China's Higher Education Expansion and its Labor Market Consequences

Shi Li
Chunbing Xing

May 2010

DISCUSSION PAPER SERIES

I Z A

China's Higher Education Expansion and its Labor Market Consequences

Shi Li
*Beijing Normal University
and IZA*

Chunbing Xing
*Beijing Normal University
and IZA*

ABSTRACT

China's Higher Education Expansion and its Labor Market Consequences

Using a 1/5 random draw of the 1% census of 2005, we investigate how China's higher education expansion commenced in 1999 affects the education opportunities of various population groups and how this policy affects the labor market. Treating the expansion as an experiment and using a LATE framework, we find that higher education expansion increased the probability of go to college tremendously. Different populations "benefit" from this policy differently however. Minority female, those from central-western region and from rural areas are less likely to benefit from it. One-child families are more responsive to this policy. Using higher education resources at the provincial level as another dimension of variation, and using a difference-in-difference strategy, we find that the education expansion decreased the within sector inequality of population with above high school (inclusive) education. This is primarily due to the increase of the income level for high school graduate. That of the college graduate deceased, but only slightly and not significantly.

JEL Classification: I2, J2

Keywords: China, higher education expansion, LATE, difference in difference, income level

Corresponding author:

Chunbing Xing
Beijing Normal University
in 2010 on leave at:
University of Western Ontario
Room 4012, Social Science Centre
London, Ontario, N6A 5C2
Canada
E-mail: xingchunbing@gmail.com

1. Introduction

China's higher education has been experiencing transformation during the whole reform period, and is still transforming. Ever since the beginning of the reform and open-up policy, the scale of higher education in terms of new entering college students and college graduates kept increasing. But the increase in the 1980s and most of 1990s is not large, which is especially dwarfed by the major increase in 1999 and thereafter. In 1999, the central government of China made a decision to enlarge the scale of higher education. As a result, the number of new students admitted to college increased by over 40% in 1999 with comparison 1998. By 2005, the number of new college students over quadrupled (4.7) that of 1998. Given the unprecedented scale of the expansion, many people term this radical policy a *great leap forward* in higher education. The questions we want to answer in this paper are twofold. First, how this radical change affects the education opportunity of different individuals? Second, what's the labor market consequence of this expansion? The empirical strategies are also twofold: one is Before-After, the other is Difference-in-Difference. The data is a 1/5 random draw from the 1% census data in 2005.

We treat the expansion policy as a natural experiment. Due to the expansion, individuals taking college entrance exams before and after the expansion policy have different probabilities of being admitted to college. Meanwhile, the *before* and *after* group are largely determined based on their ages, which are exogenous. Therefore we separate all individuals into two groups according to their ages. The younger group took college examinations after 1999 (if they take at all), and the younger group took the exams before 1998.[1] We also realize that this policy didn't affect the whole population homogeneously. High school graduate of extremely high (low) ability will (not) be able to enter college disregard of the policy. Financial constraints also play a role. Therefore, only part of the high school graduates is really affected, and these people at the margin are our focus. Therefore, we used a LATE framework developed by Imbens and Angrist (1994).[2]

We find that higher education expansion increased the probability of going to college tremendously, and that different populations "benefit" from this policy differently. In language of treatment effect literature, *compliers* are less likely to be minority female, they are more likely to come from one-child family, and they are less likely to come from central-western region. These results indicate that the economic unbalance in China makes it hard for all

---

[1] Of course, due to the fact that we don't have information of the exact year when people go to college (or graduate from high school), this separation is doomed to have errors.

[2] See also Angrist and Pischke (2008).

families to benefit from the expansion policy homogeneously. The story underlying this result is that poor families face more severe credit constraint or that the economic conditions make those from poor families less ready to go to college (due to poor academic performance in high school). This result is consistent with Wang et al, (2007). They find "evidence consistent with the growing influence of private financial constraints". The evidence that Wang et al.'s conclusion is based on are the estimated returns to education using several repeated cross section between 1988 and 2002. Our evidence is more direct.

The expansion increased the supply of college graduate, whose employment is a major problem in current China. Meanwhile, the proportion of high school graduate decline relatively. In the second part of our paper, we investigate how the expansion policy affects the relative income between high school graduate and college graduate. Instead of using individual level data, we study how the expansion affects intra-sector inequality and average income levels for high school and college graduate respectively. To identify this relationship, we introduce provincial level higher education resources before the expansion as another dimension of variation, because individuals in provinces with more tertiary education resources are much more likely to benefit from the expansion policy. We find that sector level inequality for high school graduates and college graduates is lower if there're more younger people *and* its province have more college resources. The increase of income level for high school graduate played the major role, and the slight decline in income of college graduate played an insignificant and minor role. We regard the above relationships as causal, and these causal effects have strong implications for evaluating the expansion policy.

Although we do not estimate the return to education directly, we can interpret the results as follows: expansion policy decreased the return to college. This relates our paper to the existing literature on China's college premium. It seems well established that the return to education has increased gradually in the 1980s and 1990s, and the increase in college premium is more prominent (Zhang et al., 2005; Li and Ding, 2005, among many others). However, the most updated estimates are still rare,[3] and there's no research explicitly taking the expansion policy into consideration, to our best knowledge. From methodological point of view, our paper is similar to Han and Zhang (2008), and Meng and Gregory (2007). Both studies look into how the Cultural Revolution affected the education opportunities of individuals. Han and Zhang's focus is on how people continue their education after the interruption. Meng and Gregory estimate the economic cost of the Cultural Revolution from the perspective of

---

[3] In Zhang et al, (2005), the latest data they use is for 2002. In Wang et al, (2007), the data is for 2002. At this time, the economic consequences of expansion policy are not prominent because people go to college after expansion haven't entered into the labor market yet.

forgone earnings. For studies in other country, Duflo (2000) is an extraordinary example who studies how the school construction in Indonesia increased the opportunity of education and therefore economic possibilities. Duflo also used a DID strategy to identify how school construction affect education and income.

This paper is organized as follows. Section 2 briefly described the higher education transformation in the late 1990s and our data. Section 3 studies how expansion policy affects education opportunities, in particular, we characterized the compliers of this natural experiment. In section 4, we investigate how the expansion policy affect intra-sector inequality between college graduates and high school graduates. Section 5 concludes.

2. China's Great Leap Forward in Higher Education: Background and data

2.1 Expansion, Rise in Tuition, and Decentralization

With the reform and opening up policy, China's education system also underwent major transformations. The process of transformations is well summarized in the following four major official documents:

- *Decisions on Reforming the Education System,* (Central Committee of CPC, 1985)
- *Guides for China's Education Reform and Development,* (The State Council, 1993)
- *On Deepening the Reform of the Higher Education System*, (The Ministry of Education[4], 1995)
- *Higher Education Law of the People's Republic of China,* 1999

Closely related to our research are the following three aspects: scale expansion, abolishment of heavy subsidies (rise in tuitions), and decentralization between local and central government in the management of higher education. From 1978 to 1998, the scale of higher education kept increasing. The number of colleges increased from 598 to 1022, the number of new college students enrollment increased from 0.4 million to 1.08 million, and the number of college students increased from 0.86 million to 3.41 million. Although the scale of higher education enlarged continuously (both in terms of establishment number and in terms of the students or graduates number), the growth rate is much lower than those in and after 1999. In early 1999, the central government decided to increase the number of students admitted to tertiary education by 0.22 million. In June, the central government and the Ministry of Education suddenly made an announcement that a further 0.33 million new students will be admitted. These decisions made 1999 a historical year in the development of China's higher education. The number of new college students experienced the largest

---

[4] The former National Education Committee.

increase ever since (48%).[5] For many of the high school graduates and their families, the expansion was unexpected. Given that the college entrance exams was in early July, the announcements made in early 1999 and especially the one in June will not change the behavior of high school graduate much. The "unexpectedness" of the announcement made this policy more like an experiments.[6] In subsequent years, the number of new college students kept increasing. In 2005, the number of new college students was 5.04 millions, 4.7 times of that in 1998. Meanwhile, the total number of college students in China ranked the first all over the world, amounting to 23 million. And the gross enrollment rate of higher education increased by 11.2%, reaching 21%.

Under the traditional central planning regime, higher education was heavily subsidized. With economic reform and with the enlargement of higher education scale, Chinese government lowered the level of subsidies for higher education gradually. In "Guidelines for China's Education Reform and Development" published in 1993, the central government made it clear that *higher education is noncompulsory, and students should pay tuition in principle* (The State Council, 1993). From 1995 to 2004, the tuition fee increased from 800 RMB per person year to 5000 RMB per person year on average (Yang, 2006). The expenditure on education ranked the first in the total expenditure in the 10th 5-year-plan (CYCRC, 2007). The high tuition level makes some families facing more severe financial constraint, so that some students from poor families gave up the college opportunities.[7] This also means that the poor families would benefit less from the expansion policy. In addition to the abolishment of heavy subsidies, another major reform is to cancel the planning system for allocating college graduates to work positions. The new system resembles a two-sided market. The graduates find jobs based on mutual agreements with employers, with much less intervention from the government. With the expansion of higher education, college graduates face more and more severe competition, putting great pressure on the labor market.

---

[5] Some media described this reform as the "Great Leap Forward of Higher Education."
(http://gaokao.zhongzhao.com/article/2007-7-15/200771513252290.shtml).

[6] The initiation of higher education expansion in the late 1990s is also closely related to the economic conditions at that time. In 1997, the 15th National Congress of Communist Party of China quickened the reform of the economic regime. Large amount of formerly state owned enterprises were privatized. This caused large amount of laid-off or unemployed workers. Meanwhile, the financial crisis happened in 1997 also had a negative impact on Chinese economy, which deteriorate the employment condition. Under these backgrounds, higher education expansion was initiated as an instrument to alleviate the unemployment problem and to stimulate consumption.

[7] According to some calculation, the expenditure per student per year is around 10,000 RMB. The net per capita incomes in urban and rural areas were 9422 RMB and 2936 RMB respectively. It cost 4.2 and 13.6 years of net income for an average urban individual and rural individual respectively to support one child to finish a 4-year college education. Although some measures were taken (scholarship, subsidized loans for examples), the credit constraint problem seems not have been alleviated effectively (Yuan, 2008).

Another aspect that influences the effect of the expansion policy is China's geographic unbalance and segregation, not only in terms of economic development but also in terms of administration of higher education resources. One important aspect is the decentralization of the administration of colleges and universities in the mid 1990s. With the deepening of the reform, the higher education system has evolved into a two-layer system, with the central government (through the Ministry of Education) managing a few brand comprehensive universities, and the provincial government, municipalities, or autonomous regions managing most local universities. In effect, decentralization gives provincial government great autonomy. The admission quota to college and the setup of specialty fields that are allowed to have students can be determined at the provincial level. It's true that individuals have some freedom to select college or university throughout the whole nation. But the allocation of admission quota often biased toward local students to great extent, and this is even true for universities being managed by the central government (Peking University, Fudan University, Sun Yat-Sen University for example). Provinces with more higher education resources have more capacity to expand after 1999, and individuals from these provinces are more likely to "benefit" from the expansion policy.

2.2 Data

The data we use is a one-fifth random draw from the 1% census data of China conducted by the National Bureau of Statistics (NBS) in 2005. The sample size is around 2.3 million individuals covering 31 provinces, municipalities and autonomous regions. Because our focus is the effect of the expansion policy after 1999, we keep only those who are most likely influenced (aged 22 to 35). Because we don't have the information when people take the college entrance exams, we assume that people enter the primary school at age 6, and they go to college at age 18.[8] The theoretical year for them to take the college entrance exams is between 1988 and 2001. In fact, many of them have not taken the college entrance exams. They may have dropped out of school before they got high school degree. Even they finished the high school they may also not to have the exams. But in order to know whether they belong to the cohorts affected by the policy, we need the hypothetical college exam year. Table 1

---

[8] The typical timing of education in China is as follows: go to primary school at age 6, after 6 years of primary school, they go to junior middle school for 3 years. Completing junior middle school became compulsory from the 1980s (so called 9-year compulsory education). After junior middle school, people can choose to go to high school which takes 3 years to complete. They can also go to technical or professional school (*zhongzhuan*). Upon graduation from high school, graduates participate in college entrance examinations if they want college education. Usually people with lower scores go to 3 year professional college (dazhuan); and those with high score go to 4 year college. It's worth mentioning that, even until very recently, the education regime is different among various regions. In some areas, primary school takes 5 years only. But people go to school 1 year older. Therefore this difference will not have effect on our calculation. However, we don't pretend that there are no errors from other sources.

presents the distribution of different ages. The average age is about 29. Around 52% are females. Slightly more than 10% are minorities. 11% of them have college or above degrees.

3. Higher Education Expansion and Education Opportunity

3.1 Expansion and Education Opportunity

For every age group, we calculate their theoretical year of taking college entrance exams. Figure 1 gives the share of different education levels in each age group. We kept only those with high school degree and above for the following two reasons. First, junior middle school is compulsory according to the law. Second, the people who are affected by this expansion policy are those high school graduates. Within the age groups between 28 and 35, the number of high school graduate decreased from those aged 35 to 28. There is considerable fluctuation from those aged 28 to 22. One notable thing is that the number of high school graduate dropped significantly for the 1998 cohort. The reasons underlying this drop are unclear. We made the following conjectures. First, the total number of this cohort is relatively small. Second, some graduates who are expected to take college exams postponed their exams to the next year once they anticipated the expansion in 1999. Another possibility is that those failed in 1998 took the exams again in 1999 and succeeded. The retiming story is possible but is not consistent with the background of the policy. Once we consider the proportion instead of the number, the trend is less volatile. Almost over the whole period, the proportion of high school graduate increased gradually although the absolute number decreased. The increasing trend stopped at 1998. The above possible reasons are applicable here also.

The change in the number of (4-year) college graduate and above is the most remarkable. Before 1998, the change in the number of college graduate is small, with the number of male college graduate decreased slightly and females increased slightly. In 1999 and thereafter however, both the number of male and female college graduates witnessed huge jumps. The number of new college student increased by nearly 50% in 1999, and the sharp increase continued in subsequent years.

The two panels in the middle column of Figure 1 show the number and share of professional college (3-year professional college). As for the absolute numbers, males and females have different trend. The number of male professional graduates decreased, whereas the number for females increased. When we consider the proportion, they increased for both genders. But the trend stopped even before 1998 (1995 for male and 1997 for female). There seems little relationship between expansion policy and the change in number of professional graduates.

In the following, we investigate how the expansion affects individual's education choice using a multinomial logit model (see Table 2). Notice that we already dropped the sample of junior middle school and below. An individual with at least high school level, has three choices upon graduation: employment, profession school (3 years), or go to college (above 4 years). Therefore, we have the following model:

$$\Pr(I=0|z) = \frac{1}{1+\sum_{j=1}^{2}\exp(\beta_j+\gamma_j*policy+\delta_j*\mathrm{X})},$$

$$\Pr(I=s|z) = \frac{\exp(\beta_j+\gamma_j*policy+\delta_j*\mathrm{X})}{1+\sum_{j=1}^{2}\exp(\beta_j+\gamma_j*policy+\delta_j*\mathrm{X})} \quad (s=1,2)$$

I=0, 1, 2 represents the above three choices for a high school graduate (0 for employment, 1 for profession school, and 2 for college). I=0 is used as the reference group in the estimation. *policy* is a dummy for expansion policy (policy=1 and policy=0 for younger and older groups respectively). We can therefore calculate how the expansion policy affects individuals' choice (holding others constant):

$$RRR \equiv \frac{\Pr(I=s|z')/\Pr(I=1|z')}{\Pr(I=s|z)/\Pr(I=1|z)} = \exp(\gamma_s)$$

Suppose s=2. The probability ratio of *go to college* to *employment* increased by $\exp(\gamma_2)$ times after expansion. If $\gamma_s > 0$, that means the relative probability of going to college increases. The results are reported in part A of Table 2. The first two columns are for males. Obviously, the relative probability of going to college increased with expansion policy (RRR=EXP(0.2)>1). The relative probability of going to professional school doesn't change significantly however. In addition, we find no significant increasing trend of going to college for male (T for the year of taking college entrance exams). Column 4 and 5 are for females. Similar to male, the relative probability increased significantly with expansion, and the relative probability of going to professional school have not changed significantly around the time of expansion. Different from male, the increasing trend of going to college is significant for females.

In order to gain more confidence that the increase in relative probability is due to expansion policy, we do a placebo study. Instead of treating those having college entrance exams after 1999 as one single category, we separate them into three subgroups (1999, 2000, and 2001). Meanwhile, we also create a dummy for the 1998 cohort. It turns out the

coefficients for year 1998 is not significant. This means that the significance for dummies for 1999, 2000, and 2001 is due to the expansion policy, not due to other factors influencing the time trend.

We perform an independence test next. Imagine two scenarios. In the first one, a high school graduate chooses between employment and going to college. With the expansion policy, the probability of going to college will increase and that of seeking employment with high school degree will decrease. In the second scenario, there's a third choice (professional school). High school graduates can choose professional school in this scenario. Those planned going to professional school may also choose college now. Therefore the choice between employment and college education may be dependent on whether there is a third choice. Whether the dependence property holds also has implication for the restructuring of higher education. To test dependence property statistically, we delete those professional graduates. The results are in column 3 and 6 for male and female respectively. As there are only two choices in the estimation, the multinomial logit model degenerates to a binary logit model. For male, the coefficients in the choice of college are very close to their counterparts in the multinomial logit model. In both part A and part B, we can't reject the hypothesis that the two sets of coefficients are equal. The test on the coefficients of the reform variables (either the single policy dummy in part A or multi dummies in part B) also has the same conclusion. The result for female is a little different. Statistical test rejects that the coefficients are equal. Even though, the difference is small in magnitude of the estimates.

We have shown that the third alternative (professional college) have no significant effect on the choice between employment and college. This conclusion is important for two reasons. First, it implies that there was no restructuring with expansion. Second, it simplifies our study tremendously by allowing us to concentrate on how the expansion policy affects individuals' choice between two alternatives: employment upon graduate from high school and going to college.

3.2 Counting and Characterizing Compliers

(1) LATE framework

Not all individuals benefit from the policy even they belong to the treatment groups. For those with high ability, they can enter college even without the expansion policy; while for those with low ability, they can NOT enter college even with the policy. Then the question becomes: who benefit from the policy? How many people benefit from the policy? To make the problem simple, we divide the total sample into two groups. Those taking college exams

before 1999 (Z_i=0) and those after 1999 (Z_i=1). To keep the two groups balanced, we drop those whose college exam year is before 1995. To reemphasize, we drop those professional graduate in the subsequent analysis. Another variable D is defined as follows: if individual i have a tertiary degree, D_i=1; otherwise, D_i=0.

To reflect the influence of the expansion policy on tertiary education, we introduced more complicated symbols. D_{1i} is the education decision when influenced by the expansion policy, and D_{0i} is the decision when uninfluenced by the policy. Therefore the actual education outcome (treatment status) can be represented by:

$$D_i = D_{0i} + \left(D_{1i} - D_{0i}\right)Z_i$$

We already assume that people take college exams before 1998 or after 1999 are not different systematically except that expansion policy. In other words, the experiment is randomly assigned on the individuals. There is no retiming of tertiary education, there is neither restructuring between different types of tertiary education. The assumption can also be understood as follows: the education decision of people under different policy assignment is independent of the policy assignment.

$\left[D_{1i}, D_{0i}\right] \amalg Z_i$ (A1: independence assumption)

It's debatable that whether A1 holds. However, pretending A1 holds a fairly good starting point. Relative to A1, the following assumption is more reasonable.

$D_{1i} - D_{0i} \geq 0 \quad \forall i$ (A2: monotonicity assumption)

Assumption A2 says that anyone's possibility of going to college will not be decreased by the expansion policy. A2 also means that we have two groups of observations: those who are influenced by the expansion policy (D_{1i}=1 and D_{0i}=0) and those who are not affected by the policy (D_{1i}=0 and D_{0i}=0, or D_{1i}=1 and D_{0i}=1). We call the former group *compliers*. Under the A1 and A2 assumptions, we can calculate the share of the compliers (see appendix for more details):

$$P\left[D_{1i} > D_{0i}\right] = E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]$$

This means we can get the share of compliers in the sample by simply calculating the difference in admission rate between those with Z_i=1 and those with Z_i=0. We can also get the result by regressing D_i on Z_i (first stage).

We can also calculate the share of compliers among those with tertiary degrees (see appendix for details).

$$P[D_{1i} > D_{0i} \mid D_i = 1] = \frac{P[Z_i = 1](E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0])}{P[D_i = 1]}$$

In words, the proportion of tertiary graduates who are compliers is given by the first stage, times the probability the policy is switched on, divided by the proportion who have tertiary degrees.

Finally, we characterize the compliers. Although the complier is a clear concept, we cannot specify which individual is a complier. What we can do is to see their relative characteristics. For simplicity, we only consider the characteristics with 0-1 variations (see appendix for details).

$$\frac{P[x_{1i} = 1 \mid D_{1i} > D_{0i}]}{P[x_{1i} = 1]} = \frac{E[D_i \mid Z_i = 1, x_{1i} = 1] - E[D_i \mid Z_i = 0, x_{1i} = 1]}{E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]}$$

(2) Results

Table 3 report the share of compliers in the whole sample and in the treated sample (college graduate). Compared with those taking college exams in 1996, 1997 and 1998, the share of college graduate in those taking exams in 1999, 2000, and 2001 is over 4%. This means that 4% of our sample is those really benefit from the expansion policy under the monotonicity and independence assumptions (we already delete those with degrees below junior middle school). Next, we calculate the proportion of compliers in the treatment group. The results show that among the college graduate, nearly 12-13% of them are compliers. Without the expansion policy, they will not be able to enter college. The results for male and female are not significantly different. Deleting those with profession degrees and master degrees don't change the results much.

To minimize the difference between different age groups, we also restrict our sample to those taking exams between 1997 and 2000 (see part B of table 3) and to those between 1998 and 1999 (part C). As the sample shrink, the proportion of compliers in the whole sample and in the treatment group decreased. There are two possible reasons. First, the scale of expansion enlarged from 1999 to 2001. Second, we don't know the exact year the individuals taking college exams. What we can do is to calculate year exam year according to age. There must be errors, which will downward bias the result.

Next, we characterize the compliers. This exercise is meaningful for two reasons. First, it's necessary for us to evaluate the expansion policy. No doubt, the beneficiaries are those lies in the middle of the ability distributions. Are they different from those non-compliers in terms of personal characteristics and family background? Second, this exercise will let us know who are on the margin of tertiary education decisions. This type of information is very important for policy makers.

We consider four characteristics, namely sex, ethnicity, region, and single-or-non-single child. The first column in table 4 indicates that the proportion of female compliers is slightly higher than the proportion of females in the whole sample. The ratio is 1.02 when all samples with above high school degree are inclusive, and it turns to be 1.06 when observations with professional degrees or master graduate are dropped. Whether the policy favors minority depends on sex. For males, the proportion of minorities in the compliers is similar to the proportion of minorities in the total male sample, indicating that the expansion does not favor nor disfavor minority males. For females, however, the proportion of minority female in the compliers is significantly less than the proportion of minority females in the total female sample. Minority females are less likely to benefit from the expansion policy.

To see the regional distribution of compliers, we divide the whole sample into two regions, namely eastern region and central-western region. One difficulty is that entering college is often associated with changing of hukou status. We therefore use the region of hukou five years ago. The proportion of compliers come from eastern areas is larger than the proportion of eastern individuals in the whole sample. This is more obvious for females. These results suggest that people from the eastern benefit more from the expansion policy. This conclusion however depends on the assumption that the hukou information we use reflect accurately the hukou region before the individuals taking college entrance exams. If not, the results here may just reflect the fact that college students choose to stay at eastern regions after graduation even they come from central-western regions. This is very probable because higher education is one of the main channel through which people especially those from rural areas change their hukou status.

Finally, we consider whether an individual comes from single-child family. The results indicate that the proportion of single child compliers is much higher than the proportion of single child in the total sample. The single child families are more likely to benefit from the expansion policy. This may indicate another fact of the expansion policy: rural residents benefit less than urban residents, because the rural families are much more likely to have more than one child.

The simple exercises here indicate one important fact. The higher education policy did not benefit people of different characteristics homogeneously. Minorities, families from central-western regions, non-single-child families benefit less than their majority, eastern, and single-child counterparts. These results are not unexpected. As minorities, central-western families and non-single-child families are more likely to be poor. High school graduates from these families tend to face more credit constraints. This is especially true with the rapid rise of college tuitions. Therefore even with the expansion policy they are unable to go to college. From the perspective of college education opportunity, the expansion policy can't be regarded as a policy that reduces inequality. To create more opportunity for the poor, there should be some companion measures to alleviate financial constraints.

In the appendix, table A1 gives the numbers of colleges and population for different provinces. Obviously, higher education resources are highly unequal among different regions. Given the regional disparity of higher education resources and the decentralization of education administration, individuals in different provinces will benefit from the expansion policy differently. First, this is due to the fact that provincial admission quota has a home bias. Second, it's also due to higher transportation or communication cost when someone goes to college outside province.

Table 5 shows the results. In the regressions, dependent variable is a dummy y with y=1 representing college graduates and y=0 representing employment (high school graduate). Therefore we estimate a linear probability model. Number of college students in school per capita and number of faculty and staff in colleges per capita are used as proxies for provincial level of higher education resources. Of course, higher education resources may be related to regional economic conditions. Therefore, we controlled provincial dummies and province specific time trend. In Table 5, both the coefficients for the resource variables (number of college students and number of staff and faculty per capita) and their interactions with the expansion dummy (policy) are significantly positive. This means that people in resource abundant province are more likely to go to college, and that families in these provinces tend to benefit more from the expansion policy. One caveat is that going to college usually means the location change of household registration. Therefore, we replace the hukou location by those of five years ago. The result doesn't change much (therefore we did not report).

The result in Table 5 is also important in the following sense. Using both cross sectional and time varying variations, we can employ a difference in difference strategy, which did a better job to identify how expansion affects the labor market.

4. Higher Education Expansion and its Labor Market Consequences

4.1 Expansion and employment

Although the expansion policy was initiated in 1999, the data in 2005 are still not ideally appropriate to see the labor market consequences of the policy. As indicated by the first column of Table 6, some of the observations are still in school in 2005, especially for the cohorts that should have taken college entrance exams between 1999 and 2001. If college education takes four years, they should have graduated in 2005. However, some of them may go to graduate school, and some may go to primary school at older age, and there are still other possibilities such as repeated schooling. Unfortunately, what we can do is to *calculate* the year for taking college exams according to a specified timing of education. Next, we first investigate the short term effect of the expansion policy by giving some simple summary statistics.

We delete those who are still in school in the following columns. In the short term, the labor market outcome for cohorts that take college exams after 1999 is worse than the earlier cohorts. The labor participation rates are significantly lower for the 1999-2001 cohorts. Within the unemployed groups (and especially the unemployed college graduate), there is a larger share who are "unemployed immediately after graduation" for these policy cohorts. For college graduate, the share is over 80%. They are also more likely to depend on their family for living. There are two possible reasons. One is policy effect and the other is age effect[9]. As for age effect, it's natural that younger people have lower participation rate and depend more on their families. As they aged up, they will find better jobs and will eventually live on their own income. Unfortunately, we cannot separate these two effects. What we do is to minimize age effect. We therefore focus only on two age groups, 1998 cohort and 1999 cohort. Still, there is significant difference between these two cohorts in terms of labor participation rate, the share of "unemployment immediately after graduation", and the share of living on families. Those who find a job through on-site recruiting fairs also increased. All these indicate that the policy has brought pressure for the labor market. One thing worth noting is that all these effects are short term effect. In the long run, there will be more individuals enter the labor market. The college education may make them more better off. But we need to wait for better data.

4.2 Expansion and within sector inequality

The pressure brought by the expansion policy should also have effect on income distributions. We use $g(w)$ to represent income distribution, $h(x)$ to represent the composition distribution

---

[9] There is also cohort effect. We can think of pooling the cohort effect with age effect.

(the ratio of high school graduates to college graduates), and $f\left(w\,|\,x\right)$ to represent skill price. Therefore we have:

$$g\left(w\right) = \int f\left(w\,|\,x\right)h\left(x\right)dx$$

It is clear that the expansion policy changed the composition $h\left(x\right)$ significantly, with the share of college graduate increase and that of high school graduate decreased. When there are only high school graduates and college graduates in the sample, $f\left(w\,|\,x\right)$ represents return to college. With the composition change mentioned above, we expect the return to college to change also (probably to decline).

To investigate how expansion affects inequality, we employ a Difference-in-Difference strategy. The specification is as follows:

$$inequality_{ij} = \alpha + \beta * share_{ij}\mathrm{X}student_i + \chi * share_{ij}$$
$$+ \pi * student_i + \lambda * \mathrm{X} + u_{ij}$$

To have more observations, we aggregate our income and policy data at the city (i) *and* industry (j) level. In the above model, *inequality* refers to standard deviation of log hour income in region i and industry j, *share* is the proportion of individuals who are affected by the expansion policy in the same cell. *student* is the number of college students in school at the per capita level in 1998 in the province where region i locates. We use this variable as a proxy for the higher education resources at different provinces. We don't use the city level college student data because the both the admission quota and graduate allocation are at the provincial level. As this variable is closely related to economic conditions and other factors, we controlled some other variables (X), including GDP per capita, urbanization rate (proportion of urban population), sex ratio, region dummies (east, central and west), and industry dummies.

It's easy to understand that we can't interpret the coefficient of *share* as the causal effect of expansion on inequality. It's true that the larger the *share* variable, the more the population affected by the expansion policy. But it's also true that people at different ages will have different income levels according to human capital theory even without expansion policy. The coefficient of *student* can be taken as the causal effect neither. Although provinces with more prior resources (represented by college student) tend to expand to a larger extent, there are many other region specific factors correlated with it. In this paper, we interpret the coefficient

for the interaction of these two variables (β) as the causal effect of expansion on inequality.

That means if there are more college graduates not only because there're more younger people but also because they are in a province with more higher education resources, we interpret this as the result of expansion.

In column 1 of Table 7, the coefficient for the interaction term is -0.074, meaning that within sector inequality decreases with expansion. The effect is not significant however (standard error is 0.062). In column 4, we delete observations for Beijing, Shanghai, Tianjin, Chongqing, and Guangdong, because these 5 provinces (municipalities) are special in that much more higher education resources are concentrated in these regions, and that a large number of colleges or universities are managed by the central government (and therefore universities in these areas have much bigger proportions of student from other provinces). For the remaining provinces, the effect of the expansion becomes larger and more significant. The coefficient for the interaction term is -0.144 and significant at the 5% level, which means that expansion policy significantly decreased within sector inequality.

To show more directly that the negative effect is due to the relative increase in the number of college graduate, we use the interaction term ($share$X$student$) as the instrument for the current proportion of college student, and run a two stage least square estimation:

$$collegeshare = \delta_0 + \delta_1 * share_{ij}\mathrm{X}student_i + \delta_2 * share_{ij} \qquad \text{first stage}$$
$$+ \delta_3 * student_i + \gamma * \mathrm{X} + v_{ij}$$

$$inequality_{ij} = \alpha + \beta * \widehat{collegeshare}_i + \chi * share_{ij} \qquad \text{second stage}$$
$$+ \pi * student_i + \lambda * \mathrm{X} + u_{ij}$$

The first stage results are reported in column 2 of Table 7. In the analysis using aggregated data at the region-industry level, the coefficient on the interaction term is still significantly positive, consistent with the results in Table 5. In the second stage regression, the coefficient for the share of college students (*college*) is negative (not significant however). Therefore, we have evidence that expansion policy increased the fraction of college graduates in the labor force, which in turn decreased the intra-industry inequality. Column 5 and 6 are the results after we delete the four municipalities and Guangdong province.

The exercise from column 7 to column 12 repeated the previous exercises. The difference is that we keep only those with urban hukou. No matter OLS or 2SLS, the results become more significant with the direction of the coefficients unchanged.

4.3 Expansion and income levels

As emphasized before, income distribution (inequality) depends on two aspects: distribution of characteristics and skill prices. Even we have established the results that expansion causes inequality to decline, it's still difficult to evaluate how the expansion influences relative income levels. In Table 8, we investigate how the expansion policy affect the relative income level between high school graduates and college graduates using the same strategy as above. The observation is still at the region-industry level. When we use the difference in log wages between college graduates and high school graduates, the coefficient for the interaction term is negative (-0.168 with the standard error of 0.189). This means that the expansion policy decreased the return to college within sectors. In column 2 and 3, we use the log wages for high school graduates and college graduates as dependent variable respectively. The coefficient in column 2 for high school graduate is 0.155 (significant at the 1% level), and that in column 3 for college graduates is -0.015 (not significant). The decrease in income gap between college graduates and high school graduates is mainly due to the increase in the income levels for high school graduates. The decrease of the income level for college graduates plays a minor role. In column 4 and 5, data for municipalities and Guangdong province are dropped. The increase for high school graduates decreased a little bit, which is 0.125 significant at 10% level. That for college graduates become more negative (-0.068) but still not significant. In columns 6, 7, and 8, we delete observations that have missing wage value either for high school graduates or for college graduates to make observations in these three columns identical. Still, expansion policy cause the income gap to decrease, which is mainly due to the increase in the income level for high school graduates. In column 8, the coefficient is positive, but still not significant.

5. **Conclusion and Discussion**

The radical expansion of higher education commenced in 1999 and in subsequent years is historical and is believed to have huge impact on China's society and economy. Yet little empirical work has been done to evaluate its effects on education opportunities, and on labor market outcomes.[10] Using a random sample from the 1% census data in 2005, we find that compliers (those go to college because of the expansion policy) in the expansion are more likely to come from eastern and urban areas, and less likely to be minority female, from poor families and from families with more than one child. We also find that higher education resources at provincial level have significant impact on how the expansion affects college opportunities. All these indicate that the expansion effects are not homogeneous among

---

[10] Li et al, (2008) is a notable exception. However, their focus is on the global implications of this transformation.

regions and among different populations. One reason underlying these results is financial constraints, which is consistent with the fact that college tuitions increased by huge amount during the late 1990s and early 2000s. The other side of the same reason is the inadequacy of measures to cope with these constraints. The current scholarships or student loans programs haven't solved the problems effectively, and more work need to be done. Another reason is decentralization of education system, which gives provincial government great autonomy, including setting admission quota each year. Therefore, decentralization (with the fact that higher education resources are not evenly allocated) makes the effect of expansion different for different provinces.

The first set of the results also has implications for some related policy debates. One is the financial problem of colleges and universities. As the scale of higher education was expanded, college or universities run in shortage of spaces for teaching, studying and lodging. Therefore, many of them started construction of new buildings. One important source of finance came from the banks. According to *China's Education Blue Book in 2007*, the loan scale reached to 450 to 500 billion by the end of 2006.[11] How to pay the loans is a big issue in China's higher education (Lin, 2008).[12] Someone argue that the loans should be bored by the government because the expansion benefits *all the people*.[13] We don't want to go to detail of the debt problems, but our results invalidate the above claim.

Our second set of results confirms the great pressure on labor market (especially urban areas) induced by the expansion. College graduates have more difficulties to find employment, and the income gap between college graduates and high school graduates is lower where the expansion policy has more impact. These results also provide empirical evidence for the current evaluation on the expansion policy itself. The expansion has long been blamed for being too radical that caused serious unemployment problem for college graduates. Casual observations of the crowds in recruiting fairs support this view. However, without systematic empirical evidence, this view can be challenged easily.[14] Our empirical support the claim that expansion policy creates pressure on the labor market. However, our results don't imply that the expansion policy is unjustified. The expansion policy has other effects including

---

[11] According to the estimate of cppcc in 2007, the number is 250 billion (http://roomx.blogchina.com/ads/6359721.html).

[12] http://www.gmw.cn/content/2009-03/10/content_896312.htm

http://news.sohu.com/20090308/n262668071.shtml

http://www.infzm.com/content/25216

[13] http://news.sohu.com/20090308/n262668071.shtml

[14] There're several competing explanations for the difficult conditions faced by college graduates. Some pointed out that the difficulty is due to China's dualistic feature (http://www.gmw.cn/CONTENT/2009-12/31/content_1031238.htm); others emphasize that more unemployment for young labors is a common feature of the labor market, therefore we can't blame expansion policy for the difficult conditions faced by college graduates (http://www.ilib2.com/A-QCode~xcj200902040.html; http://news.xinhuanet.com/politics/2009-03/08/content_10966534.htm).

accelerating China's urbanization process and making the economic conditions more stable. These are beyond the scope of this paper.

References:

Ai Chunrong, and Edward C. Norton, (2003), "Interaction terms in logit and probit models", *Economics Letters*, 80 (2003) 123–129

Angrist J. and Pischke J., 2008, *The Mostly Harmless Econometrics*, Princeton University Press.

Angrist, Joshua D & Krueger, Alan B, (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?," *The Quarterly Journal of Economics*, Vol. 106(4), pages 979-1014, November.

Autor, David, Lawrence Katz, and M. S. Kearney. (2005), "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists." NBER Working Paper No. 11627.

Deng, Quheng and Gustafsson, Bjorn, "China's Lesser Known Migrants" (May 2006). IZA Discussion Paper No. 2152., http://ssrn.com/abstract=908236

Duflo, Esther. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *A.E.R,* 2000, 91(4): 795-813.

Heckman James J. and Xuesong Li, (2004), "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000", *Pacific Economic Review*, 9: 3.

Imbens, G. W. and Angrist, J. D., 1994 "Identification and estimation of local average treatment effects" *Econometrica*, 62(2), pp. 467-475.

James J. Heckman and Xuesong Li, 2004. "Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000," *Pacific Economic Review*, Blackwell Publishing, vol. 9(3), pages 155-171, October.

Li Yao, John Whalley, Zhang Shunming, and Zhao Xiliang, 2008, "The Higher Educational Transformation of China and its Global Implications", NBER Working Paper #13849.

Meng, Xin, and Robert Gregory, 2007, "Exploring the Impact of Interrupted Education on Earnings: The Educational Cost of the Chinese Cultural Revolution", Discussion Paper No. 2548

Wang, Xiaojun, Belton M. Fleisher, Haizheng Li, and Shi Li, 2007. "Access to Higher Education and Inequality: The Chinese Experiment," IZA Discussion Papers 2823, Institute for the Study of Labor (IZA).

Xing, Chunbing, 2009, "Migration, Self-selection, and Income Distributions: Evidence from Rural and Urban China," MPRA Paper 17036.

Zhang, Junsen, Zhao, Yaohui, Albert Park, and Song Xiaoqing, (2005), "Economic returns to schooling in urban China, 1988 to 2001," *Journal of Comparative Economics*, 33, 730-752.

References in Chinese:

National Education Committee, 1995, *On Deepening the Reform of the Higher Education System*, http://www.gzzb.gov.cn/gb/ShowArticle.asp?ArticleID=8992.

State Council, 1993, *Guidelines for China's Education Reform and Development* http://www.moe.edu.cn/edoas/website18/level3.jsp?tablename=208&infoid=3334.

Li Shi and Ding Sai, 2003, "Long-term Change in Private Returns to Education in Urban China", Social Science in China, No. 6.

Lin Li, 2008, "Fighting the Debt Crisis in Universities: Issues and Approaches", in *Blue Book of Education in China*, 2008.

Sun Zhijun, Du Yuhong, 2009, "Does Education System Affect Rural Resident's Schooling and Earnings? Evidence from Survey Data in Rongan County,Guangxi Province", *Chinese Journal of Population Science*, No.4.

Yang Dongping, 2006, *Transition and Development of China's Education*, Social Science Academic Press.

China Youth and Child Research Center (CYCRC), 2007, *Report on the Development of the China Youth During the 10th and 11th 5-Year-Plan*.

Central Committee of CPC, 1985, *CCCPC'S Decisions on Reforming the Education System*. http://news.xinhuanet.com/ziliao/2005-02/06/content_2554936.htm

Figure 1 Number and proportion of different education levels
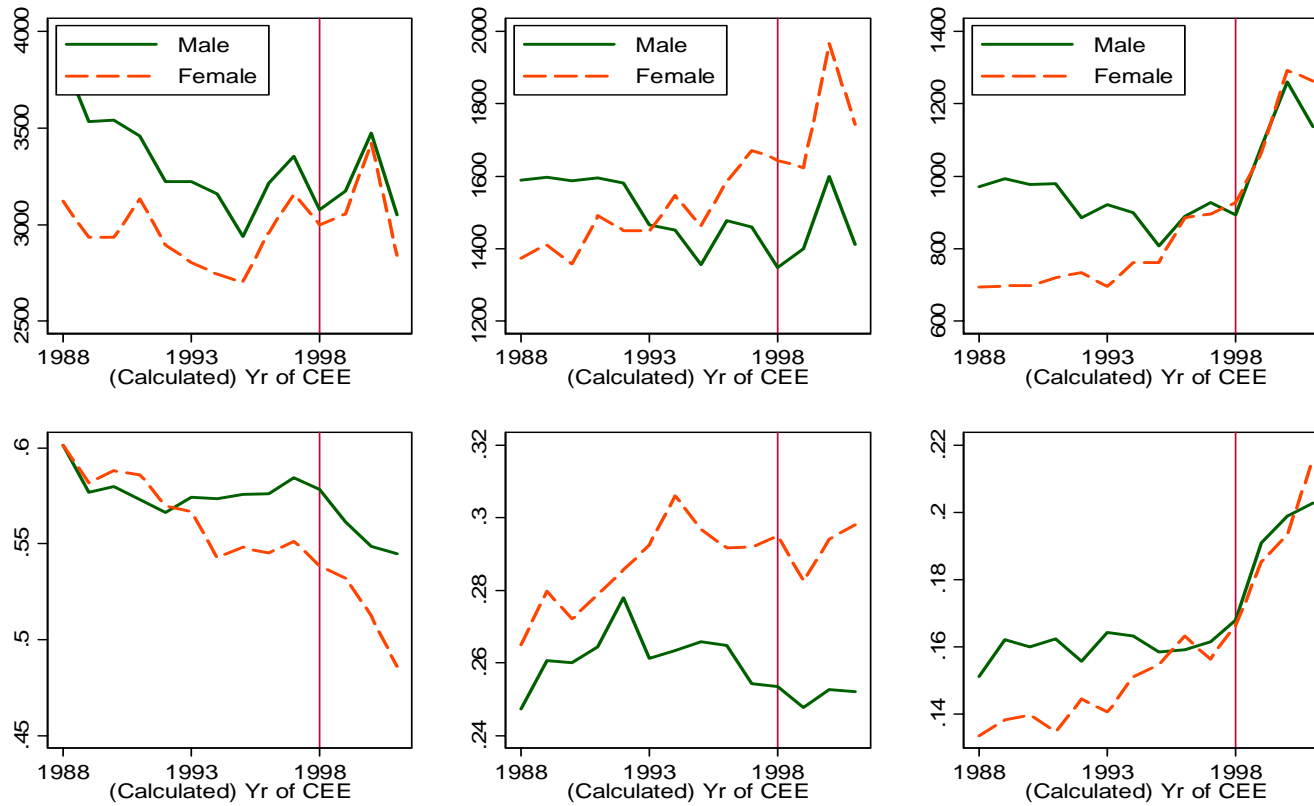
Note: from left to right, each column for high school, 3-year college, and 4-year college (and above) respectively. The upper panels are for absolute numbers and the lower panels are for proportions.

Table 1 Summary statistics

Variables	weighted	unweighted
Age	29.14	29.06
Female (%)	52	52
Minority(%)	10	12
Eastern(%)	48	53
One Chile(%)	33	34
Education Level(%)		
Below Primary	2.65	3.28
Primary	18.27	18.93
Middle School	52.87	49.97
High School	15.07	15.65
3 Year College	7.16	7.59
4 Year College	3.65	4.16
Master and above	0.33	0.41
Age Groups(%, theoretical year of taking exams in "( )")		
35 (1988)	9.82	9.47
34 (1989)	8.94	8.75
33 (1990)	8.51	8.37
32 (1991)	8.21	8.15
31 (1992)	7.84	7.83
30 (1993)	7.13	7.1
29 (1994)	6.74	6.75
28 (1995)	5.95	5.95
27 (1996)	6.23	6.29
26 (1997)	6.32	6.38
25 (1998)	5.94	6.08
24 (1999)	5.99	6.11
23 (2000)	6.68	6.84
22 (2001)	5.7	5.94
Obs	562,313	

Table 2 Expansion and education choice in mlogit model, employment as reference group

	Male			Female		
	Model 1		Model 2	Model 3		Model 4
	（1）	（2）	（3）	（4）	（5）	（6）
	professional	college	college	professional	college	college
A:						
T	0.004	0.003	0.003	0.022***	0.031***	0.030***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
policy	-0.007	0.200***	0.202***	-0.028	0.152***	0.156***
	(0.029)	(0.034)	(0.034)	(0.028)	(0.035)	(0.034)
Constant	-8.915	-7.020	-6.803	-44.379***	-62.359***	-61.514***
	(5.833)	(6.965)	(6.950)	(5.936)	(7.559)	(7.551)
R2_p	0.005		0.008	0.006		0.013
N	80835		59911	75582		53804
chi2(4)	6.06			11.48		
Prob > chi2	0.1948			0.0217		
chi2(1)	0.88			6.59		
Prob > chi2	0.3494			0.0102		
B:						
T	0.006*	0.001	0.001	0.023***	0.028***	0.027***
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
yr1998	-0.054	0.016	0.016	-0.033	0.011	0.011
	(0.039)	(0.046)	(0.046)	(0.037)	(0.046)	(0.046)
yr1999	-0.051	0.163***	0.164***	-0.088**	0.095**	0.100**
	(0.040)	(0.046)	(0.046)	(0.039)	(0.047)	(0.047)
yr2000	-0.013	0.224***	0.225***	-0.032	0.141***	0.144***
	(0.041)	(0.047)	(0.047)	(0.040)	(0.048)	(0.048)
yr2001	-0.015	0.250***	0.250***	0.013	0.277***	0.282***
	(0.045)	(0.051)	(0.051)	(0.043)	(0.052)	(0.052)
Constant	-12.835*	-3.935	-3.788	-45.610***	-57.038***	-56.164***
	(6.740)	(8.085)	(8.069)	(6.932)	(8.882)	(8.871)
R2_p	0.005		0.008	0.006		0.013
N	80835		59911	75582		53804
chi2(7)	5.92			14.92		
Prob > chi2	0.549			0.037		
chi2(4)	0.99			8.44		
Prob > chi2	0.9109			0.0769		

Note: ***, **, * represent the significance level of 1%, 5%, and 10% respectively. Standard errors are in (). We controlled region dummies in the regression.

Table 3 The Share of Compliers

| | Samples with above high school degree | |
	Male (1)	Female (2)
A: 96/97/98(Z=0) vs 99/00/01(Z=1)		
P[Z=1]	0.516	0.523
E[D|Z=0]	0.219	0.229
E[D|Z=1]	0.264	0.280
P[D=1]	0.242	0.256
P[D1>D0]	0.044	0.051
P[D1>D0|D=1]	0.095	0.103
B: 97/98(Z=0) vs 99/00(Z=1)		
P[Z=1]	0.521	0.526
E[D|Z=0]	0.221	0.229
E[D|Z=1]	0.260	0.267
P[D=1]	0.241	0.249
P[D1>D0]	0.040	0.038
P[D1>D0|D=1]	0.086	0.081
C: 98(Z=0) vs 99(Z=1)		
P[Z=1]	0.517	0.512
E[D|Z=0]	0.225	0.236
E[D|Z=1]	0.254	0.258
P[D=1]	0.240	0.248
P[D1>D0]	0.029	0.022
P[D1>D0|D=1]	0.062	0.046

Table 4 Characterizing compliers

Characteristics (X)	Male=1	East=1		Minority =1		Single child=1	
		Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
96/97/98(Z=0) vs 99/00/01(Z=1)							
E[D|Z=0]	0.22	0.22	0.23	0.22	0.23	0.22	0.23
E[D|Z=1]	0.27	0.26	0.28	0.26	0.28	0.26	0.28
E[D|Z=0,x=1]	0.22	0.25	0.25	0.17	0.25	0.34	0.37
E[D|Z=1,x=1]	0.26	0.29	0.32	0.23	0.28	0.42	0.47
P[x=1|D_{1i}>D_{0i}]/ P[x=1]	0.94	1.03	1.28	1.37	0.55	1.70	1.88
97/98(Z=0) vs 99/00(Z=1)							
E[D|Z=0]	0.22	0.22	0.23	0.22	0.23	0.22	0.23
E[D|Z=1]	0.26	0.26	0.27	0.26	0.27	0.26	0.27
E[D|Z=0,x=1]	0.22	0.25	0.25	0.18	0.27	0.35	0.38
E[D|Z=1,x=1]	0.26	0.28	0.30	0.22	0.24	0.40	0.45
P[x=1|D_{1i}>D_{0i}]/ P[x=1]	1.02	0.97	1.17	0.96	-0.78	1.41	1.74
98(Z=0) vs 99(Z=1)							
E[D|Z=0]	0.23	0.23	0.24	0.23	0.24	0.23	0.24
E[D|Z=1]	0.26	0.25	0.26	0.25	0.26	0.25	0.26
E[D|Z=0,x=1]	0.23	0.25	0.26	0.18	0.25	0.36	0.38
E[D|Z=1,x=1]	0.25	0.28	0.29	0.25	0.24	0.39	0.44
P[x=1|D_{1i}>D_{0i}]/ P[x=1]	1.13	0.97	1.37	2.36	-0.43	1.10	2.48

Table 5 Prior education resource and expansion effect (Linear Probability Model)

	Male		Female	
	(1)	(2)	(3)	(4)
Policy	-0.008	-0.018	0.004	-0.015
	(0.012)	(0.016)	(0.012)	(0.016)
faculty and staff # in 98	0.099***		0.102***	
	(0.003)		(0.003)	
faculty and staff # in 98*policy	0.017***		0.024***	
	(0.005)		(0.005)	
# of student in college 98		0.137***		0.139***
		(0.004)		(0.004)
# of student in college 98*policy		0.020***		0.031***
		(0.006)		(0.006)
Constant	0.010	-0.060***	-0.009	-0.078***
	(0.006)	(0.008)	(0.007)	(0.008)
R2_a	0.043	0.042	0.05	0.049
F	539	523	568	550
N	59911	59911	53804	53804

Note: ***, **, * represent the significance level of 1%, 5%, and 10% respectively. Standard errors are in (). Province dummies and time trend are controlled.

Table 6 Employment effect of expansion

Age groups (the year of taking exams)	In school	Never work	Unemployed After graduation	Main source of income			Whether looking for a job within last 3 month				
				Labor	Families	Other	HR Agent	Relatives & friends	On-site recruitment	Other	No searching
	(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Above high school											
1993	0.61	12.78	6.72	86.49	10.37	3.14	5.21	14.34	3.02	10.94	66.49
1994	0.89	13.56	10.01	85.86	11.13	3.01	6.21	13.71	3.57	11.72	64.79
1995	1.06	13.70	13.54	85.49	11.59	2.92	8.02	11.66	4.68	13.20	62.44
1996	1.57	14.44	17.08	84.93	12.33	2.74	8.00	13.42	4.90	13.42	60.26
1997	1.80	14.57	23.38	84.75	12.67	2.58	9.12	15.71	4.56	13.68	56.93
1998	4.05	16.12	30.83	83.18	14.28	2.54	10.41	16.40	6.52	13.47	53.20
1999	6.75	16.39	42.91	82.87	14.84	2.29	11.27	15.20	8.84	13.48	51.21
2000	12.85	19.57	51.79	79.66	17.99	2.35	13.86	16.32	10.90	13.09	45.83
2001	26.23	21.47	62.65	77.84	19.80	2.36	15.61	18.97	10.69	12.70	42.03
College graduate											
1993		2.97	11.63	96.54	2.28	1.18	11.63	6.98	9.30	9.30	62.79
1994		4.44	13.85	94.74	3.42	1.84	7.69	10.77	7.69	18.47	55.38
1995		5.18	35.21	94.53	4.16	1.31	12.68	7.04	11.27	22.53	46.48
1996		3.89	33.33	96.05	3.37	0.58	15.00	6.67	16.67	11.66	50.00
1997		6.34	48.00	93.41	5.45	1.14	11.00	14.00	15.00	17.00	43.00
1998		7.53	60.00	91.94	6.81	1.25	26.09	12.17	16.52	19.13	26.09
1999		11.97	75.88	87.73	11.00	1.27	17.17	12.12	24.24	14.65	31.82
2000		17.68	80.73	81.74	16.85	1.41	15.27	9.45	28.36	12.37	34.55
2001		22.53	88.76	77.22	22.41	0.37	20.22	10.67	28.09	8.44	32.58

Table 7 Expansion effect on within region-industry inequality

	All provinces			Municipalities and guangdong deleted			All provinces			Municipalities and guangdong deleted		
	OLS	2SLS		OLS	2SLS		OLS	2SLS		OLS	2SLS	
		First	Second		First	Second		First	Second		First	Second
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
policyXstudent	-0.074	0.185***		-0.144**	0.143***		-0.149**	0.223***		-0.254***	0.178***	
	(0.062)	(0.050)		(0.068)	(0.053)		(0.072)	(0.055)		(0.079)	(0.060)	
College			-0.397			-1.007			-0.668*			-1.425**
			(0.362)			(0.622)			(0.368)			(0.664)
Policy	0.162	-0.356***	0.020	0.298**	-0.281***	0.015	0.368**	-0.440***	0.074**	0.567***	-0.359***	0.056
	(0.128)	(0.103)	(0.024)	(0.138)	(0.109)	(0.030)	(0.149)	(0.114)	(0.030)	(0.162)	(0.122)	(0.041)
student	0.011	0.016	0.017	-0.005	-0.006	-0.011	0.020	0.012	0.029	0.007	-0.009	-0.006
	(0.015)	(0.012)	(0.021)	(0.016)	(0.013)	(0.019)	(0.017)	(0.013)	(0.023)	(0.018)	(0.013)	(0.024)
Gdp_pc	-0.007***	-0.002	-0.007***	-0.009***	0.003	-0.006	-0.006**	-0.003	-0.007**	-0.009***	0.004	-0.004
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.005)
Urban rate	-0.053***	0.155***	0.008	-0.048**	0.158***	0.111	-0.017	0.144***	0.079	0.006	0.147***	0.215**
	(0.018)	(0.014)	(0.059)	(0.019)	(0.015)	(0.101)	(0.024)	(0.018)	(0.059)	(0.026)	(0.019)	(0.105)
Sex ratio	0.002	-0.007	-0.001	0.003	-0.014	-0.011	-0.005	-0.007	-0.009	-0.009	-0.014	-0.030
	(0.017)	(0.014)	(0.019)	(0.018)	(0.014)	(0.025)	(0.020)	(0.015)	(0.023)	(0.021)	(0.016)	(0.033)
R-squared	0.072	0.519	.	0.077	0.521	.	0.069	0.521	.	0.068	0.525	.
N	4499	4499	4499	4069	4069	4069	4071	4071	4071	3656	3656	3656

Note: ***, **, * represent the significance level of 1%, 5%, and 10% respectively. Standard errors are in (). Regional dummies and industry dummies are controlled.

Table 8 Expansion effect on income level

	Depdent var=							
	ln(coll wg)-ln(hs wg)	ln(hs wg)	ln(coll wg)	ln(hs wg)	ln(coll wg)	ln(coll wg)-ln(hs wg)	ln(hs wg)	ln(coll wg)
	All provinces			Municipalities and guangdong deleted		Municipalities and guangd and missing values deleted		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
policyXstudent	-0.168	0.155**	-0.015	0.125*	-0.068	-0.254	0.366**	0.112
	(0.189)	(0.064)	(0.122)	(0.067)	(0.127)	(0.213)	(0.157)	(0.181)
Policy	0.388	-0.368***	-0.095	-0.312**	0.032	0.577	-0.826**	-0.249
	(0.393)	(0.130)	(0.255)	(0.136)	(0.263)	(0.436)	(0.323)	(0.372)
student	0.128	-0.019	0.137**	-0.086**	-0.025	0.083	-0.189**	-0.106
	(0.093)	(0.036)	(0.062)	(0.038)	(0.064)	(0.101)	(0.075)	(0.086)
Gdp_pc	0.010	-0.001	0.017**	-0.011*	-0.007	-0.000	-0.006	-0.006
	(0.010)	(0.005)	(0.008)	(0.006)	(0.009)	(0.012)	(0.009)	(0.010)
Urban rate	-0.218***	0.076***	-0.077	0.082***	0.014	-0.174*	0.168**	-0.006
	(0.081)	(0.023)	(0.065)	(0.025)	(0.067)	(0.091)	(0.067)	(0.077)
Sex ratio	0.000	-0.137***	-0.009	-0.148***	-0.039	-0.007	-0.069	-0.076
	(0.067)	(0.023)	(0.043)	(0.024)	(0.043)	(0.073)	(0.054)	(0.062)
Constant	0.701***	2.136***	2.637***	2.295***	3.041***	0.850***	2.407***	3.257***
	(0.248)	(0.084)	(0.181)	(0.089)	(0.187)	(0.272)	(0.201)	(0.232)
R-squared	0.030	0.279	0.147	0.252	0.079	0.025	0.166	0.078
N	1358	3927	1662	3517	1419	1124	1124	1124

Note: ***, **, * represent the significance level of 1%, 5%, and 10% respectively. Standard errors are in (). Regional dummies and industry dummies are controlled.

Appendix A:

1. Counting compliers in the population:

$$P\left[D_{1i} > D_{0i}\right] = E\left[D_{1i} - D_{0i}\right] \qquad (monotonicity)$$

$$= E\left[D_{1i}\right] - E\left[D_{0i}\right]$$

$$= E\left[D_{1i} \mid Z_i = 1\right] - E\left[D_{0i} \mid Z_i = 0\right] \quad (Independence)$$

$$= E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]$$

2. Counting compliers in the college graduate:

$$P\left[D_{1i} > D_{0i} \mid D_i = 1\right] = \frac{P\left[D_i = 1 \mid D_{1i} > D_{0i}\right] P\left[D_{1i} > D_{0i}\right]}{P\left[D_i = 1\right]}$$

$$= \frac{P\left[Z_i = 1 \mid D_{1i} > D_{0i}\right]\left(E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]\right)}{P\left[D_i = 1\right]}$$

$$= \frac{P\left[Z_i = 1\right]\left(E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]\right)}{P\left[D_i = 1\right]}$$

3. Characterizing compliers

$$\frac{P\left[x_{1i} = 1 \mid D_{1i} > D_{0i}\right]}{P\left[x_{1i} = 1\right]} = \frac{P\left[D_{1i} > D_{0i} \mid x_{1i} = 1\right]}{P\left[D_{1i} > D_{0i}\right]}$$

$$= \frac{E\left[D_i \mid Z_i = 1, x_{1i} = 1\right] - E\left[D_i \mid Z_i = 0, x_{1i} = 1\right]}{E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]}$$

Table A1 Population and number of universities and colleges in 1998 in different provinces

Province	# of college	population	Province	# of college	population
Beijing	63	1240	Henan	51	9243
Tianjin	20	953	Hubei	54	5873
Hebei	46	6525	Hunan	47	6465
Shanxi	23	3141	Guangdong	43	7051
Inner Mongolia	19	2326	Guangxi	28	4633
Liaoning	61	4138	Hainan	5	743
Jilin	41	2628	Chongqing	22	3042
Heilongjiang	38	3751	Sichuan	43	8430
Shanghai	40	1457	Guizhou	20	3606
Jiangsu	66	7148	Yunnan	26	4094
Zhejiang	32	4435	Tibet	4	248
Anhui	34	6127	Sha'anxi	42	3570
Fujian	29	3282	Gansu	17	2494
Jiangxi	31	4150	Qinghai	6	496
Shandong	49	8785	Ningxia	5	530
			Xinjiang	17	1718