t^{\rho - \eta} (\theta_{ct} C_t + \theta_{ht} H_t^{\rho})^{\frac{\alpha - 1}{\rho}} H_{jt}^{\eta - 1}
\]

In the steady state, the first equation reduces to:

\[ r + \delta = \alpha A k^{\alpha - 1} \left( \theta_{ct} C_t + \theta_{ht} H_t^{\rho} \right)^{\frac{1-\alpha}{\rho}} \]

### 3.4 Equilibrium Definition

A stationary equilibrium in this economy is a set of value functions and decision rules for the households: \( \{c_t(j, e, b, a), a_{t+1}(j, e, b, a), \eta(j, e, b, a), V_t(j, e, b, a)\} \), firms’ optimal decision \( C_t, H_t, K_t, \) the capital rental rate \( r, \) and the wage rate for college and non-college labor, \( w^C_{jt}, w^H_{jt}, \) such that

1. Given the labor income tax \( \tau, \) and initial endowments, individuals’ decision rules and value functions solve problems (2).

2. Firms’ decisions solve the corresponding problem (5).

3. Asset market clears

\[ K = \sum_{j=1}^{T} \mu_j \int a(j, e, b, a^{-1})dF(b, s, a^{-1}), \tag{6} \]

4. College labor market for each cohort clears, \( j = 1, \ldots, T_w: \)

\[ C_j = \mu_j \int \eta(b, a)dF(b, a) \tag{7} \]
5. High-school labor market for each cohort clears, \( j = 1, \ldots, T_w \):

\[
H_j = \mu_j \int (1 - \eta(b, a))dF(b, a)
\]  

(8)

6. The government chooses college entry threshold \( B^t \) and the budget constraint satisfies

\[
\sum_{e=C,H} \sum_{j=T_w+1}^{T} \mu_j \int f^r w^e_{T_w} \phi(e, b) dF(b) + \sum_{j=1}^{4} \mu_j \int y_{col} dF(b) = \tau \mu_j \sum_{j=5}^{T_w} w^C_{jt} N(j, b, C) \eta(b, a) dF(b, a) + \tau \mu_j \sum_{j=1}^{T_w} w^H_{jt} N(j, b, H)(1 - \eta(b, a)) dF(b, a)
\]  

(9)

3.5 Model Mechanism

In this section we will discuss the general equilibrium effects of an expansion in the college entry class. In the initial steady state, personal college enrolment decision depends on the comparison of life-time value out of college and high school education. As shown in Figure 4, individuals with little initial assets would prefer go to work early to consume more early on, and wealthy agents would prefer to go to college and earn a positive college wage premium in their life-time. There is also a positive correlation between individual ability level and the probability to go to college, since more able workers could generally reap more benefits from the college education, as shown by each of the panels in Figure 5.
Figure 4: Value Functions over Initial Asset Levels for College and High School Education

Figure 5: Value Functions over Ability Levels

Value Function for High Initial Wealth  Value Function for Low Initial Wealth

We could further graph the education choice outcome of an individual characterized by both ability and initial wealth level as in Figure 6. The college admission was determined by both individual preferences and the college entrance ability threshold set by the government.
As college expands by a substantial magnitude, the increased college labor supply will drive down the college wage premium where less workers would choose to enrol. Thus, for the same amount of assets, individuals need to be more endowed with more ability to be willing to go to college, and for the same level of ability, they should be wealthier to attend. The indifference curve of college education choice is driven to the right in the mid-panel of Figure 6.

4 Descriptive Analysis

This section presents (1) the time series of college wage premiums and the effective college over high-school labor supplies, for different age and residential status groups, (2) strong negative correlations between the change in college wage premiums and the labor supply in the periods immediately after the college expansion in 1999, and (3) a second-stage estimation of elasticities of substitution between age and education groups.

4.1 College Wage Premiums

College Wage Premiums by Age Groups  Table 3 presents the estimated college wage premiums for five-year age groups, taken at each survey year in CHNS from 1993 onwards. The college wage premiums are estimated in separate regression models for each age group in each survey year, using samples of full-time workers with exactly a high school or exactly a college degree. Each regression includes a dummy for college graduates, a linear age term, and dummies for other demographic properties, gender and residential status.

An important feature of the college wage premium in Table 3 is that they are based on differences in earnings between individuals of the same age with a college degree or a high school diploma. This measure compares individuals who attended elementary and secondary schooling together, and faced the identical scenarios when they made decisions to attend college or not. Depending on our interest in explaining the effects of college expansion policy on college entrance choices and the systematic age effects in our econometric models, this measure will outweigh the potential disadvantage of ignoring different labor market experience in the same age groups who have different level of schooling. The entries in Table 3 provide a variety of information on the evolution of college-high school wage differentials Comparisons across the rows of Table 3 reveal the age profile of the college-high school wage differentials in the particular survey year. Across the years, the age profiles generally show a hump-shape with
Table 3: College-High School Wage Differentials by Age and Year

<table>
<thead>
<tr>
<th>Age Range</th>
<th>21-25</th>
<th>26-30</th>
<th>31-35</th>
<th>36-40</th>
<th>41-45</th>
<th>46-50</th>
<th>51-55</th>
<th>56-60</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>-0.149</td>
<td>-0.347</td>
<td>-0.004</td>
<td>-0.108</td>
<td>0.149</td>
<td>-0.244</td>
<td>0.674</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.028)</td>
<td>(0.028)</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.040)</td>
<td>(0.039)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>1996</td>
<td>-0.096</td>
<td>-0.059</td>
<td>0.186</td>
<td>0.089</td>
<td>0.281</td>
<td>0.292</td>
<td>-0.096</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.011)</td>
<td>(0.027)</td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>1999</td>
<td>0.158</td>
<td>0.096</td>
<td>-0.066</td>
<td>0.110</td>
<td>0.146</td>
<td>0.138</td>
<td>0.089</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.032)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>2003</td>
<td>0.215</td>
<td>0.172</td>
<td>0.309</td>
<td>0.156</td>
<td>0.250</td>
<td>0.388</td>
<td>0.194</td>
<td>0.315</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.013)</td>
<td>(0.017)</td>
<td>(0.013)</td>
<td>(0.009)</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>2005</td>
<td>0.197</td>
<td>0.295</td>
<td>0.315</td>
<td>0.224</td>
<td>0.493</td>
<td>0.489</td>
<td>0.359</td>
<td>0.324</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.021)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>2008</td>
<td>0.075</td>
<td>0.208</td>
<td>0.263</td>
<td>0.161</td>
<td>0.240</td>
<td>0.531</td>
<td>0.570</td>
<td>0.441</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.012)</td>
<td>(0.017)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>2010</td>
<td>0.047</td>
<td>0.272</td>
<td>0.267</td>
<td>0.296</td>
<td>0.411</td>
<td>0.383</td>
<td>0.395</td>
<td>0.487</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.008)</td>
<td>(0.010)</td>
</tr>
</tbody>
</table>

the maximum obtained around age 41-50. And comparing down a column of the table shows the changing college premium for a particular age group.

Panel A of Figure 7 graphs the time series of college wage premium across three representative age groups. Among the youngest group, 21-25 years olds, the college wage premium increases until 2003, when the first cohort of college students who is impacted by the college expansion policy in 1999 graduates, and immediately decreases from that year onwards. For older groups, their college wage premium generally rose until later than 2008 then slowly levelled off. The college wage premium trends across different age groups in Figure 7 shows that the college wage premiums for specific age groups can rise or fall independently of other groups, suggesting the potential importance of the age effects on the college wage premium.
College Wage Premium by Household Registration Status  Using the detailed household registration information from CHNS\textsuperscript{18}, Table 4 documents the college premium by individual’s household residential status across survey years. Rural workers have lower college wage premiums until 1996, it grew sharply and overtook the college wage premium of the corresponding urban workers from 1996 to 1999 when tuition fees increased. After 2003, the college wage premium generally levelled off. While for urban workers, the college wage premiums stay below the corresponding rural premiums since 1999, and it began to decrease after 2005, as shown in the panel A of Figure 8.

**Table 4: College-High School Wage Differentials by Residential Status**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>0.092</td>
<td>0.049</td>
<td>-0.229</td>
<td>0.054</td>
<td>0.189</td>
<td>0.554</td>
<td>0.503</td>
<td>0.523</td>
<td>0.478</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.017)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.008)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Urban</td>
<td>0.038</td>
<td>0.022</td>
<td>0.105</td>
<td>0.100</td>
<td>0.165</td>
<td>0.247</td>
<td>0.377</td>
<td>0.315</td>
<td>0.354</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

4.2 Relative Supply

We turn next to an overview of the relative supplies of college-educated labor by age and residential status. Following Katz and Murphy (1992), I measure efficient supply of labor by each demographic group as weighted total hours worked in that group, where the weighting being the average relative wage across all the survey years. I then take the log ratio of the effective college over high school hours as the relative supply of college versus high-school labor.

Panel B of Figure 7 shows the evolution of the relative college labor in three representative age groups: 21-25 years olds, 26-30 years olds and 46-50 years olds. For the 21-24 years-old group, relative college labor supplies grow most sharply from 1999 to 2003, while for older groups, the relative supplies trended upwards fairly steadily after 1999.
The relative college labor supply by household registration status is shown in Figure 8. The general college labor supply is higher across the years for urban hunk-holders. To compare the growth rate of relative college labor supply, I use the y-axis on the left to show the relative college labor supply for rural workers, the corresponding urban labor supply is presented using the y-axis on the right. The relative college labor supply grows at around the same rate for both residential groups from 1996 to 2003, when the college tuition grows fast. After 2003, the rural college labor grows strictly faster.

Using the detailed individual education history from CGSS surveys, we construct the rural student’s ratio over all the college applicants (those who took college entrance examination in their senior years of high school), and the rural students’ proportion over all the actual college enrolment population, by the years when they took the National College Entrance Exams, as in Figure 9. For both the application and enrolment trends, the relative ratio of rural students didn’t grow until after 1999, this confirms the change of relative college labor supply among different residential groups. And one possible reason might be that rural residents are restricted by their available family resources to finance the substantial increase in college tuition and fees. From 1997 to 1999, the rural students’ college enrolment ratio dropped even more than their corresponding application ratio.

4.3 Elasticity Estimation

Figure 10 graphs the times series of college wage premium and the corresponding relative labor supply. The overall series for the college labor supply is upward sloping, the supply grows faster than average after 1996, especially after 2003 when the first cohort of college labor graduates out of the college expansion in 1999. In the meantime, the college wage premium began to fall after 2005, it appears that fluctuations in supply growth has the potential to explain observed college wage premium changes, especially after 2005. To further examine the negative correlation between the college wage premiums and the relative college labor supplies, we present the changes in these two variables from 2003 to 2008, in 64 distinct age, gender and education groups in Figure 11.

Thus, using the first order conditions from the firm’s problem, as in Card and Lemieux (2001),

\[ \text{When individuals surveyed multiple times throughout the survey series, the earliest available household registration status information is taken.} \]
and applying a linear time trend to proxy the technology shock ($\log(\theta_{ct}/\theta_{ht})$) and the aggregate supply effect ($((\frac{1}{\sigma_E} - \frac{1}{\sigma_A}) \log(\frac{C_t}{H_t}))$, we get the first stage estimates (REG1) for the elasticity for age-group specific college labor supply. The estimated year effect, as shown in the first column of Table 5 is positive and significant.

$$\log \left( \frac{w^C_{jt}}{w^H_{jt}} \right) = \log \left( \frac{\beta_j}{\alpha_j} \right) + \log \left( \frac{\theta_{ct}}{\theta_{ht}} \right) - \left( \frac{1}{\sigma_E} - \frac{1}{\sigma_A} \right) \log \left( \frac{C_t}{H_t} \right) - \frac{1}{\sigma_A} \log \left( \frac{C_{jt}}{H_{jt}} \right) + e_{jt}$$

**Table 5: Estimated Models for the College-High School Wage Gap**

<table>
<thead>
<tr>
<th>Fit Statistics</th>
<th>REG1</th>
<th>REG2</th>
<th>REG3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.745</td>
<td>0.746</td>
<td>0.748</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**Table 6: Elasticity Estimates Comparison**

<table>
<thead>
<tr>
<th>Elasticities</th>
<th>China(CHNS,CGSS)</th>
<th>U.S.(Katz &amp; Murphy)</th>
<th>U.S.(Card &amp; Lemieux)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-group ($\sigma_A$)</td>
<td>4.4-7.6</td>
<td>4-6</td>
<td></td>
</tr>
<tr>
<td>Col-High ($\sigma_E$)</td>
<td>1.1-1.9</td>
<td>1.4</td>
<td>1.1-1.6</td>
</tr>
</tbody>
</table>
The second and third column of Table 5 presents the estimates of the second-stage models that include both age-group specific relative college labor supplies and the aggregate labor supply. Where the relative productivity efficiency effect ($\log(\beta_j/\alpha_j)$) is estimated on the first stage age-group elasticity, and the aggregate supplies of college and high school labor time series is constructed assuming perfect (as in REG2) and imperfect (as in REG3) substitution across age groups with same education. The estimated elasticity of substitution between college and high school labor is in the range of 1.1 to 1.9, comparable to the estimates in Katz and Murphy (1992) and Card and Lemieux (2001)\textsuperscript{19}. And my estimation of elasticity of substitution between different age groups are in the range of 4.4 to 7.6, generally higher than the estimates range of 4 to 6 from Card and Lemieux (2001). The estimates of age-group elasticities from the second-stage are close to the first-stage results, and the year effects across different specification show a steeply rising returns.

5 Calibration

Given the elasticities estimation in Section 4, I calibrate the parameters by choosing an initial steady state which is consistent with the pre-expansion policy data. Assuming that after the college expansion, only the ability threshold for college entry changes, I present the implications of the calibrated model for college wage premiums and relative college labor supply for different subgroups, both in comparative statics and the simulated transition path.

5.1 Parameter Choices

**Exogenously Determined Parameters** A model period corresponds to one year of calendar time. Individuals enter the economy at age 21 and retire at age 55 (working periods $T_w = 35$). Retirement lasts for 15 years, and everyone dies at age 70. The net interest rate, $r$, is set equal to 2%. Since there is no leisure decision involved, I use the conventional power utility specification of preferences, and the risk aversion for utility function is set equal to 2.0. The labor share of the Cobb-Douglas production technology is set at 0.50, broadly consistent with the existing empirical evidence\textsuperscript{20}. Table 7 shows all the exogenously determined parameters.

\textsuperscript{19}I compare my estimation results with the elasticity estimates from United States of Katz and Murphy (1992) and Card and Lemieux (2001) in Table 6

\textsuperscript{20}As in BAI and QIAN (2010) and Chi and Qian (2013)
TABLE 7: Exogenously Determined Parameters of the Baseline Economy

<table>
<thead>
<tr>
<th>Description</th>
<th>Parameter</th>
<th>Value</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk aversion</td>
<td>(\sigma)</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>Interest rate</td>
<td>(r)</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td>(\beta)</td>
<td>(1/(1 + r))</td>
<td></td>
</tr>
<tr>
<td>Labor Share</td>
<td>(\alpha)</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>Survival Rate from (j) to (j+1)</td>
<td>(\zeta)</td>
<td>—</td>
<td>WHO(2000)</td>
</tr>
<tr>
<td>Els of substitution btw age groups</td>
<td>(\sigma_A)</td>
<td>7.60</td>
<td>Estimation from data</td>
</tr>
<tr>
<td>Els of substitution btw edu groups</td>
<td>(\sigma_E)</td>
<td>1.40</td>
<td>Estimation from data</td>
</tr>
<tr>
<td>Working periods</td>
<td>(T_w)</td>
<td>35</td>
<td>years: 21-55</td>
</tr>
<tr>
<td>Retirement periods</td>
<td>(T_r)</td>
<td>15</td>
<td>years: 56-70</td>
</tr>
</tbody>
</table>

**Distributions: Ability, Initial Wealth for Rural and Urban Groups**  
Agents enter the economy with perfect knowledge of two individual-specific attributes: ability and initial assets endowment. Accounting for the sizeable gap between the rural and urban household, not only in income and wealth level, but also in the education resources\(^{21}\), I assume different distribution parameters for rural and urban groups.

The initial asset is assumed to follow log-normal distributions for both groups, and the ability is assumed to be uniformly distributed. Thus the distribution of ability and initial assets yield six parameters to be calibrated: (i) the standard deviation of ability for both groups, (ii) the mean ability level for rural groups, and (iii) the mean and standard deviation of initial wealth level for both groups (the average rural ability and rural wealth level are simply set to a computationally convenient level).

**Data Targets**  
I use the wage series from micro-econometric evidence of the panel data to pin down the ability distributions for both age groups. For initial wealth distribution for rural and urban individuals, I use the target of aggregate household disposable income as well as the corresponding Gini coefficients\(^{22}\). The detailed list of steady state targets are listed in

\(^{21}\)As documented by Li, Whalley, Zhang, and Zha (2008) and Willmann and Schucher (2005), there are huge gaps in educational funding, teacher qualifications, and school conditions between rural and urban schools

\(^{22}\)Income statistics from NBS, various years
Table 8.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Target Data</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (Initial)</td>
<td>A_1</td>
<td>College admission rate before the expansion</td>
<td>0.60</td>
</tr>
<tr>
<td>SD of ability (Rural)</td>
<td>σ^R</td>
<td>SD of log wage for rural workers</td>
<td>0.29</td>
</tr>
<tr>
<td>Average Ability (Rural)</td>
<td>E[b]^R</td>
<td>Mean log wage for rural workers</td>
<td>0.5</td>
</tr>
<tr>
<td>SD of ability (Urban)</td>
<td>σ^U</td>
<td>SD of log wage for urban workers</td>
<td>0.29</td>
</tr>
<tr>
<td>Average Ability (Urban)</td>
<td>E[b]^U</td>
<td>Mean log wage for urban workers</td>
<td>0.65</td>
</tr>
<tr>
<td>SD of initial asset distr. (Rural)</td>
<td>σ^R_a</td>
<td>Rural Gini-Coefficient</td>
<td>0.7</td>
</tr>
<tr>
<td>Mean of initial asset distr. (Rural)</td>
<td>E[a]^R</td>
<td>Rural Household Disposable Income</td>
<td>1.0</td>
</tr>
<tr>
<td>SD of initial asset distr. (Urban)</td>
<td>σ^U_a</td>
<td>Urban Gini-Coefficient</td>
<td>2.0</td>
</tr>
<tr>
<td>Mean of initial asset distr. (Urban)</td>
<td>E[a]^U</td>
<td>Urban Household Disposable Income</td>
<td>2.65</td>
</tr>
</tbody>
</table>

Age Efficiency Schedule and the Pension Scheme  
To make sure labor income sequences generated by the model be consistent with the empirical evidence on the dynamics of wages found in the panel data, we need to select the functional forms for the age efficiency schedules and the retired pension scheme. Specifically, I assume the age efficiency follows an exponential function of a quadratic polynomial on individual age $f(j)$, and an individual component that follows a linear function of age, with the coefficient determined by individual ability and intercept by education level$^{23}$.

$$
N(j, b, e) = \exp \left[ f(j) + bj + g(e) \right]
$$

We model the pension system in China as a defined benefits plan, as in Song, Storesletten, Wang, and Zilibotti (2012), the replacement ratio is assumed to be 60%. And this is also general in line with the retirement labor income in the survey data, so we approximate the pension scheme to be 60% of the average life-time labor income. The subsidy for college education $y_{col}$ is assumed to be 30% of the current average wage, which is closely assembles the average living expenses for a college students.

$^{23}$As in Guvenen, Kuruscu, and Ozkan (2009), the regression residual of a raw wage on a polynomial in age is assumed to follow a linear trend in age, with coefficients and intercepts determined by a pair of perfectly correlated learning abilities.
5.2 Quantitative Results

In this section, I begin by presenting the comparison of two steady states, before and after the college expansion policies. I then proceed to give out the simulation results on the transitional path of the key variables of interest.

The Long-term Impacts of a College Expansion: Steady State Comparison  As shown in Table 9, the as ability threshold lowered by the government, college wage premiums are decreased for both groups by similar magnitudes, where the college labor supply increased much more in rural than urban groups. This is also shown by an 35% increase in the admission rate for rural students, where only less than 20% increase for the corresponding urban groups in the long run. It seems in the long run, the ability effects dominate. Think of rural students are more concentrated around the a relatively lower level of ability, with less initial wealth. As ability threshold decreases, more rural students could access the college education with an ability level below the initial threshold and above the later one. On the other hand, urban students’ ability are generally higher, so they are not affected as much by the expansion.

Table 9: Steady State Comparison

<table>
<thead>
<tr>
<th></th>
<th>Average College Wage Premium</th>
<th>Log Relative College Labor Supply</th>
<th>Admission Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline(Rural, Before)</td>
<td>1.09</td>
<td>-1.19</td>
<td>0.21</td>
</tr>
<tr>
<td>Baseline(Urban, Before)</td>
<td>1.10</td>
<td>0.14</td>
<td>0.50</td>
</tr>
<tr>
<td>Baseline(Rural, After)</td>
<td>0.28</td>
<td>-0.33</td>
<td>0.47</td>
</tr>
<tr>
<td>Baseline(Urban, After)</td>
<td>0.29</td>
<td>0.65</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Transitional Dynamics  Figure 12 displays the transition path of the college enrolment rates for urban and rural students in Panel A. There is a clear overshooting trend that the initial admission rates increased by over 40%, then gradually decrease over time as more college labor supplied and college wage premium driven down. Panel B shows the relative college labor supplied, it doesn’t increase much in the first four periods before the first cohorts of students

---

24 This is in line with our estimation of rural college enrolment ratio using CGSS data in Figure 9 and as documented by Li, Whalley, Zhang, and Zha (2008).
affected by the expansion policy graduates, then it also overshoots to a higher level than
the new steady state college labor supply, and gradually decreases afterwards. Rural college
labor increase much more than the corresponding urban one. Panel C presents the percentage
decrease in college wage premiums, college wage premium decreases most four years after the
expansion, and increases gradually after that. The general patter of transitional dynamics are
in line with our empirical trends in Section 4.

6 Counterfactual Experiments

In Nov 1998, the chief economist Dr. Min Tang, at the Beijing representative office of the
Asian development bank, wrote a proposal to the Premier Minister Rongji Zhu on expanding
the college admission by one fold in three years25. The initial proposal was to increase the
college entry class by 25% annually until it expand the college admission population by one
fold in three to four years. The suggestion was quickly taken into consideration by the Ministry
of Education, and the initial target in early 1999 was set to increase the college entry class
by 21% from the previous year. However, in June of 1999, the official policy by the State
Council increased the college admission quota by 47.4%26, which largely exceeded the initial
expansion plan.

In this section, I conduct counterfactual analysis using the calibrated model in Section 5. The
policy experiment is on an alternative conservative expansion policy, closely resembles the
initial proposal by Dr. Min Tang, and the plan of the Ministry of Education in the early
1999. Comparisons are made between the actual policy expansion and the alternative policy
experiments with welfare analysis on different residential groups.

Mild College Expansion Policy  In view of the recent decreasing in the college wage
premium, and especially in the young college graduates, as shown in Figure 7, there are heated
debate on the possible oversupply of college graduates in China. In the first counterfactual
experiment, I examine the alternative effects on enrolment decisions and the corresponding
relative college labor supply and college wage premiums, for different residential groups. I
increase the college entry class by 20% each year for 3 consecutive years, which will generate
an increase of total enrolment population by only one fold, and I keep the admission rate

25Published in the Economic News on Feb. 19th, 1999 under title of Education to Promote Consumption,
constant after the expansion. Comparing to the actual college expansion starting from 1999 in China, which increase the total enrolment population by five folds in just 6 years, the policy experiment is a conservative and mild expansion policy.

Panel A of Table 10 presents the college enrollment results, with just a mild expansion, the overall admission rate is increased by only 30%, and although rural enrollment rate grows faster than the urban groups, the overall relative college labor supply increases by only 57 log points, comparing to an increase of 86 log point in the college expansion in 1999. As for the urban groups, the relative college labor supply increases by 9 log points less than the expansion in 1999. Thus a mild college expansion would generate less college labor supply, especially for the rural groups. A long-term welfare comparison from the Panel C shows that the overall welfare will increase by less than the actual college expansion, and this is specially phenomenal for the rural group, with a slightly increase of 6.87% comparing to an original 22.70% of increase in their welfare. Since the overall growth in college labor supply in urban groups is less than rural students, a mild college expansion will increase their welfare by curbing the decrease of the college wage premium.
Table 10: Experiment I: Mild College Expansion

Panel A: Admission Rates

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Urban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Expansion</td>
<td>0.33</td>
<td>0.50</td>
<td>0.21</td>
</tr>
<tr>
<td>Base Expansion</td>
<td>0.56</td>
<td>0.70</td>
<td>0.47</td>
</tr>
<tr>
<td>Mild Expansion</td>
<td>0.43</td>
<td>0.61</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Panel B: Relative College Labor Supply

<table>
<thead>
<tr>
<th>Change (log points)</th>
<th>All</th>
<th>Urban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Expansion</td>
<td>66</td>
<td>51</td>
<td>86</td>
</tr>
<tr>
<td>Mild Expansion</td>
<td>46</td>
<td>45</td>
<td>57</td>
</tr>
</tbody>
</table>

Panel C: Long-term Welfare Comparison

<table>
<thead>
<tr>
<th>Change (%)</th>
<th>All</th>
<th>Urban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Expansion</td>
<td>19.10</td>
<td>12.08</td>
<td>22.70</td>
</tr>
<tr>
<td>Mild Expansion</td>
<td>11.46</td>
<td>13.81</td>
<td>6.87</td>
</tr>
</tbody>
</table>

7 Conclusion

In this paper, I developed an empirically grounded dynamic overlapping-generation general equilibrium model with endogenous college education choice. It provides a better understanding of the college labor market by jointly considering an exogenous policy threshold and an endogenous ability selection conditional on the initial wealth. I allow for heterogeneity in ability and initial wealth, and produce a model that is consistent with the main features of life-cycle wage growth and the discrepancies between rural and urban workers that are the central features of the recent China labor market.

My empirical analyses suggest that there is significant heterogeneity in workers’ response to the substantial college expansion in 1999. The urban students respond immediately to the college expansion with an overall higher initial asset level. However, as the college wage
premium driven down by the increasing supply of young college graduates, intention for college enrolment decreased. On the other hand, rural students are constrained by their average low level of household income, the college labor supply increased gradually, and the overall growth in rural college labor exceeds the urban college labor supply. Also, the college-high school wage gaps of different age groups have not moved together. In particular, the college premium for young graduates fall immediately after the first cohort affected by the college expansion policy graduated in 2003, while the premium for older workers didn’t level off until five to ten years later.

Following a model that incorporates imperfect substitution between different age and education groups, I found that the evolution of the college wage premiums accounts well for the patterns in data. The dramatic exogenous policy change in 1999 serves as a sharp identification strategy, to estimate the production demand elasticities on different categories of labor input. As a result, the estimated elasticities are generally comparable with the previous literature, with a slightly higher elasticities for different age groups.

Finally, I apply the calibrated model to a set of alternative government policies on college admission. This lays the foundation for the welfare analysis on different residential subgroups and generations. And the huge college expansion policy, as well as its effects on different subgroups, constitutes a good case study on acceleration in the supply of college educated workers across the developing countries\textsuperscript{27}.

\textsuperscript{27}Mehta and Shiela (2007) documented the age-profile of college wage premiums in India, Philippines and Thailand, and Oostendorp and Doan (2013) the case in Vietnam.
Figure 6: College Enrolment with Ability Thresholds
FIGURE 7: College Wage Premium and Relative Efficient College Labor Supply by Age Groups

Panel A: Estimated College Wage Premium

Panel B: Efficient College over High School Labor Supply
**Figure 8:** College Wage Premium and Relative Efficient College Labor by Residential Status

Panel A: Estimated College Wage Premium: Rural — Urban

Panel B: Efficient College over High School Labor Supply: Rural — Urban

**Figure 9:** Rural Students’ Share for College Enrolment and Application
Figure 10: College Wage Premium and Relative College Labor Supply

Figure 11: Price and Quantity Changes for 64 Groups, 2003-2008
Figure 12: Transition Path by residential groups

Panel A: College Enrolment Rate

Panel B: Effective College over High School Labor Supply

Panel C: College Wage Premium
References


