China's Higher Education Expansion and its Labor Market Consequences

Very preliminary draft!!
September, 2009

Shi LI
Beijing Normal University
and IZA

Chunbing XING
Beijing Normal University

Abstract

In the late 1990s, the number of students in tertiary education increased dramatically in China. Using a random 1/5 subsample of the 2005 1% population census, we find that high school graduate from different regions, ethnicities, and family backgrounds benefit from this policy differently. The compliers are less likely to be minority female, they are less likely to come from families with more than one child, and they are less likely to come from central-western families or that the college graduates are less likely to return to central-western regions after graduation. We also find some evidence on the short term effect of the education expansion policy. With the expansion of higher education, the employment pressure of new graduates increased. Those unemployed immediately after graduation and those relying on their families when unemployed increased.

China's Higher Education Expansion and its Labor Market Consequences

1. Introduction

Higher education is of great importance for both individuals and the whole nations. For individuals, it's related with significantly higher income and job stability. In the last decades, with the deepening of technological progress, the demand for high skilled labor increased dramatically. This further increases the price of college graduate both in China and all over the world (see Acemoglu, 2002, and Autor et al, 2005 for example; and see Heckman and Li, 2004 for the case of China). In a transition economy like China, the labor market is often segmented in terms of geography, ownership, and occupation. Higher education is often regarded as an essential channel through which individuals move upward.[1]

As a developing and transitional country, the share of individuals having tertiary education is relatively low. After China initiated the "reform and opening up" policy, the number of college graduate increased gradually. In particular, enhancing higher education is viewed as essential for the country's development strategy. Despite the gradual increase, China didn't experience significant increase in the number of college graduate until the late 1990s. The most remarkable change happened in 1999. The number of new student entering tertiary education was significant enlarged in 1999 and in subsequent years. The higher education expansion policy is so significant that it signaled the transformation of China's higher education system (Li, et al, 2008).

The most direct consequence of this policy is that more high school graduates are admitted to pursue tertiary education. The questions we want to answer in this paper are twofold. First, who benefit from this policy? Second, what's the labor market consequence of this expansion?

To answer these questions, we regard the expansion policy as a natural experiment. Individuals taking college entrance exams before and after the expansion policy have different probabilities of being admitted to college. And at the same time, the *before* and *after* group are largely determined based on their age, which can be regarded as exogenous. 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. Only part of the high school graduates is really affected. Therefore, we used a LATE framework developed by Imbens and Angrist (1994).[2]

Using a random 1/5 subsample of the 2005 1% population census, we find that high school graduate from different regions, ethnicities, and family backgrounds benefit from this policy differently. We also find some evidence on the short term effect of the education expansion policy.

---

[1] For example, the Deng and Gustafsson (2005) and Xing (2009) point out the higher education is a formal channel for rural residents to obtain official urban hukou.

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

In section 2, we introduce the background of the higher education expansion policy. Section 3 introduces the data. Section 4 investigates the policy effects on education. In section 5, we count and characterize the compliers of the policy using the LATE framework. In section 6, we investigate the labor market consequences of the policy. Section 7 concludes.

2. Higher Education Expansion in Late 1990s

Ever since the initiation of Reform and Opening Up policy, China's higher education made significant progress. In the 1980s and 1990s, the education regime especially higher education regime evolved continuously. The number of new college students and college graduate also increased continuously. From 1978 to 1998, the number of higher education establishments increased from 598 to 1022, the number of new college students increased from 0.4 million to 1.08 million, and the number of college students in school increased from 0.86 million to 3.41 million. However, although the scale of higher education enlarged continuously (both in terms of number of establishments and in terms of the number of students) the growth rate is dwarfed by those in 1999 and in subsequent years.

In 1999, the central government made the strategic decision of "expand the scale of higher education". 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 to entering tertiary education. This decision made 1999 a special year in the history of China's higher education. The number of new college students experienced the largest increase ever since (48%).[3] 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 "suddenness" of the announcement made this policy more like an experiments.

In the subsequent years, the number of new college students kept increasing. In 2005, the number of new college students was 5.04 million, 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 enrollment rate of higher education increased by 11.2%, reaching 21%.

The initiation of higher education expansion in the late 1990s is also closely related to the economics 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 the 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.

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

The effect of the expansion policy starting from 1999 is profound. Until now, the policy is still hotly debated. No doubt, the policy increased the probability of pursuing higher education. At the same time, this policy increased the relative supply of skilled workers. This in turn makes the employment problem of the newly graduate college students an important issue nowadays.

It's extremely difficult to evaluate the expansion policy. But we realize that the expansion from 1999 is like an experiment. Take two groups of high school graduate for example: those taking college entrance exams in 1998 and those in 1999. Students in these two groups will have different probability of entering college simply because they take the exams in different years because they are of different ages. But the effect would be different for different people. For high school graduates with both extremely high and extremely low ability, the policy will hardly have any effect on them. Whom the policy really affected were those lying in the middle of the ability distribution. They may differ from others in terms of sex, ethnicity, family backgrounds and geography. Is it true that the policy favor those in central and western regions, those of minority ethnicity as claimed by the policy makers?

From the methodological point of view, we used the framework of Local Average Treatment Effect (LATE) developed by Imbens and Angrist (1994). The expansion policy is used as an instrument. It affects the probability of people having higher education, and affects people's labor market outcomes only through their education decisions.

This experiment is not perfect however. The strategy we use is a Before-After comparison method. If there is some time varying unobservables affecting both the education decisions and people's labor market outcomes, the estimates from the before-after method will not be consistent. To overcome this shortcoming, we controlled the time trend of enrollment rate of higher education when using observations covering different ages. An alternative strategy is to keep only those observations around the policy year. This makes the following assumption more reasonable: people taking exams before and after the policy are not different systematically.

There is the possibility that other event (or economic conditions) affects the higher education expansion policy. Another problem is that the policy may change the timing of people's higher education. If people know the expansion policy in advance, some of them may postpone their exams to the next year. We do not pretend that this possibility is nil. However, the expansion policy (especially that in 1999) was largely unexpected, this minimized the problem of rearranging.

3. Data

The data we use is a one-fifth random draw from the 1% census data of China administered 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 the people enter the primary school at age 6, and they go to college at age 18. The theoretical year for them to take the college entrance exams is between 1988 and 2001. The fact is that 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 to do this. Table 1 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.

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	

4. The Effect of Education Expansion on Education

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 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. But we have 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 as depicted in the left panel. The increasing trend stopped at 1998. The above possible reasons are applicable here also.

The two panels in the second row of figure 1 show the number and share of professional college (3-year college). As for the absolute numbers, males and females have different trend. The number of male professional college graduates decreased, whereas the number of 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).

The change in the number of (4-year) college graduate 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 jumps are also clear in terms of proportion of college graduate. We also described the number and share of those with master degrees. There is no clear trend. This may due to the fact that the number of masters in our sample is relatively small.

To see the effect of the expansion policy on higher education more rigorously, we run the following regressions: [4]

$$College_i = \beta + \delta * t + \sum_{t=1998}^{2001} \gamma_t Cohort_{ti} + \varepsilon_i$$

College is a dummy indicating whether someone has college degree. We controlled time t to control the trend of higher education. The variable we are interested is *cohort*, it's used as indicators whether the observations belong to the cohorts that take college exams in 1998 (1999/2000/2001). Under the assumption that the probability of entering college is linear with age, the effect of expansion policy is captured by the coefficients of *cohort*. We also include a cohort dummy for 1998 to see whether there is abnormal change before the policy.

[4] As in the figure 1, observations with below middle school education are dropped.

Table 2 The Effect of Expansion on Education

	3 year college		4 year College		Graduate		College (3 or 4 years)		4 year College	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A: Model										
T	0.001	0.003***	0.001	0.002***	-0.000	0.001***	0.001*	0.005***	0.001	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Cohort98	-0.012*	-0.008	0.009	0.004	-0.003	-0.000	-0.003	-0.004	0.008	0.004
	(0.007)	(0.007)	(0.006)	(0.006)	(0.002)	(0.002)	(0.008)	(0.008)	(0.008)	(0.008)
Cohort99	-0.018**	-0.024**	0.030**	0.022***	-0.002	-0.002	0.012	-0.003	0.035**	0.024***
	(0.007)	(0.008)	(0.006)	(0.006)	(0.002)	(0.002)	(0.008)	(0.008)	(0.008)	(0.008)
Cohort00	-0.014*	-0.016**	0.040**	0.031***	-0.004**	-0.006**	0.026**	0.015*	0.049**	0.039***
	(0.008)	(0.008)	(0.006)	(0.006)	(0.002)	(0.002)	(0.008)	(0.009)	(0.008)	(0.008)
Cohort01	-0.015*	-0.015*	0.046**	0.055***	-0.007**	-0.011**	0.031**	0.040***	0.056**	0.073***
	(0.008)	(0.009)	(0.007)	(0.007)	(0.002)	(0.002)	(0.009)	(0.009)	(0.009)	(0.009)
F	2	6	41	70	8	8	14	55	37	83
r2_a	0.000	0.000	0.002	0.005	0.000	0.000	0.001	0.004	0.003	0.008
B: Model										
T	3.133**	2.318**	0.503	-0.893	0.377	0.160	3.636**	1.425	1.646	-0.679
	(0.988)	(1.080)	(0.812)	(0.840)	(0.276)	(0.283)	(1.110)	(1.180)	(1.079)	(1.152)
T2	-0.001**	-0.001**	-0.000	0.000	-0.000	-0.000	-0.001**	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cohort98	0.005	0.004	0.012	-0.001	-0.001	0.001	0.017*	0.003	0.017*	0.000
	(0.009)	(0.009)	(0.007)	(0.007)	(0.003)	(0.002)	(0.010)	(0.010)	(0.010)	(0.010)
Cohort99	0.008	-0.005	0.035**	0.014	0.001	-0.000	0.043**	0.009	0.050**	0.018
	(0.011)	(0.012)	(0.009)	(0.009)	(0.003)	(0.003)	(0.013)	(0.013)	(0.012)	(0.013)
Cohort00	0.024*	0.012	0.046**	0.020*	0.000	-0.004	0.070**	0.032*	0.069**	0.031*
	(0.014)	(0.015)	(0.012)	(0.012)	(0.004)	(0.004)	(0.016)	(0.016)	(0.015)	(0.016)
Cohort01	0.035**	0.022	0.055**	0.041***	-0.001	-0.008*	0.090**	0.063***	0.083**	0.062***
	(0.018)	(0.019)	(0.015)	(0.015)	(0.005)	(0.005)	(0.020)	(0.021)	(0.020)	(0.020)
F	3	6	34	59	7	6	13	46	31	69
r2_a	0.000	0.000	0.002	0.005	0.000	0.000	0.001	0.004	0.003	0.008
N	80835	75582	80835	75582	80835	75582	80835	75582	58680	52724

Note: ***, ** and * represent the significance level of 1%, 5% and 10%; standard errors are in parenthesis.

The results are reported in table 2. For females, the probability of entering college increased continuously. There is no significant trend for males. For whatever the types of college (3 year college, 4 year college or master graduate), the 1998 cohort dummy is not significant. Consistent with figure 1, the cohorts that should have taken college exams from 1999 to 2001 have significantly higher probability of being admitted to college. The probability increased by 2-6%. In contrast, the expansion policy has no positive effect on the probability of being admitted to other tertiary educations. Especially for the professional degrees, the 1999 cohort has significantly lower probability than the time trend. The result is worrisome that the increase in the number of 4 year college is simply the results of structural change between different types of tertiary educations. In column 7 and 8, we put the 3 year college and 4 year college into one single category. The 1999 cohort dummy is not significant any more. The 2000 and 2001 cohorts are still

significantly above zero.

Panel A of table 2 relies on the assumption that the probability of entering college is linear with ages (cohort). In panel B, we assume nonlinear trends. We add a quadratic term in the model. It seems that nonlinear trend is more appropriate to describe the trend for professional education. For both females and males, both the first and quadratic terms are significant. There is sign that the admission rate first increase and then decrease. The turning point occurs before 1999, and this is consistent with the results in figure 1. Under the nonlinear trend assumption, the coefficients of cohort dummies are no longer significantly negative. For males, the policy has positive effects. Adding the quadratic term doesn't change the results much. For females however, the 1999 cohort dummy are no longer significant after adding the quadratic term. The 2000 cohort dummy is significant at 10% level, and the one for 2001 cohort dummy is significant.

It's difficult to choose between these two specifications of linear or nonlinear model. The fact may be that even without the expansion policy, different types of tertiary education would have different time trends. From the significance level and the explanatory power of the model, nonlinear trend is more appropriate to describe the time trend of profession education, while the linear one is more appropriate for college education. Finally, although there are differences between these two specifications, the message is clear. The expansion policy has significant impact on the college education entrance. Therefore in the last two columns, we delete the observations with professional degrees and master degrees, focusing our attention on high school graduate and college graduate. Compared with the results in column 3 and 4, the coefficients are more significant.

5. Counting and Characterizing Compliers

5.1 The 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 taken 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. In the subsequent analysis, we either drop those professional graduate and those with master degrees or pool them with the 4 year college graduate. 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:

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

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.

$[D_{1i}, D_{0i}] \perp\!\!\!\perp Z_i$ (A1: independence assumption)

It's debatable that whether A1 holds. However, assume A1 is a fairly good start point. Relative to A1 assumption, 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):

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

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).

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

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).

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

5.2 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 4% higher. 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 considerable errors, which will downward bias the proportion.

Table 3 the Share of Compliers

	Sample I: All with above high school degree		Sample II High School and college graduate	
	Male (1)	Female (2)	Male (3)	Female (4)
A: 96/97/98(Z=0) vs 99/00/01(Z=1)				
E[D|Z=1]	18.66	18.43	25.28	26.53
E[D|Z=0]	14.80	14.47	20.34	20.98
P[Z=1]	51.39	52.21	51.75	52.38
P[D=1]	16.79	16.53	22.90	23.89
P[D1>D0]	3.86	3.96	4.94	5.55
P[D1>D0|D=1]	11.81	12.51	11.16	12.17
B: 97/98(Z=0) vs 99/00(Z=1)				
E[D|Z=1]	18.28	17.41	24.79	25.03
E[D|Z=0]	15.09	14.42	20.61	20.92
P[Z=1]	52.03	52.38	52.19	52.61
P[D=1]	16.75	15.98	22.79	23.08
P[D1>D0]	3.19	2.99	4.18	4.11
P[D1>D0|D=1]	9.91	9.80	9.57	9.37
C: 98(Z=0) vs 99(Z=1)				
E[D|Z=1]	17.75	16.78	24.02	23.98
E[D|Z=0]	15.53	14.83	21.17	21.59
P[Z=1]	51.54	50.76	51.71	51.24
P[D=1]	16.67	15.82	22.65	22.82
P[D1>D0]	2.22	1.95	2.85	2.39
P[D1>D0|D=1]	6.86	6.26	6.51	5.37

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.

Table 4 Characterizing compliers

Characteristics (X)	**female=1**	**Minority=1**		**Eastern=1**		**Single child=1**	
		Male	**Female**	**Male**	**Female**	**Male**	**Female**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample I							
96/97/98(Z=0) vs 99/00/01(Z=1)							
E[D|Z=0,x=1]	14.47	11.55	15.60	16.67	16.11	21.92	22.34
E[D|Z=1,x=1]	18.43	15.39	17.81	20.75	20.84	28.07	29.44
E[D|Z=0]	14.64	14.80	14.47	14.80	14.47	14.80	14.47
E[D|Z=1]	18.54	18.66	18.43	18.66	18.43	18.66	18.43
$P[x=1\|D_{1i}>D_{0i}]/ P[x=1]$	1.02	0.99	0.56	1.06	1.19	1.59	1.79
Sample II							
96/97/98(Z=0) vs 99/00/01(Z=1)							
E[D|Z=0,x=1]	20.98	16.52	23.54	22.78	23.11	32.00	34.44
E[D|Z=1,x=1]	26.53	21.32	26.80	28.18	30.25	40.50	45.38
E[D|Z=0]	20.65	20.34	20.98	20.34	20.98	20.34	20.98
E[D|Z=1]	25.90	25.28	26.53	25.28	26.53	25.28	26.53
$P[x=1\|D_{1i}>D_{0i}]/ P[x=1]$	1.06	0.97	0.59	1.09	1.29	1.72	1.97

Note: Sample I use all the observations with degree above high school. For sample II, those with professional degrees or master degrees are dropped.

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, central-western families, 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, the high school graduate from these families tend to face with credit constraint. Therefore even with the expansion policy they are unable to go to college.

6. The Labor Market Consequences of the Expansion

Although the expansion policy was initiated in 1999, the data in 2005 are still not appropriate for us to see the labor market consequences of the policy. As indicated by the first column of table 5, some of the observations are still in school in 2005, especially for the cohorts that should take college exams between 1999 and 2001. If college education takes four years, they would have graduated in 2005. However, some of them may go to graduate school, and some others may go to primary school at older age, and still there are other possibilities such as repeated education. Unfortunately, what we can do is to *calculate* the year of taking college exams according to a specific timing of education. Next, we investigate the short term effect of the expansion policy.

We delete those who are still in school. In the short term the labor market outcome for the cohorts that taking 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 of "unemployed immediately after graduation" for these cohorts. They are also more likely to be dependent on their family as the main source of income. The reasons are twofold, one is policy effect and the other is age effect[5]. As for age effect, it's natural that younger cohort have lower participation rate and depend 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 recruitment 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 will make them well-off. But we need to wait for better data now.

[5] There is also cohort effect. We can thing of pooling the cohort effect with age effect.

Table4 the Economic Consequences of Expansion

Age groups (the year of taking exams)	In school	Never work		Main source of income			Whether looking for a job within last 3 month				
			Unemployed After graduation	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

7. Conclusion

The expansion of higher education in the late 1990s is dramatic. How they affected the education decision of individuals? We find that a significant share of individuals go to have tertiary education because of the expansion policy. These compliers are systematically different from the whole sample. They are less likely to be minority female, they are less likely to come from families with more than one child, and they are less likely to come from central-western families or that the college graduates are less likely to return to central-western regions after graduation.

These results have strong policy implications. Although the expansion policy is aimed at increasing the probability of having tertiary education (especially those in the less developed areas), those from the less developed regions seemed to benefit less from it. One reason that they didn't benefit may due to credit constraint. Although the threshold in terms of test scores was lowered due to the expansion policy, other cost (tuition in particular) increased at the same time. The later may be a huge financial burden for many rural households. Even without financial credit constraint, high school graduates coming from less developed regions are more likely to be at the lower positions of the test score distributions, making them less likely to benefit from the policy. Therefore, alleviating the credit constraint should have been one of the policy priorities.

We also find some negative effect on the labor market consequences of the expansion policy. With the expansion of higher education, the employment pressure of new graduates increased. Those unemployed immediately after graduation and those relying on their families when unemployed increased. To see longer term effect, we need to wait for more data.

References

Acemoglu, Daron, “Technical Change, Inequality, and the Labor Market”, *Journal of Economic Literature* 40, March 2002, 7-72

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

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

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

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.

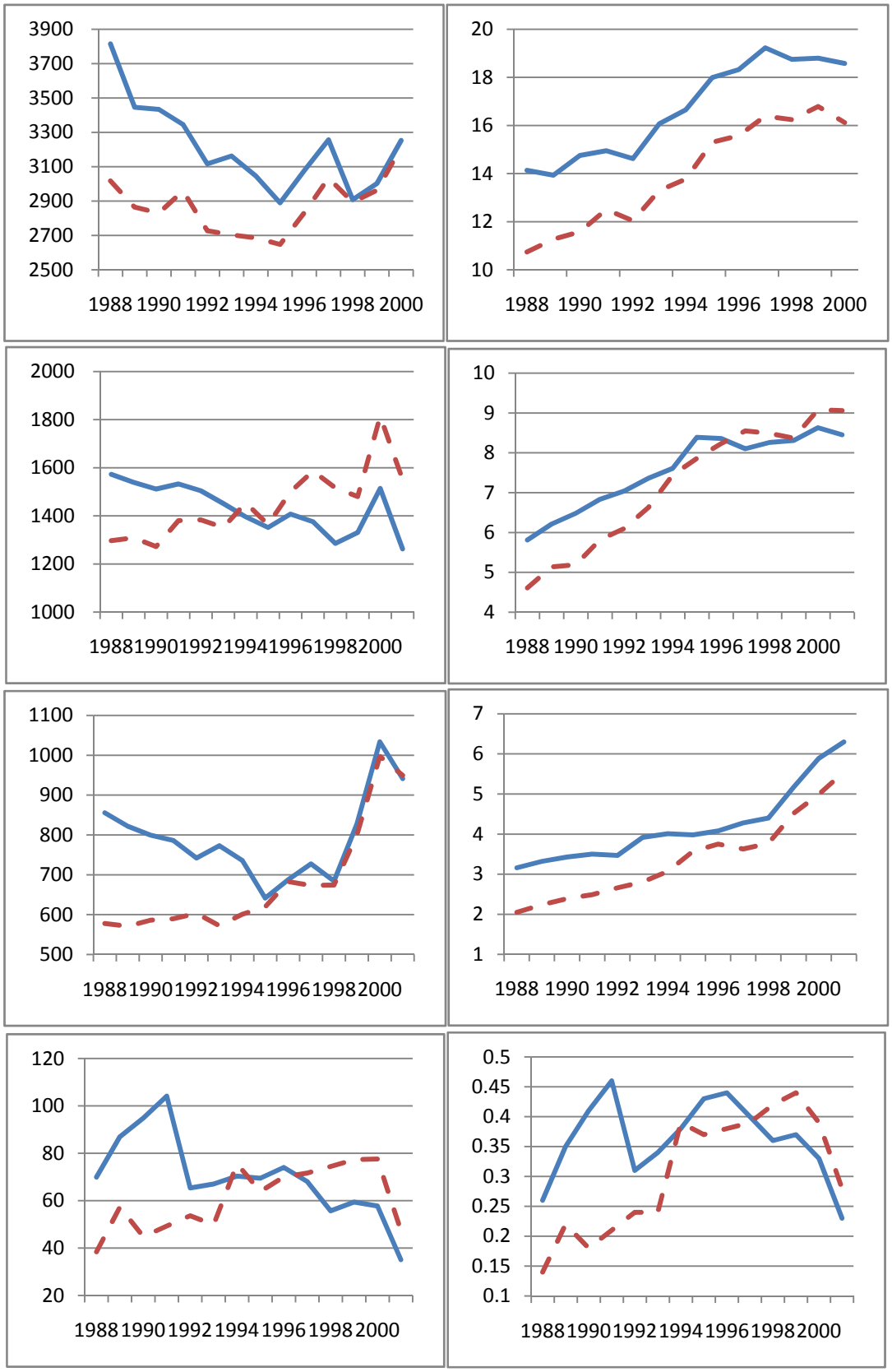
Meng, Xin, and Robert Gregory, 2007, “Exploring the Impact of Interrupted Education on Earnings: The Educational Cost of the Chinese Cultural Revolution”, Discussion Paper No. 2548

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

林莉，2008，《化解高校债务危机：途径与问题》，载于《2008 年中国教育蓝皮书》。

TBA…

Figure 1Number and proportion of different educations by age



Note: The first, second, third and fourth row are for high school graduate, professional graduate, college graduate and masters respectively. The left panel represents the number of observations, while the right one is the proportion.

Appendix A:

1. Calculating the share of compliers in the sample:

$$
\begin{aligned}
P\left[D_{1i} > D_{0i}\right] &= E\left[D_{1i} - D_{0i}\right] && \text{(monotonicity)} \\
&= E\left[D_{1i}\right] - E\left[D_{0i}\right] \\
&= E\left[D_{1i} \mid Z_i = 1\right] - E\left[D_{0i} \mid Z_i = 0\right] && \text{(independence)} \\
&= E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]
\end{aligned}
$$

2. Calculating the share of compliers in the college graduates:

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

3. Characterizing the compliers:

$$
\begin{aligned}
\frac{P\left[x_{1i} = 1 \mid D_{1i} > D_{0i}\right]}{P\left[x_{1i} = 1\right]} &= \frac{P\left[D_{1i} > D_{0i} \mid x_{1i} = 1\right]}{P\left[D_{1i} > D_{0i}\right]} \\
&= \frac{E\left[D_i \mid Z_i = 1, x_{1i} = 1\right] - E\left[D_i \mid Z_i = 0, x_{1i} = 1\right]}{E\left[D_i \mid Z_i = 1\right] - E\left[D_i \mid Z_i = 0\right]}
\end{aligned}
$$