HUMAN CAPITAL IN DEVELOPING COUNTRIES

by

Xinzheng Shi

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Economics) in The University of Michigan 2009

Doctoral Committee:

Professor David A. Lam, Co-Chair
Associate Professor Albert F. Park, Co-Chair
Professor Brian A. Jacob
Assistant Professor Dean C. Yang

To My Family

Acknowledgements

I would like to thank all the members in my family, who always support me, although most of them have never known what I have been doing. I also would like to thank Yao Lu for her persistent support.

I would like to thank Professor Albert Park. I met Albert when I was a master student in China. Without Albert, I could not study in The University of Michigan, and would not become a student (and a scholar in the future) focusing on development economics. In every stage of my Ph.D. study, Albert provided great help. Especially, when I started to write my dissertation, Albert encouraged me to participate in the field survey organized by him in Gansu Province in China. This survey not only provided me with a good data set for my dissertation, but also provided me with an exceptional experience of doing field research. Albert is always a person I can turn to when I have any questions.

I would like to thank Professor David Lam. Although I got to know David very late in my Ph.D. study (actually in my fifth year), David serves as a co-chair (with Albert) of my dissertation committee and helped me a lot especially after Albert left Michigan.

I would like to thank Professor Dean Yang. I started to work with Dean in the summer of my second year on a project, of which I became a co-author although I was mostly playing a role as a research assistant. Dean is the most energetic person I have met; he is doing empirical research very carefully. I have learned a lot about how to conduct high-quality empirical research from working and talking with Dean.

I would like to thank Professor Brian Jacob. Brian's comments on my papers helped to improve the quality of my dissertation a lot.

I would like to thank all other people, not mentioned in the above because of space limit, who helped and supported me in different ways.

Table of Contents

List of Figures

Figure

List of Tables

Table

Abstract

My dissertation investigates human capital issues, including education and health, in China. In the first chapter, I test for evidence of an intra-household flypaper effect by evaluating the impact of an educational fee reduction reform in rural China on different categories of household expenditures. Using data from Gansu Province in China, I find that educational fee reductions were matched by increased voluntary educational spending on the same children receiving fee reductions, providing strong evidence of an intra-household flypaper effect. The second chapter investigates the long-term effects of China's 1959-1961 famine. Using China's 2000 population census data, I find that women affected by the famine in the first year of life were living in smaller houses, achieved lower level of education, and provided less labor in their adulthood. But there are no long term effects on men affected by the famine in their early years of life. In the third chapter, I investigate the impact of school quality on students' educational attainment using a regression discontinuity research design that compares students just above and below entrance examination score thresholds that strictly determine admission to the best high schools in China's rural counties. Using data from Gansu Province in China, I find that attending the best high school in one's county of residence decreases the probability to take college entrance examination; increases college entrance scores and the probability of entering college.

Chapter 1
Introduction

My dissertation, “Human Capital in Developing Countries”, includes three separate papers.

Chapter 2 is the first paper which is titled “Does an intra-household flypaper effect exist? Evidence from the educational fee reduction reform in rural China”. In this chapter, I test for the existence of an intra-household flypaper effect by estimating the impact of the educational fee reduction reform in rural China on household expenditures.

Chapter 3 is the second paper which is titled “Famine, fertility and fortune in China”. In this chapter, I investigate the effects of the 1959-1961 Great Famine in China on the long term social and economic outcomes of those cohorts affected by the famine in the early years of life.

Chapter 4 is the third paper which is titled “Does school quality matter? Evidence from a natural experiment in rural China”. This chapter is based on a co-authored paper with Albert Park, Chang-tai Hsieh and Xuehui An. In this chapter, I investigate the effects of school quality on students’ academic performance in high schools by using the magnet school system in rural China as a natural experiment.

Chapter 5 is the conclusion.

Chapter 2

Does an intra-household flypaper effect exist? Evidence from the educational fee reduction reform in rural China

2.1 Introduction

Government transfers in developing countries often address the challenge posed by poverty and low levels of educational attainment among vulnerable populations, especially children. An important question in considering the effectiveness of such transfers is the extent to which transfers targeting to a child "stick" to him or her or, as suggested by many models of household decision-making (e.g., Becker, 1974), targeted individual transfers are equivalent to an increase in total household resources and are neutralized by reallocations of resources away from the target child toward other family members. The existence of an intra-household flypaper effect (IFE) thus provides an essential justification for policies targeting children. Indeed, Becker (1981) attributes the failure of compensatory education programs for minority children to the absence of a strong IFE.

In recent years, conditional cash transfer (CCT) programs like PROGRESA in Mexico have become extremely popular in developing countries.[1] Cash transfers made conditional on child enrolment can be thought of as a subsidy that reduces the price of education for each targeted child. Since the subsidy reduces household's expenditures on the targeted child, the question arises as to how the saved funds are spent, in particular whether the funds benefit (or stick to) the targeted child or whether they act similarly to a generic household income transfer. However, perhaps partly due to data limitations, no previous studies have examined this issue systematically.[2] Rather, evaluations of CCT

[1] For example, such programs exist in Bangladesh, Pakistan, Argentina, Chile, Colombia, Brazil, Mexico, Nicaragua, and Honduras.

[2] Some research focuses on the implications for intra-household allocation of the fact that the funds are transferred to women, not men. See Attanasio and Lechene (2002), Rubalcava, Teruel and Thomas (2006) and Bobonis (2004).

programs have focused on the programs' impact on educational attainment[3], health and nutrition[4], work[5], migration[6], fertility[7], and spillover effects among households in the same community[8]. If there is an IFE for conditional cash transfers, then the programs could benefit targeted children in ways that go well beyond the impact of the program on school enrolment[9]. Testing for the existence of the IFE also yields insights into the nature of intra-household decision-making in developing countries.

In this paper, I test for the first time whether an intra-household flypaper effect exists for child-targeted transfers by investigating the impact on household expenditures of an educational fee reduction reform that took place in rural China in 2005. As part of the large-scale program called the Two Exemptions One Subsidy (TEOS) program, students enrolled in primary and middle schools were exempted from school fees charged by schools, and students from poor families were exempted from textbook charges and received living subsidies if they lived in school dormitories. Since only enrolled children received any benefits from the program, the program was equivalent to conditional (on enrolment) cash transfer programs widely implemented in many developing countries.

A reform that reduces educational fees for children essentially reduces the price of education. For those households not sending children to schools before the reform because of the high educational fees, the combination of price and income effects from the reform may induce parents to send more of their children to school. However, for those households already sending most or all of their children to school before the reform, the reform has no effect on children's school enrolment. In China, primary and middle school enrolment rates were already high before the fee reduction reform, and a main motivation for the reform was to reduce the financial burden of taxes and fees for rural residents. Indeed, according to Park and Shi (2008), the estimated effect of the reform on primary school enrolment is close to zero and not statistically significant, and the

---

[3] Schultz (2004), Behrman and Sengupta (2005), de Janvry et al. (2006), Schady and Araujo (2006), Khandker, Pitt and Fuwa (2003), Filmer and Schady (2006) and Maluccio and Flores (2004).
[4] Behrman and Hoddinott (2005), Rivera et al. (2004), Gertler (2004) and Morris et al. (2004).
[5] Parker and Skoufias (2000, 2001), Gertler, Martinez and Rubio-Codina (2006), Schady and Araujo (2006) and Yap, Sedlacek and Orazem (2002).
[6] Angelucci (2004) and Stecklov et al. (2005).
[7] Schultz (2004), Stecklov et al. (2006) and Todd and Wolpin (2006).
[8] Bobonis and Finan (2005) and Angelucci and De Giorgi (2006).
[9] Park and Shi (2008).

estimated effect of the reform on middle school enrolment is statistically significant but very small[10]. Thus, for most households, with reduced educational fees spent on enrolled children, parents had more money for other expenditures, including expenditures on voluntary educational goods and other non-educational goods. If the IFE exists, we would expect parents to spend the extra money still on the education and on the target child who benefits from the fee reduction[11].

In this paper, I test the IFE by estimating the impact of TEOS reform on individual-specific household expenditures, focusing in particular on how the reform affected voluntary educational expenditures (on supplies, tutoring, etc.) and required educational expenditures (primarily school and textbook fees) on targeted children and their siblings. Detailed information on household expenditures is required for conducting such an analysis. I use Chinese rural household data from the Gansu Survey of Children and Families (GSCF), a longitudinal study following multiple cohorts of children. I helped supervise the last wave of the survey in 2007 for which I designed questions to collect comprehensive information on access to the TEOS program and on household expenditures on food (23 categories), non-food consumption (17 categories), and educational expenditures on each child (both required and voluntary).

Our identification strategy exploits the fact that household surveys were conducted before and after the reform, there was variation across schools in the amount of fee reductions due mainly to differences in the amount of fees charged prior to the reform, and access to subsidies and the amount of subsidies varied with the age (or grade level) of children in the household. The GSCF collected three waves of data in 2000, 2004, and 2007. The first two waves of survey were conducted before the reform and the third wave was conducted after the reform. I compare the spending of households with similar age children before and after the reforms living in the same village, and see how differences in the amount of educational fee reductions across schools and for children in different grades affected household expenditure patterns. In this paper, I use *intended transfer* as the main treatment variable, which is calculated based on program rules.

[10] 10% increase in the reduced educational fees leads to only 0.4% increase in the probability to be enrolled in the middle school.

[11] In my paper, if the parents spend the money received by the targeted children still on these children, it is also defined as intra-household flypaper effect, which traditionally only focuses on the money sticking to some specific categories of expenditures.

Firstly, I check the impact of the reform on household level expenditures. I find that a one *Yuan* increase in the household intended transfers per capita from the reform leads to 0.613 *Yuan* decrease in household required educational expenditures per capita and a 0.651 *Yuan* increase in household voluntary educational expenditures per capita; the absolute values of these two coefficients are not significantly different. But there is no significant impact of the reform on household total income per capita, total expenditures per capita, or other specific categories of expenditures. And then, using a household fixed effect specification to examine differences in child-specific expenditures within the same households, I find that a one *Yuan* increase in individual intended transfers from the reform leads to a 0.448 *Yuan* decrease in individual required educational expenditures and a 0.519 *Yuan* increase in individual voluntary educational expenditures; again the absolute values of these two coefficients are not significantly different. The results provide strong evidence that an IFE exists. Parents spent the saved money from the fee reductions on voluntary educational expenditures on the same child. I also find that given the same amount of intended transfers households with better educated mothers spend more on education, and parents spend more on girls, older children and children enrolled in middle school.

The only previous literature that has studied the IFE in developing countries is a small empirical literature examining the impact of school feeding programs[12]. Jacoby (2002) studied the impact of school feeding programs in the Philippines on children's caloric intake, comparing children's caloric intake on schooling days and non-schooling days. He found that daily caloric intake rose roughly one-for-one with feeding program calories. Using a similar methodology, Afridi (2005) found that 49% to 100% of nutrients provided by a mandated school meal program in India “stuck” to the children receiving meals. While both studies provide support for the existence of the IFE, their results could be driven by specific features of food consumption and school feeding program, for example it may be difficult to substitute more consumption in one meal with less consumption at other times, or home food availability may be related to other factors

[12] There exists a much larger literature on the flypaper effect in public finance (see Hines and Thaler (1995) for a review). More recent papers include Knight (2002), Choi, Laibson, and Madrian (2007),Walle and Mu (2007) , and Lalvani (2002). A number of papers in this literature find positive evidence that earmarked funds do increase spending on targeted areas.

affecting home production. In contrast, cash transfers or subsidies are fully monetized, providing more opportunity for passing on program benefits to other household members; in this sense they provide a purer test of the IFE.

One study asking a similar question to ours but in a developed country context is by Kooreman (2000), who showed that the marginal propensity to consume child clothing out of exogenous child benefits provided by a Dutch government program was much larger than the marginal propensity out of other income sources. However, the effects are identified solely from policy variation over time, so the results could be driven by time trends in expenditures that would have occurred even in the absence of the policy.

The rest of this chapter is organized as follows. The second section introduces the data and China's recent educational fee reduction reform in rural areas. The third section provides a simple conceptual framework for analysis. The fourth section describes empirical strategies. The fifth section presents the results. The sixth provides some robustness checks, followed by a concluding section.

2.2 Background

2.2.1 Gansu Survey of Children and Families

The data used in this paper was collected as part of Gansu Survey of Children and Families (GSCF). The GSCF is a longitudinal study which was conducted in Gansu province in the western part of China, one of the poorest provinces. In 2007, GDP per capita in Gansu was 9527 *Yuan*[13] (about 1389 dollars using the exchange rate on July 1st, 2008[14]), and the population was about 26 million[15]. The sample is representative of rural Gansu, excluding minority counties, and is drawn from 100 villages in 42 townships and 20 counties. Figure 2.1 shows the distribution of these 20 counties in Gansu province.

The GSCF follows a cohort of rural children aged 9-12 in the year 2000, when the first wave of the survey was conducted. The children and their families were re-interviewed in 2004 (wave 2) when the children were 13-16 years old, and in 2007

[13] This number is from http://tieba.baidu.com/f?kz=306047423.
[14] 1 Dollar = 6.86 *Yuan*. The exchange rate is from http://www.x-rates.com/d/CNY/table.html.
[15] The population number is from Gansu Bureau of statistics(http://210.72.51.4/doc/ShowArticle.asp?ArticleID=3408).

(wave 3) when they were 16-19 years old. The third wave also surveyed a new cohort of children aged 9-15 in 2007 and their families. The GSCF has a linked survey design which includes child, household, mother, father, homeroom teacher, school principal, and village leader questionnaires. The survey collected detailed information for target children and all siblings on enrolment status, grade level, and different types of educational expenditures, as well as on household income, expenditures, and wealth. Questions on income included a battery of questions on specific crops produced, livestock raised, self-employment income, wage income, prices, and inputs used in each type of production. The expenditure module consisted of 23 categories of food expenditures and 17 categories of non-food expenditures asked in each wave of the survey. Also, there are questions on over 40 different types of fixed capital and consumer durable goods. In the third wave, I designed a special section in the household questionnaire asking about fee exemptions and living subsidies received by each enrolled child in the household. School and county educational bureau questionnaires also described policy implementation during the recent period of educational fee reforms.[16]

In this paper, I use data on households with children enrolled in school. The main analysis focuses on the survey data for the years 2000 and 2007. For the sake of comparison, in each year only households having a sampled child aged from 9-12 years are used in the analysis. After dropping observations with missing values, the final sample includes 2134 households, 1629 in 2000 and 505 in 2007. In the sample, there are 4410 children (defined as survey target children and their siblings), 3498 in 2000 and 912 in 2007. Table 2.1 lists some basic characteristics of these households and children. From Panel A in Table 2.1, we can see that the average household size, number of children, and number of enrolled children are very close in the two years, but a little smaller in 2007. But the ratio of enrolled children to the total number of children is higher in 2007(0.890), than in 2000(0.830). We can also see that in 2007, mothers and fathers have more schooling years, and households have higher income and expenditures. The total enrolment rate of children in the sample increases from 85.8% in 2000 to 94.7% in 2007, and the enrolment rate of children aged 9-16 increased from 93.0% in 2000 to 98.2% in

[16] More detailed information about the GSCF is available at the project website: http://china.pop.upenn.edu/.

2007. In this paper, I restrict attention to enrolled children in the analysis. After observations with missing values are dropped, there are 3865 children remaining in the final sample, of which 3001 children are from the year 2000 data, and 864 children are from the year 2007 data. Table 2.1 also lists the characteristics of enrolled children. From the table, we can see that in both years, the average age and the ratio of girls in the sample are very close to each other. Among all enrolled children, the percentage of primary school students is a little lower in 2007 than in 2000; it is 81.1% in 2007 and 88.3% in 2000. However, the percentage of middle school students is a little higher; it is 16.4% in 2007 and 11.1% in 2000. The percentage of students enrolled in other levels of schools is 2.4% in 2007 and 0.6% in 2000.

2.2.2 Educational fee reduction reform in rural China

Under the decentralized fiscal system established after China's economic reforms began in 1979, compulsory education in China has been financed by local governments. Because of imbalanced economic development, poorer localities lacking local financial resource bases frequently have had no choice but to pass financing burdens down to farm households. The free compulsory education system espoused by official policy has never been achieved in practice, and educational fees in fact have often accounted for a large share of household expenditures, especially of the poor. Before recent reforms, individual schools enjoyed significant discretion in setting various fee levels for students. This system led to frequent complaints about excessive fee charging and heavy burdens placed on farmers.

In the past 5 years there have been a number of reforms to the system of rural educational finance in China. One of the main objectives of the recent reforms is to reduce the rural educational fees charged to students in order to reduce farmers' burdens. The focus of this paper is on the most recent of these reforms, known widely as the “Two Exemptions One Subsidy (TEOS)” policy, which targets children engaged in compulsory education, which in China includes primary school (typically grades 1-6) and lower secondary, or middle school (grades 7-9). According to government documents, TEOS was initiated nationally in the beginning of calendar year 2005 in all nationally designated poor counties (Ministry of Finance and Ministry of Education, 2005). The population of national poor counties accounts for about one fourth of China's total rural population. The policy was expanded to all of rural China in 2006, since which time it also has been called the new security system for rural compulsory education (*nongcun yiwu jiaoyu baozhan xin jizhe*)[17]. In this paper, I do not distinguish between these two names, and call the program TEOS throughout.

The policy's two exemptions refer to exemptions from paying school fees (*xueza fei*) and charges for textbooks (*keben fei*), which previously had been paid by students. The one subsidy refers to a living subsidy to partially cover the costs of students living in school dormitories. The three components differ somewhat in their coverage and financing. The exemption of school fees targeted all rural children, and funding for this exemption is shared proportionally by the central and local governments. The textbook charge exemption was targeted at the children of poor households only, and the central government took full responsibility for funding. The subsidy for living expenses was also targeted at the poor, but was financed by local governments who were asked to put their programs in place by the end of 2007. If a student was fully exempted from all three types of expenses (school fees, textbook fees, and dormitory living costs), based on field visits in Gansu the only remaining fee he or she would typically be asked to pay by the school would be to purchase notebooks.

In Tables 2.2-2.3, I report evidence on the timing of the implementation of the TEOS reform based on county and household questionnaires from the Gansu survey. Among the 20 counties surveyed, 15 percent reported implementing the school fees exemption in primary schools in the fall of 2004, another 50 percent reported starting in spring 2005, and the rest reported starting in spring 2006 (Table 2.2). The textbook fee exemptions began a little bit earlier, with 30 percent of counties starting in fall 2004, and another 45 percent in spring 2005. Just for the school fees exemption, by spring 2006 all counties had implemented the textbook fee exemption. The results are nearly identical for middle schools (Table 2.2). Implementation of the living subsidies was much slower, especially in primary schools. Only 10 percent of counties reported providing living subsidies in spring 2005, 65 percent of counties reported having started the program by spring 2006 when both exemption policies were fully implemented, and 30 percent of

[17] Guangming Daily (2007).

counties still had not begun providing subsidies by the time of the survey in summer 2007. Implementation was much faster in middle schools, which typically have more students living in dormitories. 60 percent of counties started providing living subsidies in spring 2005, and by spring 2006, 95 percent of counties had begun providing living subsidies.

Evidence from the household data is presented in Table 2.3. Because the coverage is based on recall, I restrict attention to students who have been in their current school since fall 2004 (primary students in grades 3-6 and middle school students in grade 9). Since the textbook fee exemption and living subsidies are targeted to children from poor households, it is possible that the coverage rates for students in the sample could be significantly lower than for schools or counties. In fact this is true for the living subsidy. No students report receiving living subsidies until spring 2005 and by spring 2007 only 0.92 percent of primary students and 8.96 percent of middle school students report receiving living subsidies (Table 2.3). In contrast, a very high proportion of students report receiving the textbook fee exemption--86 and 78 percent of primary and middle school students in spring 2007 (Table 2.3). The large difference in coverage of textbook exemptions and living subsidies likely is due to the fact that as a poor province Gansu receives a large amount of central subsidies for textbook exemptions, but lacks local resources to finance living subsidies (which are not centrally financed). Finally, a small percentage of primary school students (4 percent) and middle school students (9 percent) report not receiving the school fees exemption in 2007. This could be due to misreporting or attendance at private schools.

Prior to the reform, schools charged a single comprehensive educational fee, which included school fees, textbook charges and notebook fees. The value of the school fee exemption thus varied by county and school level, depending on the value of the single fee established prior to TEOS. I use the following formula to estimate the annual exempted school fees for each primary and middle school student:

$$SF_{i,2007} = \mathrm{Pr}e\_CEF_{s} - notebookfee_{i,2007} - textbookfee_{i,2007} \qquad (2.1)$$

Here, $\Pr e\_CEF_s$ is the comprehensive educational fee charged by school s before the reform[18], which is asked in the school questionnaire; the notebook charge is equal to 10 *Yuan* for primary students and 15 *Yuan* for middle school students; and the textbook fee is equal to 70 *Yuan* for primary school students and 140 *Yuan* for middle school students. The average value of estimated school fees exempted for primary school students is 54.26 with standard deviation equal to 54.66, and the average value of the estimated school fees exempted for middle school students is 71.75 with standard deviation equal to 56.13. Figure 2.2 shows the kernel density of exempted school fees over 100 villages. We can see that the exempted school fees are different for primary school students and middle school students, and they vary across villages.

According to the Gansu Educational Bureau, the subsidy provided to counties in 2007 for textbook purchases under TEOS was 35 *Yuan* per semester for primary school students and 70 *Yuan* per semester for middle school students. Since counties within the province buy the same textbooks, the value of the textbooks actually provided, and thus the value of the subsidy, is assumed to be uniform throughout the province. The criterion for allocating the free textbooks was the students' household income per capita[19]. Overall, 79.58 percent of primary school students and 72.69 percent of middle school students reported receiving textbook fee exemptions in the past year. The provincial education bureau recommended a living subsidy for poor students of one *Yuan* per day for 239 days per year, or 239 *Yuan* per year and 120 *Yuan* per semester (Gansu Educational Bureau, 2006). However, according to the 2007 survey data, only 0.92 and 8.96 percent of primary and middle school students received the subsidy.

Therefore, for each child enrolled in primary or middle school, total cost reductions from the reform, which I call *intended transfer*, are equal to the exempted school fees plus 70 or 140 if she is exempted from the textbook fee too; plus 239 if she gets the living subsidy. Table 2.4 shows the average value of intended transfers per capita

---

[18] In this paper, each village is assigned a nearest primary school, each township is assigned a nearest middle school. Students living in each village are assumed to attend the nearest school.

[19] But there was also variation in the number of free textbooks available to students in specific grade levels or schools, due to unexpected variation in class size or poor planning. Also, variation in the number of needy students across grades and schools could lead to differences in the extent to which the poorest students received the exemption. In our field research, we encountered a number of cases in which students received the textbook fee exemption in some semesters but not in others.

for households in the sample in 2007. It also summarizes descriptive statistics for household expenditures, including total expenditures per capita[20], food expenditures per capita[21], non-food expenditures per capita[22], health expenditures per capita[23], required educational expenditures per capita[24], and voluntary educational expenditures per capita[25]. From Table 2.4, we see that, for all households, the intended transfers account for about 2.971% of total expenditures. This figure is similar for the richest 30% of households (2.197%), but is greater for the poorest 30% of households (3.774%).

2.3 Conceptual framework

The impacts of the educational fee reduction reform can be analyzed under the framework of maximizing household utility. We can assume that parents are maximizing a household utility function U(X, E), where X is a vector of household consumptions, including consumptions of voluntary educational goods and other non-educational goods, and E is the number of enrolled children. I assume that U(X, E) satisfies: (1) $U_X > 0$, $U_E > 0$; (2) $U_{XX} < 0$, $U_{EE} < 0$; (3) $U_{XE} > 0$. And I also assume that both X and E are normal goods. Since the number of enrolled children cannot be arbitrarily large, we assume E cannot exceed E^*. E^* could be interpreted as the total number of children of schooling age in the family. I assume that household has income I, the vector of prices for X is denoted as P_X, and the price for E is denoted as P_E. P_E in this paper can be

---

[20] Total expenditures are equal to the sum of expenditures on food, expenditures on non-food consumption and service, expenditures on health care, expenditures on required educational items, and expenditures on voluntary educational items.

[21] Expenditures on food are the sum of expenditures on rice, flour, maize, beans, bean products, other grains, potatoes and sweet potatoes, grain products, corn, pork, beef and mutton, chicken, other meat, marine products, eggs, vegetables, edible oils, dried and fresh fruits and nuts, spices, cigarette and alcohol, soft drinks, canned food, restaurant food, and money spent on food proceeding.

[22] Expenditures on non-food consumption and service are the sum of expenditures on washing and cleaning supplies, miscellaneous household items and hardware, clothing consumption, bedding, transportation maintenance and parts, electronic appliances maintenance, transportation costs, postage, communication devices maintenance fees, rent, house decoration and renovation, electricity, fuel, cultural products, cultural service fee, personal goods (including jewelry, makeup, etc.), and personal service fees (including salon, bath, photos, etc.).

[23] Expenditures on health care are the sum of expenditures on buying medicine, and health insurance.

[24] Expenditures on required education items is the sum of tuition and textbook fee.

[25] Expenditures on voluntary educational items is the sum of expenditures on supplies of pens, exercise books, etc., supplementary tutoring, and snake in school and costs of transportation to school.

defined as required educational fees for each child. Then the household maximization problem is represented by the following:

$$\underset{X,E}{Max}\, U(X,E)$$

$$S.t. \qquad P_X * X + P_E * E \leq I$$

$$E \leq E^*$$

Figure 2.3 shows the optimal choice of households having different preferences. For type A households, parents care more about the education of their kids such that they send all their kids of schooling age to the schools. Therefore, the maximization problem has a corner solution, which is represented by A^0 in Figure 2.3. But for type B households, parents care less about the education of their kids; therefore, their maximization problem has an interior solution, which is represented by B^0 in Figure 2.3.

The educational fee reduction reform reduces P_E, which is shown in Figure 2.3 as budget line rotating rightward around point M. Since both X and E are normal goods, for type A households, the maximization problem still has a corner solution, which is represented by point A^1 in Figure 2.3. That is, the number of enrolled kids stays the same as before the reform, while the households can have more money to spend on other goods, which is represented by the higher level of other consumptions in Figure 2.3. On the contrary, for type B households, their optimal choice moves to B^1, at which more of their kids are enrolled in the schools and they are also able to consume more of other goods.

In this paper, my goal is not to study the impact of the reform on the children's school enrolment, which is studied in a separate paper (Park and Shi (2008)); therefore I take household's enrolment decision as given in the analysis. It is a reasonable assumption since the educational fee reduction reform was not found to have big effects on children's school enrolment. Actually, according to Park and Shi (2008), the primary school enrolment rate was 96.3% before the reform and 98.18% after the reform; and the middle school enrolment rate was 88.92% before the reform but 91.44% after the reform.

The increase of enrolment rates is small. Indeed, the estimated impact of the reform on primary school enrolment is close to zero and not statistically significant. Although the estimated impact of the reform on middle school enrolment is statistically significant, the magnitude is not large; 10% increase in money transferred from the reform only led to 0.4% increase in the probability of middle school enrolment. Besides, children of primary school age (6-12 years old) account for 72% of the sample but children of middle school age (13-16 years old) only account for 25% of the sample. It shows that it is reasonable to take household enrolment decision as given in the analysis. Of course, it might cause some potential bias in empirical exercises, the detail of which is discussed in section 2.4.1.

Based on this assumption, parents' maximization behavior can be divided into two steps. In the first step, they pay required educational fees for their enrolled children, the total amount of which is $P_E * E^*$; in the second step, parents choose the optimal value of X using the remaining money $I - P_E * E^*$ to maximize the utility. For the sake of analysis, I divide X into two parts: voluntary educational goods, E_V, and all other non-educational goods, $\bar{X}$. The household maximization problem can be rewritten as the following:

$$\underset{\bar{X}, E_V}{Max}\, U(\bar{X}, E_V, E^*)$$

$$S.t. \qquad P_{\bar{X}} * \bar{X} + P_V * E_V \le I - P_E * E^* \qquad (2.2)$$

Here, $P_{\bar{X}}$ is a vector of prices of non-educational goods and P_V is a vector of the prices of voluntary educational goods. We assume that the educational fee reduction reform reduces the required educational fees to $\bar{P}_E$; then the maximization problem becomes:

$$\underset{\bar{X}, E_V}{Max}\, U(\bar{X}, E_V, E^*)$$

$$S.t. \quad P_{\bar{X}} * \bar{X} + P_V * E_V \le I - P_E * E^* + (P_E - \bar{P}_E) * E^* \qquad (2.3)$$

Comparing (2.2) and (2.3) in the above, we can see that reducing the required educational fee does not change the relative prices between the non-educational goods and voluntary educational goods, but is equivalent to a cash transfer to households equal to $(P_E - \overline{P}_E)*E^*$. The other thing we can see is that the amount of the transfer depends on the price of education households pay before the reform, i.e. P_E. The higher is P_E, the greater is the implicit cash transfer.

Figure 2.4 shows the maximization problem described above. In Figure 2.4, voluntary educational goods are put on the X-axis, and all other non-educational consumptions are on the Y-axis. Since the educational fee reduction reform did not change the relative prices between the voluntary educational goods and non-educational goods, the reform shifts the budget line outward. Before the reform, households choose to consume A. After the budget line is shifted outward, optimal consumption will be at point B if households redistribute the money among all expenditures, but will be at point C if households continue to spend the extra money on education, which is defined as the IFE in this paper. One could similarly interpret the X-axis as measuring the non-educational goods purchased for the targeted children, and the non-educational goods for other enrolled children on the Y-axis.

The above discussion implies that, if the IFE exists, then the impact of the reform on household voluntary educational expenditures should be positively significant, and much greater than the impact on other household expenditures. And individual-specific voluntary educational expenditure should increase for the children receiving benefits from the reform.

One caveat needs mentioning. If household preference over voluntary educational goods and other non-educational goods has some special features, for example, the utility function is quasi-linear, then even the IFE does not exist, with the increase of the money available for voluntary educational goods and all other non-educational goods, we can still see that all the extra money is spent on voluntary educational goods. This alternative explanation is shown in Figure 2.5. However, we can test whether this type of household preference is consistent with the data. I regress expenditures on voluntary educational goods and expenditures on all non-educational goods separately on money available for purchasing them[26]. If the household utility function has the features described in the above, we should expect to see a much larger coefficient of voluntary educational expenditures than that of other non-educational expenditures. Table 2.A shows the regressions using household sample from year 2000. From this table, we can see that with the increase in the money available, the non-educational expenditures increase, which is shown in columns (2) and (4). But from columns (1) and (3), with the increase in the money available, the increase in voluntary educational expenditures is much smaller. These results show that it is impossible for the preference to have some special features such that the increase in the money available to be spent on voluntary educational goods and other non-educational goods only leads to the increase in the consumption of voluntary educational goods. Besides, we need to notice that quasi-linear preference plays no role in the determination of the implication of IFE that the saved money will be spent on the children receiving this money.

Another important question is how parents know how much they save from the reduced education fees and then spend it on children's voluntary educational items. Actually, there are several ways for parents to get this information. One is from schools' publicized information. In the field research, we did see schools posted the information about what kinds of items were exempted and how much was exempted. The parents can also know this information just from their own experience since most of them have children enrolled in school before 2005 when the TEOS reform started. Figure 2.A shows the distribution of the school entering year of the oldest children in the sample households. From Figure 2.A, we can see that more than 95% households had children enrolled before 2005. For the remaining households, they can also know this information from their neighborhoods.

After knowing the implications of the intra-household flypaper effect, one might ask what causes the IFE. To the best of my knowledge, there is no formal theory providing explanations for the IFE. However, one might image that parents have a fixed amount of money they would like to spend for each of their children's education, the

[26] At the same time, I control for household demographic structure, land area per capita, number of kids enrolled in different levels of schools, mother's education, father's education, and dummies for different villages. I use two measures for money available for purchasing voluntary educational goods and all non-educational goods; one is the sum of expenditures on these goods, and the second one is equal to household total income subtracting required educational expenditures.

reduction of required educational fees does not change this “habit formation”. Therefore, with the reduction in the required educational fees and the fixed amount of total money to be spent, we can observe the increase in the voluntary educational expenditures.

2.4 Empirical strategy

As discussed in the above, if the IFE exists, the impact of the reform on household voluntary educational expenditure should be significantly positive, and the impact of the reform on household other categories of non-educational expenditures should be much weaker. And the reform is also expected to have significantly positive impact on individual voluntary educational expenditures. In this section, I discuss empirical strategies used to identify the reform's impacts on expenditures in household level and in individual level.

2.4.1 Household level regressions

Before the reform, different schools in different villages charged different educational fees; therefore, when the reform removed the educational fees, households in villages having higher educational fees before the reform benefited more from the reform. We can compare educational expenditures of households, which were surveyed in 2007 when the reform had been implemented, in villages having higher pre-reform educational fees with those households in villages having lower pre-reform educational fees. However, cross-sectional comparisons cannot differentiate the impact of the reform from the village characteristics which might also be related with expenditures. Since GSCF also collected information in 2000; therefore, households surveyed in 2000 are used as a control group in this paper. In other words, I compare households in the same village but surveyed before the reform and after the reform, by which the fixed village characteristics can be canceled out; and then I compare the before-after difference in the expenditures of households in the villages having higher pre-reform educational fees with the before-after difference in expenditures of those living in the villages having lower pre-reform educational fees. The identification idea used here is a “continuous” version of difference-in-difference strategy.

A crucial identification assumption for different-in-difference strategy to be valid is that the change of educational expenditures from 2000 to 2007 in villages benefiting more from the reform should not be systematically different from those in villages benefiting less from the reform in the absence of the reform. I can test this assumption by using the first two waves (in years 2000 and 2004) of data. The identification will also be violated if there exist some time-varying and region-specific effects correlated with transfers (essentially the pre-reform educational fees) from the reform and potentially affect household educational expenditures. Unfortunately, the mechanism for schools in different villages to charge different fees before the reform is not clear. But, since GSCF collected detailed village level variables, which makes it possible to include a large number of village variables in year 2000 (interacted with year 2007 dummy) to control these time-varying and region specific effects. In the regression, I include illiterate ratio, ratio of primary school graduates in labor force, primary school enrolment rate, middle school enrolment rate, indicator for having pre-school classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises. Combined with village fixed effects, inclusion of these variables in regression function helps to correct bias due to possible endogeneity of the pre-reform educational fees charged in different villages.

Another concern is that there might be some other concurrent government programs correlated with the level of educational fees charged before the reform and potentially affected educational expenditures. Indeed, Chinese government exempted all agriculture-related taxes in 2006; this policy should not affect the estimation since it is not related with pre-reform educational fees and common for all the households in different villages. To the best of my knowledge, there are no other programs concurrent with the educational fee reduction reform in Gansu.

The following regression function is estimated:

$$EX^{m}_{ht} = \alpha_0 + \alpha_1 * ITPC_{ht} + \alpha_2 * H_{ht} + V_{v,2000} * Year_{2007} + Village_v + Year_t + \varepsilon_{ht} \quad (2.4)$$

In equation (2.4), EX^{m}_{ht} on the left hand side is a vector of outcome variables, including household income per capita, household total expenditures per capita, and specific categories of expenditures per capita, including food, non-food goods and service, health care, required educational expenditures and voluntary educational expenditures. $ITPC_{ht}$ is intended transfers per capita household h should get in year t from the reform. The value is equal to zero in pre-reform years. In the year 2007, it is equal to the sum of intended transfers to each enrolled child in the household divided by the number of family members[27]. H_{ht} is a vector of household variables, including household demographic structure[28], household endowments[29], log of household real income per capita, and the number of children enrolled in different grades[30]. Interactions of village variables in 2000 and dummy for year 2007, $V_{v,2000} * Year_{2007}$, are included in the regressions to control for time-varying and region-specific factors. I also control for village average expenditure levels and cohort average expenditure levels by including village fixed effects, $Village_v$, and year fixed effects, $Year_t$. ε_{ht} is an error term with mean equal to zero. In the estimation, robust standard errors are calculated by clustering at the village level.

---

[27] Since the eligibility for students to get textbook fee exemption and living subsidy mainly depends on household income, and the amount of exempted textbook charge and living subsidy are the same for all the kids eligible to get them. Therefore, by including household income in the regressions as a control, the variation from textbook fee exemption and living subsidy is absorbed. In other words, the variation in the variable of intended transfer comes from the variation in the extent of exempted school fees. I will test it in Section 6.1.

[28] Household demographic structure includes ratio of male from 0-5, 6-12, 13-16, 17-19, 20-29, 30-39, 40-49, 50-54, 55 and above, ratio of female from 0-5, 6-12, 13-16, 17-19, 20-29, 30-39, 40-49, 50-54, 55 and above, and total number of family members.

[29] Household endowments include land area per capita, father's schooling year, and mother's schooling year.

[30] This vector of variables includes number of children enrolled in primary school, middle school, high school, and other school levels.

In estimating functions (2.4), one important caveat bears mentioning. According to Park and Shi (2008), the fee reduction reform has no significant effects on primary school enrolment, but it does have significant positive effects on middle school enrolment. In this paper, I take the enrolment decision as given by using households having enrolled children as the sample for analysis and including the numbers of kids enrolled in different grades as control variables. The results shown in this paper should not be extended to interpret the behaviors of households without any children enrolled in schools. However, even when we interpret the results in this way; there are still ways in which the results could be biased. Households sending children (especially 13-16 years old children) to school only after the reform reduced educational fees might be less willing to invest in children's education due to unobservable characteristics that prevented them from enrolling their children in the absence of the reform. Then the estimates of the impact of the reform on voluntary educational expenditures would be biased downward. However, the bias should not be large because 9-12 years old children account for majority of the sample (67%)[31] and the impact of the reform on middle school enrolment is small as described above.

2.4.2 Individual level regressions

In order to identify the impact of the reform on individual level educational expenditures, I take advantage of two sources of variation. One is the variation in intended transfer received by children enrolled in different levels of grade but living in the same family. The other is differences in the magnitude of fee reductions in different villages. Children surveyed in year 2000 are used as a control group in both cases.

GSCF collected information on required and voluntary educational expenditures for each enrolled child, which makes it possible to exploit within-household variation to identify the impact of the reform on individual educational expenditures. Children living in the same family might get different transfers from the reform if they are enrolled in different grades. Children enrolled in high school receive no transfers, and children

[31] There are also 5% children aged from 6-8 years, although Park and Shi (2008) did not investigate the impact of the reform on them, the impact should not be significant since these children were only be likely to be enrolled in primary schools, and the impact of the reform on enrolment in primary school is not significant.

enrolled in middle school get more transfers from the reform than children enrolled in primary school. We therefore can compare educational expenditures on children living in the same family, but enrolled in different grades, to identify the impact of the reform on educational expenditures while controlling for household level unobservables. However, if just cross-sectional data in one year is used, the impact of the reform cannot be differentiated from the systematic difference in the educational expenditures on children enrolled in different grades in the absence of the reform. Fortunately, children surveyed in 2000 can be used as a control group. Intuitively, I compare educational expenditures differences between children enrolled in different grades but living in the same family before the reform with similar differences after the reform, a version of difference-in-difference strategy.

The above idea can be represented by the following specification:

$$EE_{iht} = \alpha_0 + \alpha_1 * IT_{iht} + \alpha_2 * X_{iht} + \gamma_{ht} + \varepsilon_{iht} \qquad (2.5)$$

In equation (2.5), EE_{iht} is a vector of child-specific expenditure variables, including educational expenditures (required and voluntary) of children i living in household h in year t; IT_{iht} is intended transfer received by child i in household h in year t, equal to zero in the year 2000; X_{iht} is a vector of individual characteristics, including sex and age; γ_{ht} is a household fixed effect for household h surveyed in year t. ε_{iht} is an error term with mean equal to zero. In the estimation, robust standard errors are calculated by clustering at the village level. In order to cancel out household fixed effects, I subtract household average values from every variable in equation (2.5), and then to compare before and after the reform within the same village, I add village fixed effects to the regression. Finally, the following equation is estimated:

$$EE_{iht} - \overline{EE_{ht}} = \alpha_1 * (IT_{iht} - \overline{IT_{ht}}) + \alpha_2 * (X_{iht} - \overline{X_{ht}}) + Village_v + (\varepsilon_{iht} - \overline{\varepsilon_{ht}})$$

(2.6)

By using the within-household variation to identify the effect of the reform, we can implicitly control any observable and unobservable village and household level variables. But the identification assumption should be that the systematic difference between educational expenditures on children enrolled in different levels of schools does not change from 2000 to 2007 in the absence of the reform. This assumption might be violated if the time path of educational expenditures on children enrolled in higher level school is different from the time path of educational expenditures on children enrolled in lower level school, or/and if parents' preference toward different children changes. I test this identification assumption by taking advantage of special survey design of the GSCF data. GSCF includes three waves of data, two (year 2000 and 2004) before the reform and one (year 2007) after the reform. Since children surveyed in year 2000 and 2004 were not affected by the reform, so I am able to do control experiment to test this identification assumption.

In order to exploit the variation in the extent of fee reductions in different villages, the following regression function is estimated:

$$EE_{iht} = \alpha_0 + \alpha_1 * IT_{iht} + \alpha_2 * IT_{-iht} + \alpha_3 * XX_{iht} + \alpha_4 * HH_{ht} + V_{v,2000} * Year_{2007} + Village_v + Year_t + \varepsilon_{iht} \quad (2.7)$$

In this function, EE_{iht} on the left hand side is defined the same as that in function (2.6), which is a vector of outcome variables, including required educational expenditures and voluntary educational expenditures. In order to control heterogeneity between treatment group and control group, I include a vector of individual characteristics, XX_{iht}, including dummy for girl, age, and dummies of grades; and a vector of household characteristics, HH_{ht}, which is a vector of household characteristics, including household demographic structure, household endowments, and log of household real income per capita. As discussed above, I include the interactions of village variables in 2000 and indicator for year 2007, $V_{v,2000} * Year_{2007}$, as control. I also control for average village and cohort education expenditure by including village fixed effect, $Village_v$, and year fixed effect, $Year_t$. In the function, ε_{iht} is an error term with mean value equal to zero. The most important variables in this regression function are IT_{iht} and IT_{-iht}. IT_{iht} is intended transfer child i in household h gets from the reform in year t. The coefficient before this variable, α_1, captures the impact of the reform on child i. IT_{-iht} is the sum of intended transfer of all other children living in the same family. The coefficient before this variable, α_2, captures spillover effect of intended transfer other children in the same family gets. In the estimation, robust standard error is calculated by clustering in village level. If the IFE exists, we should expect to see precisely estimated α_1 and insignificant α_2.

In order to estimate function (2.7), the identification assumptions needed are the same as those for estimating household regressions. The potential selection due to endogenous enrolment decision leads to the same bias in the estimation as it does in the household regressions. In addition, the estimated results might still be driven by some unobservable household or village changes in educational expenditures. But if so, then we should expect to see the same effects of IT_{iht} and IT_{-iht} on individual educational expenditures. Therefore, the estimation of function (2.7) itself can also provide evidence to see whether the unobservable household or village changes in educational expenditures affect the estimation.

2.5 Empirical results

2.5.1 Descriptive results

Table 2.5 describes the household expenditure patterns in 2000 and 2007. From this table, we can see that required educational expenditures per capita in 2007 is much lower than that in 2000, the difference of the average values in these two years is -71.349, and is significant by different from zero at the 1% level. This means that the educational fee reduction reform did reduce required payments for children's education. All other variables, including total expenditures, food expenditures, non-food expenditures, health expenditures, and voluntary educational expenditures, increased from 2000 to 2007, with the changes being statistically significant. But this increase might just be due to the existing time trends.

Next, I estimate a rough measurement of the impact of the reform by using the difference-in-difference strategy, and report the results in Table 2.6. In the table, the “treated” group includes villages having average intended money transferred from the reform above the median, and the “untreated” group includes villages having average intended money transferred below the median. Year 2000 is the pre-reform period and year 2007 is the post-reform period.

The table lists results for total expenditures per capita and other categories of expenditures per capita, including expenditures on food, expenditures on non-food consumption and services, expenditures on health care, required educational expenditures, and voluntary educational expenditures. From the table, we can see that compared with the untreated group, the expenditures of the treated group are always higher. One important reason might be because the treated group includes villages having higher intended transfers, i.e. higher pre-reform educational fees, which could be related to better economic status of these villages. Secondly, we can see that, compared with pre-reform period, expenditures in the post-reform period are greater, whether for the treated or untreated group. The difference-in-difference results are shown in the right-bottom cell
in each section in Table 2.6. We can see that the effects are statistically significant only for required educational expenditures per capita and voluntary educational expenditures per capita; both of them are significant at the 1% level. The difference-in-difference result for required educational expenditures per capita is -40.27, which means the average required educational expenditures per capita in the treated villages decreased by 40.27 *Yuan* more than in the untreated villages. The difference-in-difference result for voluntary educational expenditures per capita is 55.33, which means that the average voluntary educational expenditures per capita in treated village increased by 55.33 *Yuan* more than in the untreated villages. All the other results are not statistically significant. These simple comparisons provide preliminary evidence for the existence of the IFE, i.e. households used money saved from reduced required educational fees to spend on the voluntary educational items. Since it cannot control for other variables, this paper provides more reliable regression results in the following.

2.5.2 Household level regressions

Table 2.7 presents the results for the household level regressions. Section A in Table 2.7 shows the main results using data from waves of years 2000 and 2007, while results in section B are results from a control experiment using data from waves of years

2000 and 2004. For clarity of presentation, we just present the coefficients of the intended transfers per capita, and suppress reporting the coefficients on the many control variables. These full results are available upon request.

Column (1) of section A shows the impact of the fee reduction reform on household income per capita. The reform is expected to increase income if households use the money from the reform to make productive investments. However, the result shows that there is no significant impact of the reform on income, and the coefficient is negative. Column (2) shows the impact on total expenditures per capita. While the coefficient is -2.222, it is not significant either. Since household saving is equal to income minus total expenditures, we can derive from these two estimated results that there is also no impact of the reform on household savings, i.e. households do not save the money from the reduced educational fees.

Columns (3) to (7) show the impact of the reform on different categories of expenditures. Columns (3) to (7) correspond to expenditure on food, non-food goods and services, health care, required educational fees, and voluntary educational spending respectively. Firstly, from columns (3), (4) and (5), we can see that the impact of educational fee reduction reform on expenditures on food, non-food goods and service, and health care per capita are not statistically significant and are all negative. The coefficients are -0.014, -1.770 and -0.476 respectively. Only the impacts on required educational expenditures and voluntary educational expenditures are significant (see columns (6) and (7)). The coefficient of required educational expenditures is -0.613 and significant at the 1% level. This means that a one *Yuan* intended transfers per capita leads to a 0.613 *Yuan* reduction in required educational payments per capita. Since there were some other fee reductions before 2005, but intended transfers only capture the change in required educational fees due to the most recent reform starting in 2005; however, the dependent variable measures the change in required educational expenditures from 2000 to 2007, therefore, the estimated coefficient is not expected to be equal to -1. In addition, it is also possible that schools did not fully implement the policies or that parents under-reported required educational fees. However, what is relevant for testing the IFE is the comparison of the coefficient in columns (6) and (7). Column (7) reports the impact of the reform on household voluntary educational expenditures per capita. The coefficient shown in this column is 0.651, significant at the 5% level. Thus, one *Yuan* of intended transfers per capita from the reform leads to a 0.651 *Yuan* increase in voluntary educational expenditures per capita. We can see that although the magnitudes of coefficients in column (6) and column (7) are not exactly the same, they are very close to each other and have opposite signs. Indeed, the P-value of the Wald test testing the null hypothesis that the absolute values of these two coefficients are the same is 0.899, which means that the difference between the absolute values of these two coefficients is not significant. These results suggest that there are no income effects of the reform and that parents spend the money transferred from the reform on voluntary educational expenditures.

Section B shows the results from a control experiment testing the identification assumption that the time trends of household expenditures in villages benefiting more from the reform are not systematically different from villages benefiting less from the reform. The critical issue is how to generate hypothetical intended transfers for each enrolled child surveyed in year 2004. I calculate the mean value of intended transfers for children enrolled in each grade and living in each village in 2007, and then I assign this average value to children enrolled in the same grade and living in the same village in 2004. The households' hypothetical intended transfers per capita are equal to the sum of all such hypothetical intended transfers divided by the number of household members. All other control variables are defined to be the same as in section A. From section B, we can see that none of the coefficients are significant, showing that the time trends of households' expenditures in different villages are not systematically different in the absence of the reform.

From Table 2.7, we see that on average one *Yuan* in intended transfers per capita leads to a 0.651 *Yuan* increase in voluntary educational expenditures per capita. I also test the heterogeneous effects of the reform on voluntary educational expenditures. I investigate whether impacts differ by household income per capita, total number of enrolled children, mother's schooling years, and father's schooling years (columns (1) to (4) in Table 2.8). The coefficient on the interaction of log value of household income per capita and intended transfers per capita is positive (0.148); but not statistically significant(column (1)). The coefficient on the interaction term with number of enrolled children is 0.274, and also not statistically significant (column (2)). Columns (3) and (4) show the heterogeneous impact of the reform in terms of mother's and father's schooling years respectively. Although both coefficients are positive, only the coefficient on the interaction with mother's schooling years (0.083) is significant (at the 5% level). The coefficient of the interaction term with father's schooling years is smaller (0.037) and not statistically significant. With higher level of education, parents might know more about the importance of children's education, since mothers have been found to care more about children's education than fathers[32], therefore, with the same amount of educational fee reductions, mothers having higher education would spend more on children's education, but fathers having higher educational would not.

2.5.3 Individual level regressions

In this paper, I am exploiting variation of different educational fee reductions received by children enrolled in different grades in the same family to identify the impact of the reform on educational expenditures on individual children. This identification idea is summarized in equation 2.6. Table 2.9 presents the estimation results. This table is also divided into two sections in the same manner as Table 2.8. Columns (1) and (2) show the results from the estimation of equation (2.6), and columns (3) and (4) show the results from the estimation of equation (2.7). As before, we only report coefficient on the individual intended transfers and the coefficient on the sum of all other children's intended transfers in the same family.

We find that one *Yuan* increase in intended transfers leads to a 0.493 *Yuan* decrease in required educational expenditures(column (1) in section A) and a 0.578 *Yuan* increase in voluntary educational expenditures(column (2)). These two coefficients have opposite signs and similar magnitudes. Both are statistically significant at the 1% level. I test the null hypothesis that the absolute values of these two coefficients are equal. The P-value of the Wald test is 0.794, suggesting that there is no significant difference between the magnitudes of these two coefficients. Column (3) also shows the impact of the reform on individual required educational expenditures, estimated from regression function (2.7). The coefficient on the individual level intended transfers is -0.422 and

[32] Thomas, 1990, and Duflo, 2000.

statistically significant at the 1% level, while the coefficient on the sum of all other children's intended transfers is -0.085 and statistically not significant at all. Similarly, for voluntary educational expenditure, the coefficient on individual intended transfers is 0.586 and statistically significant at the 5% level, while the coefficient on the sum of other children's intended transfers is 0.012 and statistically not significant. As before, I also am unable to reject the null hypothesis that the absolute values of the coefficients on individual intended transfer in columns (3) and (4) are equal (the P-value of the Wald test is 0.550). In addition, as I point out in Section 4.2, if the results are driven by unobservable village or household changes in education expenditures, the individual intended transfers, IT_{iht}, and the sum of other children's intended transfers, IT_{-iht}, should have same effects on education expenditures. But the results in Table 2.9 show that they are different.

Section B in Table 2.9 shows the results from the control experiment that uses data from year 2000 and year 2004. Columns (1) and (2) show the results testing the identification assumption that the difference of educational expenditures between children enrolled in different grades does not change across the time. Both coefficients in these two columns are not significant, which supports the assumption. The last two columns in section B test whether the time trends of household educational expenditures are different in different villages in the absence of the reform. From columns (3) and (4), we can also see that no coefficients are statistically significant. This provides evidence that there are no different time trends of individual educational expenditures in different villages in the absence of the reform.

In this paper, I also investigate the heterogeneous impact of the reform on voluntary educational expenditure in terms of individual characteristics, including gender, birth order, and grade enrolled. All the heterogeneous tests are based on regression function (2.7). Table 2.10 presents the results. Columns (1) to (3) correspond to tests for heterogeneous impacts in terms of gender, birth order, and grade enrolled respectively. The coefficient on the interaction term of the girl dummy and individual intended transfers is 0.585, and is statistically significant at the 1% level. Given one *Yuan* increase in intended transfers, girls can get 0.585 *Yuan* more than boys in terms of voluntary education expenditures from their parents. One explanation for it might be because parents did not spend enough money on girls before the reform due to budget constraint; therefore, after the reform parents would like to make up for the gap. Parents also spend more on older children. The coefficient on the interaction of birth order and individual intended transfer is -0.471 and statistically significant at the 1% level. Consistent with this result, the coefficient on the interaction of individual intended transfer and middle school dummy is 0.864 and statistically significant at the 5% level (column (3)). Thus, with one *Yuan* increase in intended transfer from the reform, parents would spend 0.864 *Yuan* more on voluntary educational expenditures on children enrolled in middle school than on children enrolled in primary school. For children enrolled in middle schools, either because they need more expenditures on school supplies, et.al., or because parents would like to reinforce their competence to enter better high school, or because parents think that the return to the educational investment on middle school kids is higher since they are about to enter the labor market, therefore, with the same amount of money saved from the reform, parents spend more on the voluntary educational expenditure on children enrolled in the middle schools than those enrolled in the primary schools.

2.6 Robustness checks

2.6.1 Do schools “evade” mandated fee reductions through greater voluntary educational expenditures?

One could argue that after the educational fees are exempted, schools may evade the regulations by requiring so-called “voluntary” expenditures; in other words, perhaps the voluntary expenditures are not actually voluntary. An obvious example is that schools might require students to take more tutoring classes organized by schools in order to charge additional tutoring fees after the regular school fees were exempted. Or the school could pass on the costs of services previously provided by the school for free and describe the payments as voluntary. In this section, I provide evidence against this possibility.

Table 2.11 shows the reform's impact on different components of household voluntary educational expenditures. The first column still shows the impact of the reform on household voluntary educational expenditures per capita. Columns (2) to (4) list the impacts of the reform on the different components of voluntary educational expenditures by increasing order of the magnitudes of the estimated impacts. The impact of the reform on tutoring expenditures is the smallest. The coefficient is equal to -0.007 and not significant at all (column (2)). It shows that the schools did not charge additional fees by requiring children to attend more tutoring classes, which is the most feasible way schools can use if they would like to charge more fees from the students.

The impacts of the reform on other components (including expenditures on school supplies, and children's food and snack in school and transportation) are positive and statistically different from zero. The coefficient on expenditures on school supplies is 0.193 and statistically significant at the 5% level (column (3)). The coefficient on expenditures on children's food and snack in school and transportation is 0.464 and statistically significant at the 5% level (column (4)). This suggests that parents used the money saved from the reduced required educational fees to buy more or better quality school supplies and food or snack for their children. It is not possible for schools to require students to buy these from the school since charging students with extra fees is politically sensitive especially in rural China, the school principals would take a big risk if parents complaint about it. Indeed, in our field research, we have never seen such a case that schools sell school supplies or food to students.

In addition, after the educational fees were exempted, the schools were subsidized by the central and local governments. In 2005, the total money used to subsidize schools in Gansu was 216 million *Yuan*[33]. At the same time, all the information is publicized, and parents can complaint about the overcharged fees. Therefore, schools do not have incentive to charge additional fees from students after the regular educational fees were exempted.

2.6.2 Do variations from textbook exemption and living subsidy drive the results?

In this paper, the variation of intended transfer also comes from the variation of getting textbook fee exemption and living subsidies. In principle, the eligibility to get textbook fee exemption and living subsidies depends on household income, however in practice, whether a child could get textbook fee exemption or living subsidies might also depend on other unobservable variables, for example personal connection, corruption,

[33] March 23, 2005, China Education Daily

et.al, which might contaminate the estimation even after household income is controlled in the regressions. In this section, I am going to investigate whether these possible unobservable variables lead to biased OLS estimates. I assume that all children enrolled in primary schools and middle schools are eligible to be exempted from textbook fee charge, but no one gets living subsidy. Then I construct a new variable, which I call *intended transfer II*. It is equal to school fees exemption, which is defined in equation (2.1) in section 2.2, plus 70 (140) if the child is enrolled in primary school (if the child is enrolled in middle school.). The household intended transfer II per capita is equal to the sum of intended transfer II of all the children living in the household divided by the number of total family members. By doing so, the variation of intended transfer II is totally from the variation in pre-reform school fees. Then, I re-estimate functions (2.4), (2.6) and (2.7) by replacing intended transfer with newly constructed intended transfer II. The results are shown in Table 2.12 and Table 2.13.

Table 2.12 shows the results from household level regression. From this table, we can see that only the coefficients shown in columns (6) and (7) are significant. The coefficient in column (6) is equal to -0.608, significant at the 1% level. We should compare this coefficient with that shown in column (6) in Table 2.7, which is equal to -0.613. We can see that they are almost the same. Indeed, I test the hypothesis that these two coefficients are equal, and P-value of Wald test is 0.880, which means that these two coefficients are not significantly different. Then we can turn to the coefficient shown in column (7) in Table 2.12. This coefficient is equal to 0.627, significant at the 5% level. The corresponding coefficient in column (7) in Table 2.7 is equal to 0.651. These two coefficients are very close to each other, and the P-value of Wald test testing their equality is 0.755, which also shows that these two coefficients are not significantly different.

Table 2.13 shows the results from individual level regressions, but using newly defined intended transfer II as treatment variable. The first two columns show the results from estimating function (2.6). The coefficient shown in column (1) is equal to -0.464, significant at the 1% level. However, the coefficient shown in column (2) is equal to 0.586, also significant at the 1% level. Columns (3) and (4) show the estimated results of function (2.7). Coefficient in column (3) is equal to -0.421, significant at the 1% level, but the coefficient in column (4) is equal to 0.537, significant at the 10% level. We can compare these four coefficients in Table 2.13 with those corresponding coefficients shown in Table 2.9. We can see that the estimated impacts of the reform on required educational fees are almost the same. The P-values of Wald test are 0.282 and 0.975 for coefficient shown in columns (1) and (3) respectively. The estimated impacts of the reform on voluntary educational expenditures in Table 2.13 and in Table 2.9 are also close to each other. The P-values of Wald test are 0.944 and 0.519.

The results discussed in above show that the variations from textbook fee charge and living subsidy do not have any effects on the estimation.

2.6.3 The impacts of required educational fees on expenditures: IV strategy

In this section, I investigate the impact of required educational expenditures on other expenditures directly. In order to exploit the variation in required educational expenditures due to the educational fee reduction reform, I use intended transfer from the reform as IV for required educational expenditures. Table 2.14 shows the estimated coefficients. Columns (1) to (6) in the table show the results of household level regressions, and column (7) in the same table shows the result of individual level regression.

Let's firstly look at the results shown in columns (1) to (6) in Table 2.14. In all these columns, household intended transfer from the reform per capita is used as an IV for household required educational expenditures per capita. The first stage regression has been shown in Table 2.7. The F-value of the first stage regression is equal to 20.10. It is larger than 10, which means that household intended transfer from the reform per capita is not a weak IV for household required educational expenditures per capita[34]. We can see that only the coefficient in column (6), i.e., the regression of household voluntary educational expenditures per capita, is significant. The coefficient is equal to -1.061, significant at the 5% level. It means that one *Yuan* decrease in required educational expenditures per capita leads to 1.061 *Yuan* increases in voluntary education expenditure per capita. Although the coefficient is not exactly equal to -1, the P-value of F test testing the hypothesis that the coefficient is equal to -1 is 0.608, which means that the coefficient

[34] Bound, Jaeger and Baker, 1995.

is not significantly different from -1. The result estimated here is also consistent with that shown in Table 2.7. In Table 2.7, one *Yuan* intended transfer per capita from the reform leads to 0.613 *Yuan* decreases in required educational expenditures per capita and 0.651 *Yuan* increases in voluntary educational expenditures per capita, from which we can derive that one *Yuan* decrease in required educational expenditures per capita leads to 1.062 *Yuan* increases in voluntary educational expenditures per capita. In Table 2.14, all other coefficients shown in columns (1) to (5) are not significant, which means that the change in the required educational expenditures due to the reform does not affect household expenditures except for household voluntary educational expenditures.

Column (7) in Table 2.14 shows the result of individual level regression. In this column, the individual intended transfer from the reform and the sum of other children's intended transfer from the reform in the same family are used as IVs for individual required educational expenditures and the sum of other children's required educational expenditures in the same family. I do not present the regression result for the first stage in Table 2.14, but the F-values are 26.83 and 60.12 in the first stage regression of individual required educational expenditures and the sum of all other children's required educational expenditures in the same family. These two F-values are larger than 10, which mean that the IVs used are not weak[35]. In column (7), the coefficient before individual required educational expenditures is -1.312, significant at the 10% level, while the coefficient before the sum of other children's required educational expenditures is not significant. It means that one *Yuan* decrease in individual required educational expenditures leads to 1.312 *Yuan* increases in voluntary educational expenditures on this child. I also test the hypothesis that this coefficient is equal to -1. The P-value of F test is equal to 0.643, which means that the coefficient is not significantly different from -1. The result shown in Table 2.14 is consistent with that shown in Table 2.9. As shown in columns (3) and (4) in Table 2.9, one *Yuan* increase in the intended transfer from the reform leads to 0.422 *Yuan* decreases in required educational expenditures and 0.586 *Yuan* increases in voluntary educational expenditures, which means that one *Yuan* decrease in required educational expenditures leads to 1.389 *Yuan* increases in voluntary educational

[35] Bound, Jaeger and Baker, 1995.

expenditures, which is slightly larger but still very similar to the estimates shown in Table 2.14.

2.6.4 Measurement error

Measurement error in household income as control variable could lead to attenuation bias in the income coefficient, while could bias other coefficient estimates as well, especially if incomes are correlated with intended transfers. In order to see the effects of measurement error on the coefficient of interest, I use the average value of income of all other households living in the same village in the same year as an IV. A critical assumption for the validity of this IV is that measurement errors in the incomes of different households are independent. Table 2.15 shows the results for the household level (columns (1) to (6)) and individual level (columns (7) to (8)).

We find that the results do not change in any noticeable may when we use instruments. The coefficients shown in columns (1) to (4) at Table 2.15 are not significant, the same as in section A at Table 2.7. The coefficients shown in columns (5) and (6) are significant. The coefficient shown in column (5) is -0.601, significant at the 1% level; and the coefficient shown in column (6) is 0.692, significant at the 5% level. Comparing with the corresponding coefficients shown in Table 2.7, we can see that although the estimated impact of the reform on household required educational expenditures is slightly smaller in Table 2.15, but the estimated impact of the reform on household voluntary educational expenditures is slightly larger in Table 2.15, both coefficients shown in Table 2.7 and Table 2.15 are very close to each other. Then, we can see the results in columns (7) to (8). The one shown in column (7) is -0.498, significant at the 5% level; the one shown in column (8) is 0.527, significant at the 10% level. Compared with the coefficients shown in columns (3) to (4) in Table 2.9, although the estimated impact of the reform on individual required educational expenditures is larger, but the estimated impact of the reform on individual voluntary educational expenditures is smaller in Table 2.15, both coefficients shown in Table 2.9 and Table 2.15 are very close to each other too.

From the above discussion, we can see that although there might be some measurement errors in the variables, the estimation of the reform's impact is not affected.

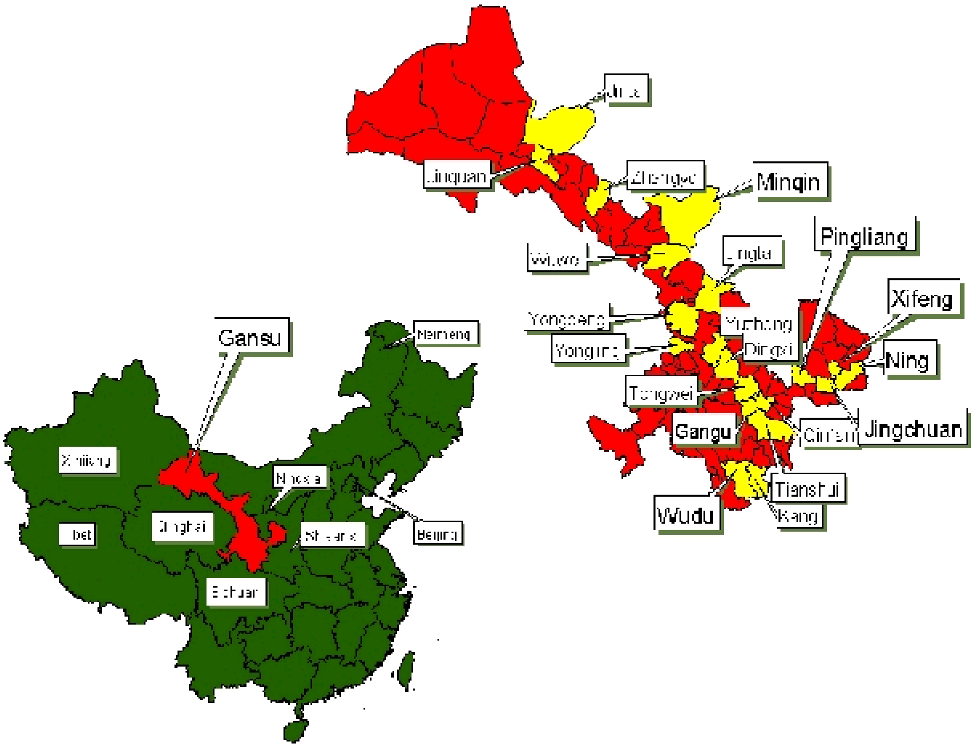
2.7 Conclusion

This paper tests for the existence of an intra-household flypaper effect by investigating the impact of an educational fee reduction reform in rural China on household expenditures.

Using household survey data from the Gansu Survey of Children and Families, I first investigate the impact of the reform on household level expenditures. I find that a one *Yuan* increase in the household intended transfers per capita from the reform leads to a 0.613 *Yuan* decrease in household required educational expenditures per capita, and a 0.651 *Yuan* increase in household voluntary educational expenditures per capita. The absolute values of these two coefficients are not statistically different. But there are no significant impacts of the reform on household total income per capita, total expenditures per capita, or other categories of expenditures, including expenditures on food, non-food consumption and services, and health care. And I also find that a one *Yuan* increase in individual intended transfers leads to a 0.422 *Yuan* decrease in the individual required educational expenditures, but leads to a 0.586 *Yuan* increase in the individual voluntary educational expenditures. The absolute values of these two coefficients are also not significantly different. In addition, I find that with the same amount of increase in the intended transfers from the reform, households having more educated mothers spend more on voluntary educational items, and parents spend more on girls, older kids and kids enrolled in middle schools. Overall, this paper provides a strong evidence for the existence of an intra-household flypaper effect.

The findings in this paper also have policy implications. Besides the commonly recognized positive impacts of governments' educational subsidy programs on targeted children's enrolment, this paper shows that these programs also increase parents' investment on enrolled children's education, which increases their short- and long-run welfare. It also shows that when evaluating governments' educational subsidy programs, we should include the flypaper effects into benefit-cost analysis.

Figure 2.1 Gansu Province and GSCF counties



Map 1. Gansu Province, GSCF Counties Marked

Source: Hannum (2001)

Figure 2.2 Kernel density of exempted school fees

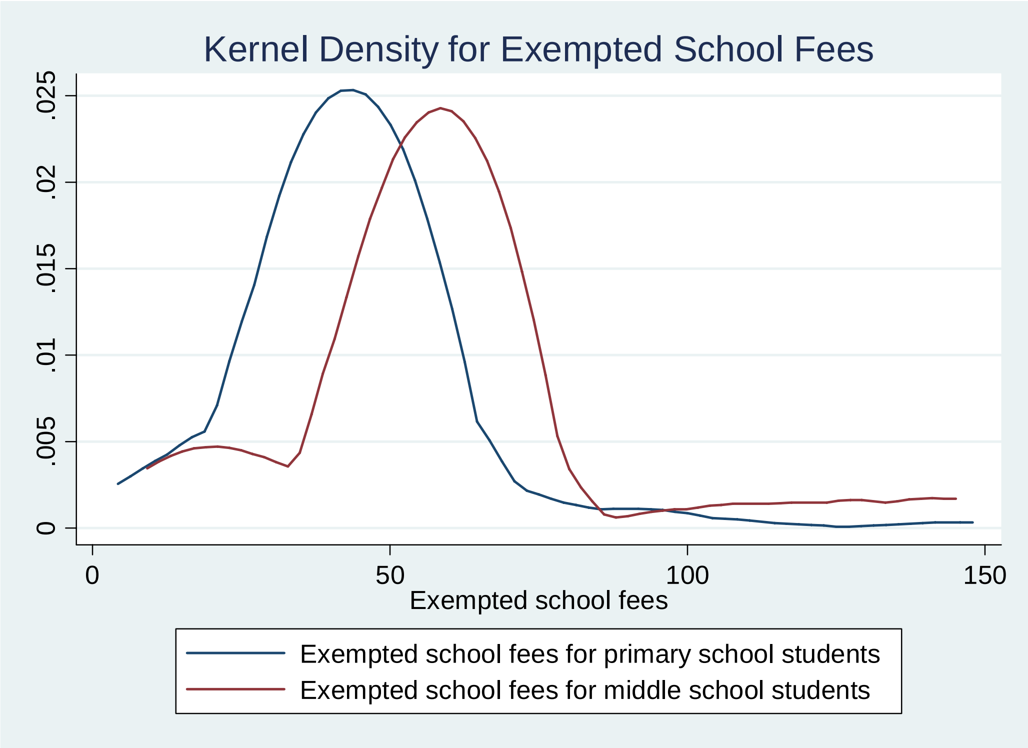
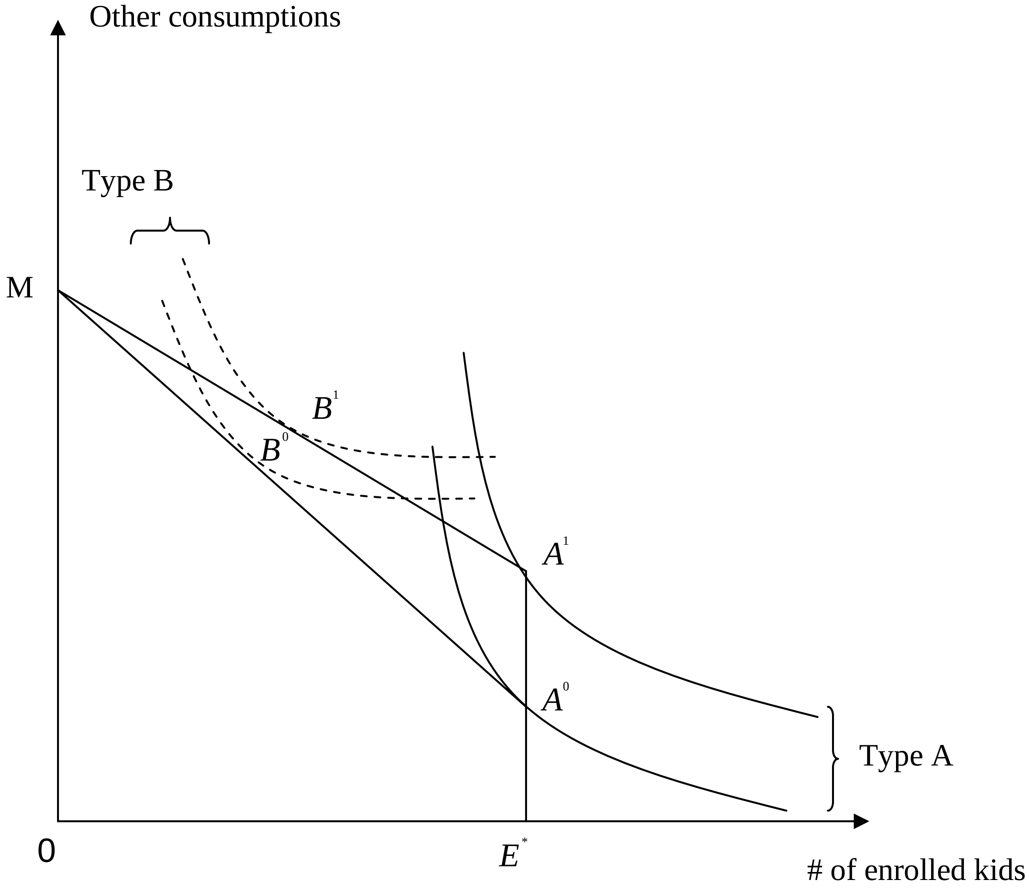


Figure 2.3



Other consumptions
Type B
M
B1
B0
A1
A0
Type A
0
E*
# of enrolled kids

Figure 2.4

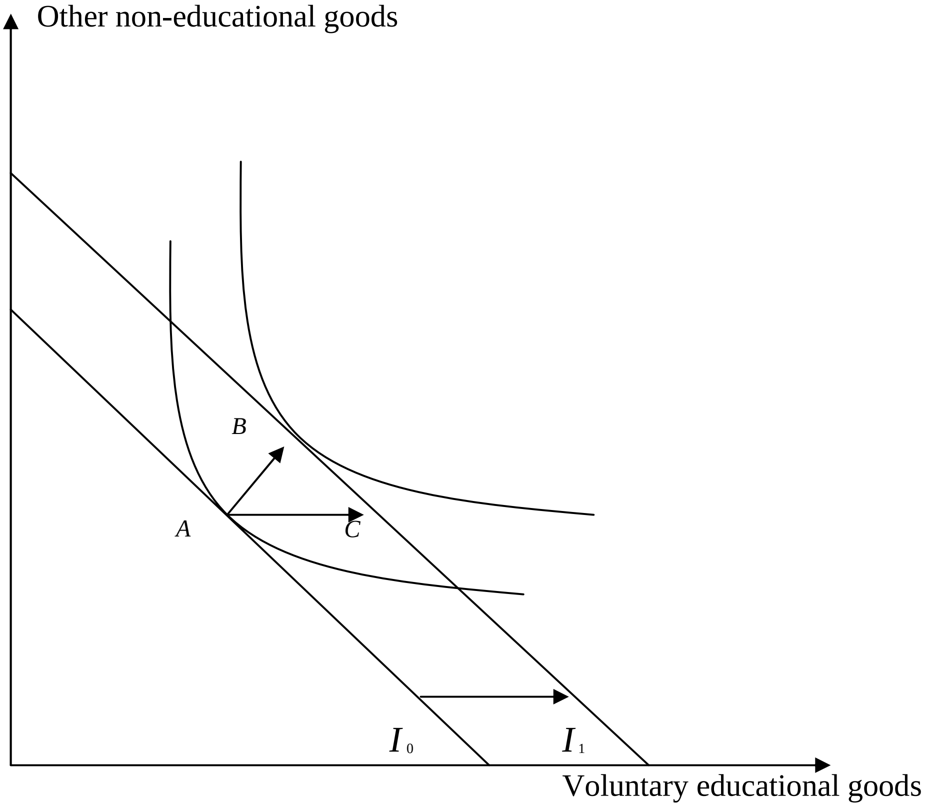

Other non-educational goods
B
A
C
I_0
I_1
Voluntary educational goods

Figure 2.5

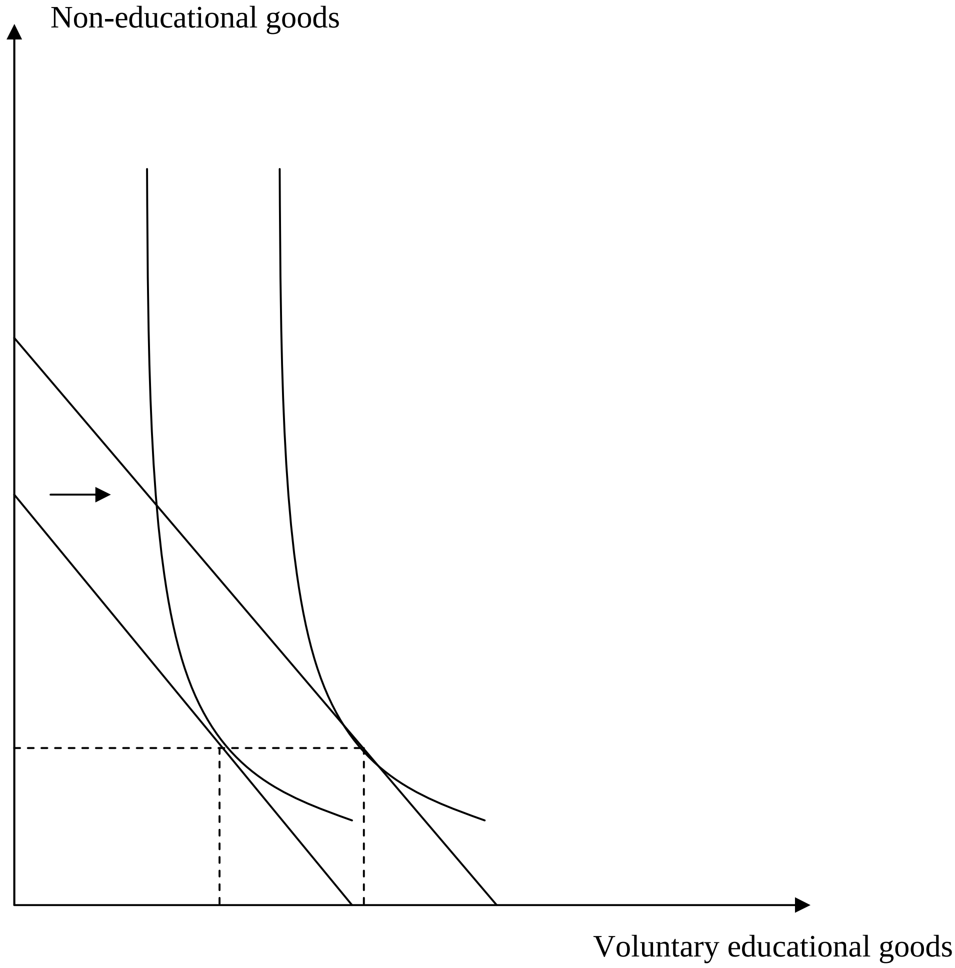
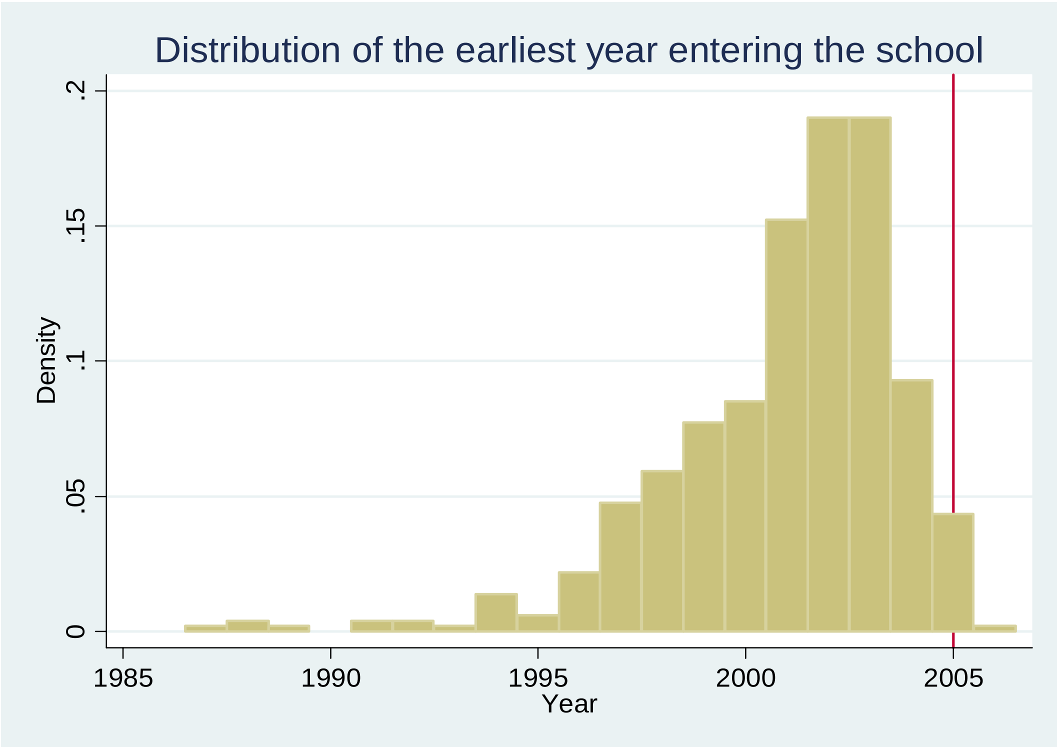


Figure 2.A



Distribution of the earliest year entering the school
Density
0
.05
.1
.15
.2
1985
1990
1995
2000
2005
Year

Table 2.1 Characteristics of households and children in the sample

	2000 Mean	2000 S.D.	2007 Mean	2007 S.D.
Panel A Households characteristics				
Household size	4.489	1.084	4.265	1.140
Number of children	2.274	0.711	2.002	0.768
Number of enrolled children	1.847	0.716	1.733	0.741
Number of enrolled children/number of children	0.830	0.235	0.890	0.233
Father's schooling year	6.953	3.566	7.251	2.899
Mother's schooling year	4.192	3.513	5.275	3.436
Land area per capita (Mu)	2.075	1.497	2.060	1.457
Total income per capita (Yuan)	1052.819	1359.057	1375.746	1659.385
Total expenditure per capita (Yuan)	1065.753	1101.159	1318.828	1001.087
Observation	1629	1629	505	505
Panel B Children characteristics				
Age	11.419	2.570	11.805	2.512
Female	0.488	0.500	0.486	0.500
Total enrolment rate	0.858	0.349	0.947	0.223
Enrolment rate of children aged 9-16	0.930	0.255	0.982	0.131
Observation	3498	3498	912	912
Of which, enrolled children				
Age	11.347	2.010	11.797	2.268
Female	0.482	0.500	0.484	0.500
Indicator for being enrolled in primary school	0.883	0.322	0.811	0.391
Indicator for being enrolled in middle school	0.111	0.314	0.164	0.371
Indicator for being enrolled in other levels of school	0.006	0.079	0.024	0.154
Observation	3001	3001	864	864

All the money has been deflated to year 2000 value

Table 2.2 Percentage of counties claiming to implement TEOS in each semester (%)

<u>Primary School</u>

	Fall, 2004	Spring, 2005	Fall, 2005	Spring, 2006	Fall, 2006	Spring. 2007
School Fees Exemption	15	65	65	100	100	100
Textbook Exemption	30	75	75	100	100	100
Living Subsidy	0	10	15	65	70	70

<u>Middle School</u>

	Fall, 2004	Spring, 2005	Fall, 2005	Spring, 2006	Fall, 2006	Spring. 2007
School Fees Exemption	15	65	65	100	100	100
Textbook Exemption	30	80	80	100	100	100
Living Subsidy	0	60	65	95	95	95
Num. of Counties	20	20	20	20	20	20

Table 2.3 Percentage of children claiming to have TEOS in each semester (%)

Primary School

	Fall, 2004	Spring, 2005	Fall, 2005	Spring, 2006	Fall, 2006	Spring. 2007	Total student number
School Fees Exemption	7.32	44.24	52.10	61.94	92.07	95.73	1311
Textbook Exemption	7.70	40.12	48.97	60.34	80.70	85.74	1311
Living Subsidy	0.00	0.31	0.31	0.69	0.84	0.92	1311

Middle School

	Fall, 2004	Spring, 2005	Fall, 2005	Spring, 2006	Fall, 2006	Spring. 2007	Total student number
School Fees Exemption	8.53	40.30	49.47	56.29	88.06	91.04	469
Textbook Exemption	8.10	41.36	49.89	56.50	74.63	78.04	469
Living Subsidy	0.21	3.20	3.20	4.90	7.25	8.96	469

(1) The total student sample is constructed according to the following (a) children must be in school, (b) children reporting policy time later than "liu ji" time were dropped; (c) children having drop-out experience were dropped; (d) primary school student sample includes students in grade 3 to grade 6 in 2007; middle school student sample includes student in grade 3.

(2) Those reporting policies earlier than the year when the policies existed or later than the time when the surveyed was conducted were dropped.

Table 2.4 Intended transfer and household expenditures in 2007

	All		Richest 30%		Poorest 30%	
Variables	Mean	Percentage (%)	Mean	Percentage (%)	Mean	Percentage (%)
Total expenditure per capita	1318.828	100	1871.353	100	1041.442	100
In which :						
Food expenditure per capita	425.418	32.257	644.890	34.461	301.357	28.937
Non-food expenditure per capita	586.485	44.470	825.368	44.105	478.213	45.918
Health expenditure per capita	202.250	15.336	272.060	14.538	165.480	15.889
Required educational expenditure per capita	17.874	1.355	20.206	1.080	19.710	1.893
Voluntary educational expenditure per capita	86.800	6.582	108.830	5.816	76.682	7.363
Intended transfer per capita	39.183	2.971	41.115	2.197	39.305	3.774
OBS	505	505	152	152	152	152

Note: .

(1) Percentages in this table are percentage of total expenditure per capita

(2) All money is deflated to year 2000 value

Table 2.5 Household expenditures in years 2000 and 2007

	2000	2007	2007-2000
Total expenditure per capita	1065.753	1318.828	253.075***
	(1101.159)	(1001.087)	(52.238)
Food expenditure per capita	264.245	425.418	161.173***
	(251.209)	(327.658)	(15.853)
Non-food expenditure per capita	501.916	586.485	84.569**
	(923.528)	(688.433)	(38.237)
Health expenditure per capita	157.725	202.250	44.525**
	(323.759)	(390.756)	(19.149)
Required educational expenditure per capita	89.223	17.874	-71.349***
	(106.120)	(43.939)	(3.277)
Voluntary educational expenditure per capita	52.643	86.800	34.157***
	(79.675)	(134.694)	(6.311)
Number of households	1629	505	

Standard deviations are in parenthesis

Table 2.6 Descriptive results

	Total expenditure per capita			Expenditure on food per capita		
	Treated	Untreated	Treated-Untreated	Treated	Untreated	Treated-Untreated
2000	1178.58	922.65	255.93	301.36	218.53	82.83
2007	1516.85	1080.65	436.20	457.37	351.80	105.57
2007-2000	338.27	158.00	180.27	156.01	133.27	22.74
	Expenditure on non-food goods and service per capita			Expenditure on health care per capita		
	Treated	Untreated	Treated-Untreated	Treated	Untreated	Treated-Untreated
2000	544.75	434.88	109.87	157.89	157.86	0.03
2007	677.13	489.13	188.00	235.13	170.64	64.49
2007-2000	132.38	54.25	78.13	77.24	12.78	64.46
	Required education expenditure per capita			Voluntary education expenditure per capita		
	Treated	Untreated	Treated-Untreated	Treated	Untreated	Treated-Untreated
2000	114.72	65.03	49.69	59.86	46.34	13.51
2007	23.86	14.44	9.42	124.66	55.81	68.84
2007-2000	-90.87	-50.59	-40.27***	64.80	9.47	55.33***

* 10% significant level; ** 5% significant level; *** 1% significant level.

In this table, "Treated" group includes the villages where the average saving from the reform is above the median value. "Untreated" group includes the villages where the average saving from the reform is below the median value.

Table 2.7 Impacts of education fee reduction reform on household expenditure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household income per capita	Total expenditure per capita	Expenditure on food per capita	Expenditure on non-food consumption and service per capita	Expenditure on health care per capita	Expenditure on required education items per capita	Expenditure on voluntary education items per capita
Section A 2000-2007							
Intended transfer/total family member	-0.677	-2.222	-0.014	-1.770	-0.476	-0.613	0.651
	(2.869)	(2.214)	(0.464)	(1.539)	(0.700)	(0.137)***	(0.264)**
Observations	2134	2134	2134	2134	2134	2134	2134
R-squared	0.25	0.22	0.50	0.13	0.11	0.45	0.53
Wald Test: H0: Absolute value of coefficients in columns (6) and (7) are equal; P-value=0.899							
Section B 2000-2004							
Hypothetical intended transfer /total family member	4.498	0.844	-0.227	-0.007	0.699	0.149	0.231
	(5.665)	(2.013)	(0.542)	(1.080)	(1.259)	(0.109)	(0.198)
Observations	2991	2991	2991	2991	2991	2991	2991
R-squared	0.21	0.28	0.43	0.17	0.06	0.63	0.58
Village variables in year 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log value of household income per capita	No	Yes	Yes	Yes	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. of kids enrolled	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household demographic structure	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses, standard errors are calculated clustering over villages; * significant at 10%; ** significant at 5%; *** significant at 1%

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) household endowments include land area per capita, mother's schooling year and father's schooling year.

(3) number of kids enrolled includes number of kids enrolled in primary school, number of kids enrolled in middle school, number of kids enrolled in high school and number of kids enrolled in other level of schools

(4) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19, ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5, ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.8 Heterogeneity of the reform's impacts on household expenditure on voluntary educational items per capita

	(1) Expenditure on voluntary education items per capita	(2) Expenditure on voluntary education items per capita	(3) Expenditure on voluntary education items per capita	(4) Expenditure on voluntary education items per capita
Intended transfer from the reform per capita(*)log value of household income per capita	0.148			
	(0.121)			
Intended transfer from the reform per capita(*)total number of enrolled kids		0.274		
		(0.202)		
Intended transfer from the reform per capita(*)mother's schooling year			0.083	
			(0.040)**	
Intended transfer from the reform per capita(*)father's schooling year				0.037
				(0.051)
Intended transfer from the reform per capita	-0.348	-0.081	0.048	0.353
	(0.828)	(0.443)	(0.399)	(0.529)
Village variables in year 2000	Yes	Yes	Yes	Yes
Log value of household income per capita	Yes	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes	Yes
Num. of kids enrolled	Yes	Yes	Yes	Yes
Household demographic structure	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes	Yes
Observations	2134	2134	2134	2134
R-squared	0.53	0.36	0.53	0.53

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) household endowments include land area per capita, mother's schooling year and father's schooling year.

(3) number of kids enrolled includes number of kids enrolled in primary school, number of kids enrolled in middle school, number of kids enrolled in high school and number of kids enrolled in other level of schools

(4) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19, ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5, ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.9 Impacts of education fee reduction reform on individual education expenditures

	(1) Required education expenditure	(2) Voluntary education expenditure		(3) Required education expenditure	(4) Voluntary education expenditure
Section A 2000-2007					
Deviation of individual intended transfer from the reform from household mean value	-0.493	0.578	Individual intended transfer from the reform	-0.422	0.586
	(0.049)***	(0.160)***		(0.142)***	(0.255)**
			Sum of other kids' intended transfer from the reform	-0.085	0.012
				(0.090)	(0.118)
Observations	3865	3865		3865	3865
R-squared	0.47	0.44		0.38	0.49
Wald test: H0: |(1)|=|(2)|, P-value:	0.794		Wald test: H0: |(3)|=|(4)|, P-value:	0.550	
Section B 2000-2004					
Deviation of individual intended transfer from the reform from household mean value	-0.241	0.332	Individual intended transfer from the reform	0.021	0.328

	(0.155)	(0.356)		(0.122)	(0.251)
			Sum of other kids' intended transfer from the reform	0.043	-0.016
				(0.095)	(0.139)
Observations	5342	5342		5342	5342
R-squared	0.653	0.397		0.66	0.52
Deviation of individual characteristics from household mean value	Yes	Yes	Village variables in 2000 interacted with year 2007 dummy	Yes	Yes
Village fixed effect	Yes	Yes	Individual characteristics	Yes	Yes
			Dummies for grade enrolled	Yes	Yes
			Household endowments	Yes	Yes
			Household demographic structure	Yes	Yes
			Household income per capita	Yes	Yes
			Year fixed effect	Yes	Yes
			Village fixed effect	Yes	Yes

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%;

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) individual characteristics include indicator for female and age

(3) dummies for grade enrolled include indicator for being enrolled in grade1-grade6 in primary schools, grade1-grade3 in middle schools, grade1-grade3 in high schools and other levels of schools

(4) household endowments include land area per capita, mother's schooling year and father's schooling year

(5) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19 ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5 ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.10 Heterogeneity of the reform's impacts on individual voluntary expenditure on educational items

	(1) Individual voluntary education expenditure	(2) Individual voluntary education expenditure	(3) Individual voluntary education expenditure
Individual intended transfer from the reform*indicator for female	0.585		
	(0.219)***		
Individual intended transfer from the reform*birth order		-0.471	
		(0.166)***	
Individual intended transfer from the reform*middle school dummy			0.864
			(0.369)**
Individual intended transfer from the reform	0.343	0.934	-0.053
	(0.266)	(0.295)***	(0.327)
Household aggregate saving from the reform	Yes	Yes	Yes
Birth order	No	Yes	No
Village variables in 2000	Yes	Yes	Yes
Individual characteristics	Yes	Yes	Yes
Dummies for grade enrolled	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes
Household demographic structure	Yes	Yes	Yes
Household income per capita	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes
Observations	3865	3865	3865
R-squared	0.49	0.49	0.46

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) individual characteristics include indicator for female and age

(3) dummies for grade enrolled include indicator for being enrolled in grade1-grade6 in primary schools, grade1-grade3 in middle schools, grade1-grade3 in high schools and other levels of schools

(4) household endowments include land area per capita, mother's schooling year and father's schooling year

(5) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19 ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5 ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.11 Decomposition of the reform's impact on voluntary educational expenditures

	(1)	(2)	(3)	(4)
	Expenditure on voluntary education items per capita	Expenditure on tutoring	Expenditure on school supplies	Expenditure on snack in school and transportation
Intended transfer/total family member	0.651	-0.007	0.193	0.464
	(0.264)**	(0.027)	(0.095)**	(0.210)**
Village variables in year 2000	Yes	Yes	Yes	Yes
Log value of household income per capita	Yes	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes	Yes
Num. of kids enrolled	Yes	Yes	Yes	Yes
Household demographic structure	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes	Yes
Observations	2134	2134	2134	2134
R-squared	0.53	0.20	0.48	0.47

Robust standard errors in parentheses, standard errors are calculated clustering over villages; * significant at 10%; ** significant at 5%; *** significant at 1%

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) household endowments include land area per capita, mother's schooling year and father's schooling year.

(3) number of kids enrolled includes number of kids enrolled in primary school, number of kids enrolled in middle school, number of kids enrolled in high school and number of kids enrolled in other level of schools

(4) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19, ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5, ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.12 Robustness check using intended transfer II in household level regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household income per capita	Total expenditure per capita	Expenditure on food per capita	Expenditure on non-food consumption and service per capita	Expenditure on health care per capita	Expenditure on required education items per capita	Expenditure on voluntary education items per capita
Intended transfer II/total family member	-1.914	-2.505	-0.294	-1.980	-0.250	-0.608	0.627
	(2.123)	(2.421)	(0.452)	(1.878)	(0.643)	(0.129)***	(0.265)**
Wald test: H0: (6) in table 12 = (6) in section A in table 7; (7) in table 12 = (7) in section A in table 7; P-value:						0.880	0.755
Village variables in year 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log value of household income per capita	No	Yes	Yes	Yes	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. of kids enrolled	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household demographic structure	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2134	2134	2134	2134	2134	2134	2134
R-squared	0.56	0.22	0.50	0.14	0.11	0.45	0.53

Robust standard errors in parentheses, standard errors are calculated clustering over villages; * significant at 10%; ** significant at 5%; *** significant at 1%

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) household endowments include land area per capita, mother's schooling year and father's schooling year.

(3) number of kids enrolled includes number of kids enrolled in primary school, number of kids enrolled in middle school, number of kids enrolled in high school and number of kids enrolled in other level of schools

(4) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19, ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5, ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.13 Robustness check using intended transfer II in individual level regression

	(1)	(2)		(3)	(4)
	Required education expenditure	Voluntary education expenditure		Required education expenditure	Voluntary education expenditure
Deviation of individual intended transfer from the reform from household mean value	-0.464	0.586	Individual intended transfer from the reform	-0.421	0.537
	(0.050)***	(0.164)***		(0.152)***	(0.283)*
			Sum of other kids' intended transfer from the reform	-0.054	0.019
				(0.080)	(0.101)
Observations	3865	3865		3865	3865
R-squared	0.47	0.44		0.38	0.49
Wald test: H0: coefficients in this table is equal to coefficients in section A in table 2.9; P-value:	0.282	0.944		0.975	0.519

Deviation of individual characteristics from household mean value	Yes	Yes
Village fixed effect	Yes	Yes

Village variables in 2000 interacted with year 2007 dummy	Yes	Yes
Individual characteristics	Yes	Yes
Dummies for grade enrolled	Yes	Yes
Household endowments	Yes	Yes
Household demographic structure	Yes	Yes
Household income per capita	Yes	Yes
Year fixed effect	Yes	Yes
Village fixed effect	Yes	Yes

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%;

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) individual characteristics include indicator for female and age

(3) dummies for grade enrolled include indicator for being enrolled in grade1-grade6 in primary schools, grade1-grade3 in middle schools, grade1-grade3 in high schools and other levels of schools

(4) household endowments include land area per capita, mother's schooling year and father's schooling year

(5) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19 ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5 ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

Table 2.14 Robustness check of change in the specification of regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household income per capita	Household total expenditure per capita	Expenditure on food per capita	Expenditure on non-food consumption and service per capita	Expenditure on health care per capita	Expenditure on voluntary education items per capita	Individual voluntary education expenditure
Expenditure on required education items per capita (Household intended transfer from the reform per capita as IV)	1.748	3.626	0.023	2.888	0.776	-1.061	
	(3.524)	(3.471)	(0.757)	(2.426)	(1.139)	(0.510)**	
Individual required education expenditure (Individual intended transfer from the reform as IV)							-1.312
							(0.671)*
Sum of other kids' required education expenditure(Sum of other kids' intended transfer in the same family as IV)							0.074
							(0.120)

F-test: H0: (6) =-1 ; (7) =-1; P-value:						0.608	0.643
Village variables in year 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log value of household income per capita	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. of kids enrolled	Yes	Yes	Yes	Yes	Yes	Yes	No
Household demographic structure	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	No	No	No	No	No	No	Yes
Dummies for grade enrolled	No	No	No	No	No	No	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2134	2134	2134	2134	2134	2134	3865
R-squared	0.55	0.20	0.50	0.08	0.08		

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) household endowments include land area per capita, mother's schooling year and father's schooling year.

(3) number of kids enrolled includes number of kids enrolled in primary school, number of kids enrolled in middle school, number of kids enrolled in high school and number of kids enrolled in other level of schools

(4) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19, ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5, ratio of female aged 6-12, ratio of female aged 13-16, ratio of female

aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

(5) individual characteristics include indicator for female and age

(6) dummies for grade enrolled include indicator for being enrolled in grade1-grade6 in primary schools, grade1-grade3 in middle schools, grade1-grade3 in high schools and other levels of schools

(7) The F-value of the first stage regression is 20.10 for IV estimates in the household level regressions; The F-value of the first stage regression is 26.83 for individual saving as IV, and it is equal to 60.12 for sum of other kids' saving as IV in individual regression.

Table 2.15 Impacts of educational fee reduction reform on expenditures after correcting for measurement error

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total expenditure per capita	Expenditure on food per capita	Expenditure on non-food consumption and service per capita	Expenditure on health care per capita	Expenditure on required education items per capita	Expenditure on voluntary education items per capita	Individual required education expenditure	Individual voluntary education expenditure
Intended transfer/total family member	-2.012	0.059	-1.661	-0.501	-0.601	0.692		
	(2.359)	(0.601)	(1.572)	(0.709)	(0.139)***	(0.310)**		
Individual intended transfer							-0.498	0.527
							(0.232)**	(0.293)*
Sum of other kids' intended transfer							-0.157	-0.012
							(0.147)	(0.173)
Village variables in year 2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log value of household income per capita	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household endowments	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num. of kids enrolled	Yes	Yes	Yes	Yes	Yes	Yes	No	NO
Household demographic structure	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual characteristics	No	No	No	No	No	No	Yes	Yes
Dummies for grade enrolled	No	No	No	No	No	No	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2134	2134	2134	2134	2134	2134	3865	3865
R-squared	0.06	0.22	0.07	0.09	0.30			0.01

Robust standard errors in parentheses, standard errors are calculated clustering over villages; * significant at10%; ** significant at 5%; *** significant at 1%.

(1) village variables in 2000 include illiterate ratio, ratio of primary school graduates in labor force, primary school enrollment rate, middle school enrollment rate, indicator for having preschool classes, distance of primary school from the village seat, log value of money given to schools by village, indicator for having primary school run by the village, indicator for having middle school run by the village, indicator for having railway through the village, indicator for having bus stop in the village, log value of average agricultural income per capita, log value of average industrial income per capita, ratio of households engaged in non-agricultural work, ratio of households running household enterprise, log value of wage for men doing non-agricultural work, log value of wage for women doing non-agricultural work, log value of wage to hire labor for agricultural production, indicator for having enterprises owned by county, indicator for having enterprises owned by township, indicator for having private enterprises.

(2) household endowments include land area per capita, mother's schooling year and father's schooling year.

(3) number of kids enrolled includes number of kids enrolled in primary school, number of kids enrolled in middle school, number of kids enrolled in high school and number of kids enrolled in other level of schools

(4) household demographic structure includes ratio of male aged 0-5, ratio of male aged 6-12, ratio of male aged 13-16, ratio of male aged 17-19, ratio of male aged 20-29, ratio of male aged 30-39, ratio of male aged 40-49, ratio of male aged 50-54, ratio of male aged above 54, ratio of female aged 0-5, ratio of female aged 6-12, ratio of female aged 13-16, ratio of female aged 17-19, ratio of female aged 20-29, ratio of female aged 30-39, ratio of female aged 40-49, ratio of female aged 50-54 and total number of family members.

(5) individual characteristics include indicator for female and age

(6) dummies for grade enrolled include indicator for being enrolled in grade1-grade6 in primary schools, grade1-grade3 in middle schools, grade1-grade3 in high schools and other levels of schools

Table 2.A Testing household utility function

	(1)	(2)	(3)	(4)
	HH voluntary educational expenditure	HH non-educational expenditures	HH voluntary educational expenditure	HH non-educational expenditures
Total expenditure - required educational expenditures	0.009	0.991		
	(0.002)***	(0.002)***		
Total income - required educational expenditures			-0.001	0.170
			(0.001)	(0.019)***
Land area per capita	-2.650	2.650	-2.640	9.406
	(6.392)	(6.392)	(6.450)	(89.534)
Mother's schooling years	-2.908	2.908	-2.220	55.713
	(2.320)	(2.320)	(2.341)	(32.495)*
Father's schooling years	1.168	-1.168	1.431	22.471
	(2.129)	(2.129)	(2.148)	(29.822)
Num of kids enrolled in primary schools	52.041	-52.041	50.182	-295.185
	(15.171)***	(15.171)***	(15.307)***	(212.484)
Num of kids enrolled in middle schools	226.332	-226.332	231.360	391.294
	(20.452)***	(20.452)***	(20.619)***	(286.229)
Num of kids enrolled in high schools	930.718	-930.718	934.668	-45.104
	(62.558)***	(62.558)***	(63.177)***	(877.021)
Num of kids enrolled in other levels of schools	2,109.959	-2,109.959	2,119.839	316.783
	(134.436)***	(134.436)***	(135.876)***	(1,886.208)
Household demographic structure	YES	YES	YES	YES
Village fixed effects	YES	YES	YES	YES
Observations	1629	1629	1629	1629
R-squared	0.54	1.00	0.53	0.23

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

References:

[1] Afridi, Farzana (2005). *Child welfare programs and child nutrition: evidence from a mandated school meal program*. Working paper, Syracuse University.

[2] Al-Samarrai, Samer and Hassan Zaman (2000). *Abolishing school fees in Malawi: the impact on education access and equity*. Working Paper, University of Sussex and World Bank.

[3] Angelucci, M. (2004). *Aid and migration: An analysis of the impact of Progresa on the timing and size of labor migration*. Discussion papers 1187, IZA.

[4] Angelucci, M., and G. De Giorgi (2006). *Indirect effects of an aid program: The case of Progresa and consumption*. Mimeo.

[5] Attanasio, O. and V. Lechene (2002). *Tests of income pooling in household decisions*. Review of economic dynamics, 5, pp. 720-748.

[6] Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L. Linden, and Francisco Perez (2007). *Using conditional transfers in education to investigate intra family decision: evidence from a randomized experiment*. Working paper, World Bank, University of Chicago, Columbia University, and Ministry of Education, Colombia.

[7] Becker, Gary S. (1974). *A theory of social interactions*. Journal of Political Economy, vol. 82, (November-December), pp. 1063-1093.

[8] Becker, Gary S. (1981). Treatise on the family. Cambridge, MA: Harvard University Press.

[9] Behrman, Jere R. and Piyali Sengupta (2005). *Progressing through PROGRESA: an impact assessment of a school subsidy experiment in rural Mexico*. Economic Development and Cultural Change, pp. 238-275.

[10] Behrman, Jere R. and J. Hoddinott (2005). *Program evaluation with unobserved heterogeneity and selective implementation: The Mexican Progresa impact on child nutrition*. Oxford Bulletin of Economics and Statistics, 67, pp. 547-569.

[11] Bobonis, G.J. (2004). *Income transfers, marital dissolution, and intra-household resource allocation: Evidence from rural Mexico*. Mimeo.

[12] Bobonis, G.J. and F. Finan (2005). *Endogenous peer effects in school participation*. Working paper, University of Toronto, Ontario, Canada and UC-Berkeley, CA.

[13] Browning, Martin, Francois Bourguignon, Pierre-Andre Chiappori, and Valerie Lechene (1994). *Incomes and outcomes: a structural model of intrahousehold allocation*. Journal of Political Economy, December, 102(6), po. 1067-1096.

[14] Choi, James, David Laibson, Brigitte C. Madrian (2007). *The flypaper effect in individual investor asset allocation*. Working Paper, Yale University, Harvard University, and Harvard University.

[15] De Janvry, A., Sadoulet Finan F., R. Vakis (2006). *Can conditional cash transfers serve as safety nets to keep children out of school and out of the labor market*. Journal of Development

Economics, 79, pp. 349-373.

[16] Del Boca, Daniela and Christopher J. Flinn (1994). *Expenditure decisions of divorced mothers and income composition*. Journal of Human Resources, Summer, 29(3), pp. 742-761.

[17] Duflo, E. (2000). *Child health and household resources in South Africa: Evidence form the old age pension program*. American Economic Review, 90, pp. 393-398.

[18] Filmer, D., N. Schady (2006). *Getting girls into school: Evidence from a scholarship in Cambodia*. Working paper 3910. World Bank Research, Washington DC.
[19] Gertler, P.J. (2004). *Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment*. American Economic Review, 94, pp. 336-341.
[20] Gertler, P.J., S. Martinez, and M. Rubio-Codina (2006). *Investing cash transfers to raise long term living standards*. Working paper series 3994, Policy Research, The World Bank.
[21] Guangming Daily (2007).
[22] Hannum, E. (2001). *Investigating Children's Schooling in the Interior: The Gansu Survey of Children and Families*. China Education Forum, 2(1).
[23] Hines, James R. Jr. and Richard H. Thaler (1995). *Anomalies, the flypaper effect*. Journal of Economic Perspectives, Vol. 9, No. 4, pp. 217-226.
[24] Jacoby, Hanan G. (2002). *Is there an intrahousehold flypaper effect? Evidence from a school feeding program*. Economic Journal, Vol. 112, No. 476. January, pp. 196-221.
[25] Khandker, S., M. Pitt, N. Fuwa (2003). *Subsidy to promote girls' secondary education: The female stipend program in Bangladesh*. Mimeo.
[26] Knight, Brian (2002). *Endogenous federal grants and crowd-out of state government spending: theory and evidence from the federal highway aid program*. The American Economic Review; March, 92,1; pp. 71-92.
[27] Kooreman, Peter (2000). *The labeling effect of a child benefit system*. The American Economic Review, Vol. 90, No.3. June, pp. 571-583.
[28] Lalvani, Mala (2002). *The flypaper effect: evidence from India*. Public Budgeting and Finance, fall.
[29] Lundberg, Shelly J. and Robert A. Pollak (1993). *Separate spheres bargaining and the marriage market*. Journal of Political Economy, December, 101(6), pp. 998-1010.
[30] Maluccio, J. and R. Flores (2004). *Impact evaluation of a conditional cash transfer program: The Nicaraguan red de protection social*. Discussion paper No. 184. FCND, IFPRI, Washington DC.
[31] Morris, S., R. Flores, P. Olinto, and J.M. Medina (2004). *Monetary incentives in primary health care and effects on use and coverage of preventive heath care interventions in rural Honduras: Cluster randomized trial*. Lancet 364(9450), pp. 2030-2037.
[32] Park, Albert and Xinzheng Shi (2008). *Impact of education fee reductions on enrollment in Gansu*. Memo, University of Oxford and University of Michigan.
[33] Parker, Susan W. and Emmanuel Skoufias (2000). *The impact of PROGRESA on work, leisure and time allocation*. Report submitted to PROGRESA. IFPRI, Washington DC. October.
[34] Parker, Susan W. and Emmanuel Skoufias (2001). *Conditional cash transfers and their impact on child work and schooling: evidence from the Progresa program in Mexico*. Working paper, IFPRI, October.
[35] Paxson, Christina and Norbert Schady (2007). *Does money matter? The effects of cash transfers on child health and development in rural Ecuador*. Working Paper, Princeton University and World Bank.
[36] Rivera, J.A., D. Sotres-Alvarez, J.P. Habicht, T. Shamah and S. Villalpando (2004). *Impact of the Mexican program for education, health, and nutrition (Progresa) on rates*

of growth and anemia in infants and young children. Journal of the American Medical Association, 291, pp. 2564-2570.
[37] Rubalcava, L., G. Teruel, and D. Thomas (2004). *Spending, saving and public transfers paid to women*. Paper CCPR-024004. California Center for Population Research. On-line working paper series.
[38] Schady, N., M. Araujo (2006). *Cash transfers, conditions, school enrolment, and child work: Evidence from a randomized experiment in Ecuador*. Working paper 3930. The World Bank Policy Research, Washington DC.

[39] Schultz, T. Paul (2004). *School subsidies for the poor: evaluating the Mexican Progresa poverty program*. Journal of Development Economics, 74 199-250.
[40] Stecklov, G., P. Winters, M. Stampini, and B. Davis (2005). *Do conditional cash transfers influence migration? A study using experimental data from the Mexican PGROGRA program*.
Demography, 42, pp. 769-790.
[41] Stecklov, G., P. Winters, J. Todd, F. Regalia (2006). *Demographic externalities from poverty programs in developing countries: Experimental evidence from Latin America*. Working paper 2006-01, American University, Department of Economics.
[42] Thomas, D. (1990). *Intrahousehold resource allocation: An inferential approach*. Journal of Human Resources, 25, pp. 635-664.
[43] Todd, P.E. and K.I. Wolpin (2006). *Using a social experiment to validate a dynamic behavioral model of child schooling and fertility: Assessing the impact of a school subsidy program in Mexico*. American Economics Review, 96(5), pp. 1384-1417.
[44] Walle, Dominique van de and Ren Mu (2007). *Fungibility and the flypaper effect of project aid: micro-evidence for Vietnam*. Journal of Development Economics, 84, pp.667-685.
[45] Yap, Y., G. Sedlacek, and P. Orazem (2002). *Limiting child labor through behavior-based income transfers: An experimental evaluation of the PETI program in rural Brazil*. The World Bank, Washington DC.
[46] http://tieba.baidu.com/f?kz=306047423
[47] http://www.x-rates.com/d/CNY/table.html
[48] http://210.72.51.4/doc/ShowArticle.asp?ArticleID=3408

Chapter 3
Famine, fertility, and fortune in China

3.1 Introduction

There were numerous famines during the twentieth century[36]. The largest of these was China's 1959-1961 famine, which resulted in about 30 million excess deaths[37]. Previous research has focused on estimating excess mortality; however, there were undoubtedly important effects on the livings as well[38]. In particular, children born during the famine may have suffered from malnutrition in the initial years of life, resulting in adverse long term health effects and influencing economic and social attainments as adults. In this chapter, I investigate the long term effects of China's 1959-1961 famine on the adult education, labor market performance, and wealth of rural Chinese women and men who were exposed to the famine in utero and during the first two years of life.

Because of lack of food in the famine, children suffered from insufficient nutrition intake, which damaged their health. This health shock may have had a long-term effect simply because it persists over time. The health shock could also affect other outcomes, such as educational attainment and labor market performance, which help determine long-run well-being. On the other hand, estimates of famine impacts are complicated by potential selection bias. First, parents living through the famine could alter their fertility decisions. Parents unable to provide adequate support for children could choose to postpone childbearing, so that only parents with better endowments or better family background would still have children. Second, because of the negative impacts of famine on women's health, only those in good health were able to conceive. Both of these led to positive fertility selection. Third, weaker children were more likely to die during the

---

[36] Sen, 1981; Ravallion, 1987
[37] Ashton et al., 1984
[38] Steinet et al.,1975; Barker et al., 2005

famine, leading to another positive selection with respect to children's health. The combination of negative shocks and positive selection complicates estimates of famine effects. After the famine ended, parents postponing having children in the famine started to have children, and women in worse health in the famine resumed fertility ability; both of these led to negative fertility selection for children born directly after the famine. In this paper, I provide a conceptual framework to describe the above responses to the famine, and estimate the long term effects of the famine after controlling for positive fertility selection during the famine. I also provide some suggestive evidence of negative fertility selection right after the famine.

The data requirements for conducting such an investigation are considerable. Information is required on the extent of famine experienced by individuals many years ago, as well as detailed information on adult outcomes in the present day. I use provincial excess death rates in 1959-1961 to measure famine intensity. Outcome variables are from the China 2000 population census. Because the 2000 census data contains information about every individual's birth province, as well as her year and month of birth, it is possible to link each individual's adult outcomes with the famine intensity experienced in her childhood. Additionally, the availability of birth month makes it possible to control for positive selection from fertility. I have estimated the long term effects of the famine by comparing cohorts conceived before the famine, therefore not affected by the positive fertility selection due to the famine, but affected by the famine, i.e. born between February 1957 and June 1959, with those not affected by the famine, i.e. born between January 1954 and January 1957, and comparing individuals born in different provinces. Using a sample of rural women from the China 2000 population census data, after controlling for fertility selection, I find that women exposed to the famine with 10% higher death rate than normal in the first year of life completed 0.024 fewer years of schooling, had a 0.54% lower probability of completing high school, worked 0.014 fewer days per week, and lived in houses with 0.005 fewer rooms and 0.199 square meters less housing area per capita. I do not find significant effects of the famine on men, nor do I find significant effects of the famine on individuals exposed to it in utero or in the second year. In order to test positive fertility selection in the famine, I replace cohorts born between February 1957 and June 1959 with cohorts born between July 1959 and October

1962, i.e., those affected by the famine in the initial years of life and whose parents' fertility decisions were also affected by the famine, and repeat the analyses. I find that the famine effects become much weaker, which supports positive fertility selection during the famine. I then test for negative fertility selection after the famine by adding cohorts born after the famine to the first sample as a control group, and re-estimate the famine effects; I find that the significant negative effects of the famine become insignificant or weaker. This provides some suggestive evidence for the existence of negative fertility selection immediately after the famine.

Four other papers have focused on the long term effects of China's 1959-1961 famine. Using Chinese Household and Health Survey (CHNS) data, Luo et al. (2006) found that women exposed to greater severity of famine in early life were more likely to be overweight as adults. However, they only compared mean value of weight for different groups in different years without running any regressions, such that their results could not be thought of as precisely estimated famine effects. Using the same data set, Chen and Zhou (2007) found that cohorts exposed to the famine had a lower average height, less labor supply, and less income. Without detailed information of birth month in CHNS data, their measurement of famine severity, i.e. 1960 death rate, could not capture different exposure to the famine of cohorts born in different months. By analyzing China's 1990 population census data and the CHNS data, Meng and Qian (2006) found that exposure to the famine reduced height, weight, weight-for-height, head circumference, educational attainment, and labor supply. But in their paper, they did not look at different effects of famine on men and women. Almond et al. (2007) found that cohorts exposed to higher exogenous mortality in utero were more likely to be poorer, disabled, and illiterate; these cohorts were also more likely not to work, to have worse marriage market outcomes and have daughters. They also used China's 2000 population census data. Since the Great Famine in China lasted for three years, cohorts born in the beginning of the famine were affected by the famine in the first two years of life, which are also considered as a very important period for children's physical and psychological development. Therefore, just looking at in utero effects might understate the true effects of the famine. More importantly, all the above four papers did not deal with selection problem explicitly.

I use China’s 2000 population census data, which includes detailed information of birth month, making it possible to construct a famine intensity measurement precisely reflecting different cohorts’ different exposure to the famine in utero, in the first year and in the second year. Especially, I can use the information of birth month to deal with different fertility selections due to the famine, i.e. positive fertility selection in the famine and negative fertility selection immediately after[39]. To the best of my knowledge, it is the first time in the literature to deal with fertility selection problem in the estimation of the long run impacts of the famine. Furthermore, I investigate in-utero, first year, and second year effects separately, which helps to identify famine effects more precisely, and provides some evidence about which stages in early life are more important in children's development. Finally, I examine the impacts of the famine on men and women separately, which provides some evidence on gender inequality.

This paper is related to other research about the effects of famines in other countries. For example, Stein et al. (1975) investigated the long term consequences of malnutrition on 40,000 children conceived and born during the 1944-1945 Dutch famine, and found no lasting effects of prenatal malnutrition on health status or mental performance. However, they found negative effects of more prolonged malnutrition in postnatal life. Barker et al. (2005) found that survivors of the 1944-1945 Dutch famine were more likely to have cardiovascular disease. This paper is also related to research using non-Chinese data on the impact of early life conditions on adult outcomes. Almond and Chay (2003) found that better conditions in early life for U.S. black women led to better health in adulthood and higher birth weight of the women's children. Behrman and Rosenzweig (2004) used U.S. data to exploit intrauterine nutrient intake differences between monozygotic female twins. They found a strong impact of fetal growth on schooling attainment and height. Royer (2005) documented long-run and intergenerational effects of birth weight differences between twins. Almond (2006) found that U. S. cohorts who were in utero during the 1918 influenza pandemic had worse adult outcomes as they aged than cohorts born just before and after the pandemic in terms of educational attainment, physical disability, socioeconomic status and mortality.

[39] GØrgens et al. (2005) estimated the impacts of the Great Famine in China on height of cohorts after controlling for selection from children death. But they did not address selection from fertility.

Lindeboom et al. (2006) showed that poor macroeconomic conditions in early life reduced longevity in the Netherlands. Alderman et al. (2006) found that rainfall shocks and exposure to war affected early-life nutrition and later height and schooling levels of young adults in Zimbabwe. Maccini and Yang (2006) found that higher early-life rainfall had a positive effect on the adult outcomes of women but not of men, using Indonesian data.

The rest of this chapter is organized as follows: Section 3.2 provides background and description of the famine. Section 3.3 provides a conceptual framework for how the famine affected individuals' adult outcomes and how different selections in and after the famine affect the estimates. Section 3.4 describes the data sources, sample and variables used. Section 3.5 presents the empirical strategy and results. Section 3.6 discusses, and Section 3.7 concludes.

3.2 1959-1961 famine

After the People's Republic of China was founded in 1949, in order to quickly catch up to western nations, the government adopted a heavy-industry-oriented development strategy in 1952. In the same year, it started agricultural collectivization, replacing traditional family farms with collectively managed production teams. From 1952 to 1958, agricultural and industrial outputs increased continuously and dramatically. Prompted by this success and a desire to surpass developed countries, Chairman Mao Zedong launched the Great Leap Forward to accelerate economic growth. In rural areas, the People's Commune movement was also launched on a full scale in the summer of 1958. The People's Commune created huge collectives and eliminated all private ownership; it also provided free food through large commune dining halls. The Communist Party told the people that China would soon enter the communism stage, when people could get whatever they needed; however, only one year later, the country was in crisis.

Beginning in the winter of 1958, starvation was observed in Sichuan and Anhui provinces. By the spring of 1959, starvation became widespread. The estimated daily availability of food energy per capita during this period decreased considerably to about

1800 calories, reaching a low point of only 1500 calories in 1960[40]. A study using demographic data released after the start of economic reforms concluded that this crisis resulted in about 30 million excess deaths[41]. From Figure 3.1, we can see that the crude death rate[42] dramatically increased between 1959 and 1960. Facing such a severe famine, parents might decide not to have children or delay having children, and weaker women might not be able to conceive. According to Peng (1987), total fertility up to age 39 was about 5.6 births per woman in pre-famine years, but it dropped to its lowest historic level, 3.06, in 1961. From Figure 3.1, we can also see that the crude birth rate[43] dropped sharply between 1959 and 1961. Both the increase in death rate and the decrease in birth rate led to a dramatic decrease in population between 1959 and 1961, shown in Figure 3.2.

Although the famine occurred nationally in 1959-1961, there was large regional variation[44]. Table 3.1 shows the distribution of provincial death rates. We can also see this geographic variation in Figure 3.3. In 1959, the most severe famine occurred in the Sichuan province, where the death rate was 4.7%; however, the lowest death rate, which was only 0.78%, was observed in Shanghai City. In 1960, when the famine was the most severe nationally, variation across provinces was still large. The Anhui province became the province with the highest death rate (6.86%), while Shanghai City still had the lowest death rate (0.69%). In 1961, when the famine was nearing its end, the Sichuan province had the highest death rate of 2.94%, and Shanghai City still had the lowest death rate of 0.77%.

In January 1962, the government abandoned its radical policies. The emphasis was shifted from steel production back to agriculture. As a result, grain output started to grow in 1962; in the same year, the famine ended. From Figure 3.1, we can see that the

---

[40] Ashton et al., 1984

[41] Ashton et al., 1984

[42] Crude death rate is defined as number of deaths per 1000 people. I call it death rate for simplicity in this paper. The source and construction are discussed later.

[43] Crude birth rate is defined as number of births per 1000 people. I call it birth rate for simplicity in this paper. The source and construction are discussed later.

[44] Due to the government's preferential treatment of urban residents through a grain rationing system and the maintenance of government-controlled stockpiles, the lack of food was much more devastating in rural than in urban areas. Because of the variation in the proportion of rural population, population density, exposure to natural disaster, and provincial response to food shortages, exposure to the famine also varied greatly across provinces. (Ashton et al., 1984; Lin and Yang, 2000)

death rate resumed its normal level immediately when the famine ended in 1962. However, the birth rate did not return to its normal level right away. There was a dramatic increase in the birth rate in 1962, and it reached the highest point in 1963. There are two possible reasons: parents postponing having children in the famine might have started to have children when the famine ended, and those women not able to conceive during the famine recovered their fertility when they had enough nutritional intakes after the famine. This dramatic increase after the famine can be thought of as compensatory birth. After 1963, the birth rate decreased and went back to its normal level.

3.3 Conceptual framework

Starting with Grossman (1972), individual health is typically considered as a function of an initial health endowment and other factors[45]. And the initial health endowment is determined by genetic characteristics, which are determined at conception, and environmental conditions experienced in early life. In biology, the idea that the environmental conditions in a certain sensitive level period of life may have irreversible effects is known as “critical-period programming.” Biological[46] and economic studies[47] provide evidence for this idea. Since health status plays an important role in determining individuals' outcomes such as education, labor market performance and income[48]. Therefore, I am actually investigating the reduced relation between environmental conditions in early life and the adult outcomes.

In this context, special attention should be given to likely directions of any selections. Individuals could only be included in the dataset used (the China 2000 population census) if they were still alive in 2000. A potential worry is that the famine might affect the likelihood of survival through 2000, and those whose survival was induced by the famine could have different initial characteristics from the overall population of births in a locality in a particular year. The famine affected survival of children through three channels. First, during the famine, parents might have postponed having children if they felt unable to provide enough support for children. Second,

[45] Other factors might include historical health inputs, demographic variables, the time histories of community infrastructure, and the disease environment.
[46] See studies cited in (Barker, 1998; Diamond, 1991).
[47] See the papers cited in section 1.
[48] Behrman and Deolalikar, 1988; Currie and Madrian, 1999; Currie and Hyson, 1999.

women might not be able to conceive because of "famine amenorrhea"[49] due to insufficient nutritional intake and subsequent poor health during the famine[50]. Third, after children were born in the famine, weaker children might have died due to lack of nutritional intake. When the famine ended, parents with lower endowments and those who had postponed having children during the famine started to have children, and weaker women not able to conceive during the famine could conceive when they got enough nutrition after the famine. In other words, the famine induced sample selectivity, biasing the regression estimates. As it turns out, the famine led to positive selection for cohorts conceived in the famine and negative selection for cohorts conceived immediately after the famine.

The joint distribution of married couples' endowments, women's health and newborn children's health is denoted by $F(\phi, \mu, \psi)$. Higher values of ϕ represent more endowments. There is a threshold ϕ^*, only those parents with endowments higher than ϕ^* would like to have children, since they would be able to provide enough support for their children. Higher values of μ represent better women's health. We also have a threshold μ^*, only those women with health better than μ^* are able to conceive. Higher values of ψ represent better newborn children's health. There also exists a health threshold ψ^* such that children with health below this threshold could not survive to adulthood. In addition, there exists another threshold g above which children have "good" health. Then, the proportion of children born and surviving to adulthood λ is:

$$\lambda = P(\phi \geq \phi^*, \mu \geq \mu^*, \psi \geq \psi^*)$$

Suppose that an inter-temporal health linkage exists, so that better infant health leads to better adult health, and only those with initial health above g have good health in adulthood. Then, the fraction of adults in good health, $Health_{good}$, is:

---

[49] Forster and Ranum, 1975.

[50] See (Frisch and McArthur, 1974; Frisch, 1978, 2002) for review of biological evidence. See (Ford et al., 1989; Langsten, 1985; SG et al., 2007; Jowett, 1991) for description of "famine amenorrhea" in Bangladesh, China and the Netherlands.

$$Health_{good} = \frac{P(\phi \geq g)}{P(\phi \geq \phi^{*}, \mu \geq \mu^{*}, \psi \geq \psi^{*})}$$

During and after the famine, both the distributions and thresholds of parents' endowments, women's health and children's health were likely to change. The changes can be represented by Figure 3.4. The solid curve in Figure 3.4 can be thought of as the distribution of parents' endowments, or women's health, or newborn children's health.

As discussed above, the famine reduced parents' endowments, worsened women's health and also new-born children's health, which can be represented by the shift of the solid curve to the dashed curve in Figure 3.4. From Figure 3.4, we can see that the tail to the left of the thresholds, ϕ^{*}, μ^{*}, and ψ^{*}, becomes thicker. It means that more parents decided not to have children, more women could not conceive because of amenorrhea due to insufficient nutrition and subsequent worse health, and more new-born children could not survive because of worse health in the famine. Meanwhile, parents with "marginal" endowments might choose not to have children in tougher environments since the costs of having children might increase because it was expected to become harder to get necessary resources in the famine. Women with "marginal" health might not be able to conceive in the famine, since exogenous environments deteriorated so much that women needed better health to conceive. As well, infants with "marginal" health died in the famine, since exogenous environments deteriorated so much such that infants need better health to survive. All these three can be represented by the shifting of the thresholds in Figure 3.4. That is, the thresholds shift from ϕ^{*} to ϕ^{**}, μ^{*} to μ^{**}, and ψ^{*} to ψ^{**}.

The shift in these distributions led to decreased parents' endowments and women's health, both leading to worse children's health since family background and mothers' health were positively correlated with children's health. Combined with the direct effects of famine on children's health, the fraction of children in good health decreased, i.e. $P(\phi \geq g)$ decreased. However, the shift in the thresholds had different implications. With the shift of the threshold of parents' endowments, richer parents could still have children. In addition, the shift in threshold of women's health made women in better health able to conceive. Higher endowments of parents and better health of mothers led to better health of children. In addition, the shift in the threshold of children's initial health caused children in better health to remain alive. We can see from the expression of the fraction of individuals with good health in the population, that the shifts in the thresholds led to increases in ϕ^*, μ^*, and ψ^*, so $P(\phi \geq \phi^*, \mu \geq \mu^*, \psi \geq \psi^*)$ decreased. The effects of shifts in the thresholds i.e. decrease of $P(\phi \geq \phi^*, \mu \geq \mu^*, \psi \geq \psi^*)$ in the denominator, offset the effects of shifts in distributions (i.e. decrease of $P(\phi \geq g)$ in the numerator). Therefore, for us to observe that more severe famine experienced in early life is associated with worse health (and worse adult outcomes) in later life, deterioration in the distribution of initial health must overwhelm the second effects (selection effects). If we do find that more severe famine in early life is negatively associated with later-life health and other adult outcomes, then the existence of these selection effects leads these effects to be a lower bound of the true causal effects.

When the famine ended, the distributions and thresholds of parents' endowments, women's health, and newborn children's health should shift back. That is, the dashed curve in Figure 3.4 should shift back to the solid curves, and all three thresholds, ϕ^{**}, μ^{**} and ψ^{**}, should also shift back to ϕ^*, μ^* and ψ^*, respectively. For those parents who were supposed to have children in the post-famine period, the proportion of children with "good" health, $Health_{good}$, should be the same as that before the famine. However, those poorer parents postponing having children in the famine might decide to have children, and weaker women who could not conceive because of insufficient nutrition intake in the famine could conceive after the famine. Their children are expected to have worse health. Therefore, compared with the pre-famine period, there should be more "low-quality" children born immediately after the famine. This is negative fertility selection. However, there was no selection due to death of children during this period, since the threshold above which children could survive in the post-famine period was the same as that in the pre-famine period.

3.4 Data sources and sample composition

3.4.1 Population census data

The primary data used in this paper is from a 0.095% random sample drawn from the China's 5th National Population Census conducted by the China National Bureau of Statistics in 2000. This sub-sample includes 1,180,111 observations, covering all of China's 31 provinces[51]. There are 604,050 (51%) men and 575,769 (49%) women in the sample.

Because of the government's preferential treatment of urban residents, the effects of the famine on urban residents were not nearly as pronounced as on rural residents. In addition, children's educational attainment was disrupted by the Great Proletarian Cultural Revolution[52] from 1966 to 1976 in urban regions[53]. I therefore restrict the study to individuals born in rural areas. However, although the census data includes information about individuals' birth provinces, it does not record whether these individuals were born in rural or urban regions. I therefore restrict the sample to those individuals who had a rural *hukou* (household registration booklet) in 2000[54], which accounts for 75.2% of the original sample.

Chongqing was an independent municipality directly under the jurisdiction of the central government in 2000, but it was a city of the Sichuan province before 1997. Therefore, in this paper, I treat Chongqing as a part of Sichuan province, but not as an independent municipality. Because of limitations in death rate data, I removed people born in Hainan and Tibet provinces. Individuals in these two provinces only account for 0.78% of the total sample, so dropping them should not lead to a selection problem. Since I am investigating the effects of early life conditions on individuals' adult outcomes, only

---

[51] The population census does not include Hongkong, Macao, and Taiwan.

[52] The Great Proletarian Cultural Revolution in the People's Republic of China was a struggle for power within the Communist Party of China that manifested as wide-scale social, political, and economic chaos, and grew to include large sections of Chinese society; it eventually brought the entire country to the brink of civil war. (See http://en.wikipedia.org/wiki/Cultural_Revolution)

[53] Gregory and Meng, 2002a, 2002b; Giles et al., 2007.

[54] This restriction excludes those who changed their *hukou* from rural to urban, leading to selection bias. However, it is very hard to change *hukou* in China; only upper class people, such as college graduates, are likely to change their *hukou*. As shown below, the famine has negative effects on cohorts affected, which means the ratio of upper class people in the cohorts not affected by the famine is higher, then hypothetically there are more upper class persons changing their *hukou* in cohorts not affected by the famine; therefore, the exclusion of those changing their *hukou* leads to downward bias of estimated famine effects.

those born between January 1954 and December 1966, i.e. 34 to 46 years old in 2000, are used for analysis.

The final sample used in this paper therefore includes individuals born between January 1954 and December 1966 having a rural *hukou* in 2000, excluding individuals from Hainan and Tibet. Table 3.2 shows the distribution of the sample by birth province.

Population census data includes individuals' birth province, birth year, and birth month. I use the birth province to link adults' outcomes with provincial famine shocks in early life. The birth year and birth month can be used to more precisely identify the timing of when individuals were affected by the famine. In this paper, I investigate the effects of the famine on three groups of outcome variables. The first group is human capital variables, including years of schooling and high school indicator; the second group is labor market performance, including employment indicator and work days per week of those employed; since there is no variable of individual income in population census data, the third group of outcome variables include household wealth per capita, measured by average number of rooms per capita and average housing area per capita. For specific variable definitions, see the Data Appendix.

3.4.2 Famine intensity data

The measurement of famine intensity is generated from provincial death rates in different years. I obtain provincial death rates from *China Compendium of Statistics: 1949-2004*, which was compiled and published by the China National Bureau of Statistics. In this paper, I use excess death rate (EDR) to measure famine intensity in 1959-1961. In order to estimate excess death rate for each province-year from 1959-1961, I first calculate the average death rate in 1954-1958, $\overline{death\_rate}_{5458}$, and the average death rate in 1962-1966, $\overline{death\_rate}_{6266}$, for each province. I then calculate provincial predicted death rates for the three famine years in the following way:

$$provincial\_predicted\_death\_rate_{jz} = \overline{death\_rate}_{5458,j} + \frac{\overline{death\_rate}_{6266,j} - \overline{death\_rate}_{5458,j}}{1962-1958} * (z-1958)$$

Here, j represents province j, z=1959, 1960 and 1961.

The excess death rates in 1959-1961 are defined as the percentage by which the death rate in each province exceeds provincial predicted death rate. Specifically, the variable is the natural log of death rate in the famine year minus the natural log of provincial predicted death rate, as shown in the following:

$$EDR_{jz} = \ln(DR_{jz}) - \ln(provincial\_predicted\_death\_rate_{jz})$$

EDRs in other years are zero, as defined. Table 3.3 shows the estimated excess death rate for different provinces in 1959-1961. From the bottom row, we can see that, nationally, the highest excess death rate was in 1960 (15.41%), and the lowest excess death rate is 5.17%, in 1961. This is consistent with the famine pattern, i.e. the most severe famine happened in 1960, and the famine was coming to its end in 1961.

Cohorts born in different months experienced the famine at different times, even within the same year. I therefore calculate famine intensity experienced in utero, the first year and the second year, respectively, for cohorts born in province j in month m of year t and experiencing famine in year z and z+1 using the following method:

$$FI_{jmt} = \frac{(months\_in\_year\_z)_{mt}}{N} * EDR_{jz} + \frac{(months\_in\_year\_z+1)_{mt}}{N} * EDR_{j,z+1}$$

Here, $(months\_in\_year\_z)_{mt}$ and $(months\_in\_year\_z+1)_{mt}$ are the number of months in utero (in the first, second year) spent in year z and year z+1 by cohorts born in month m of year t. N is set to 10 when I calculate famine intensity experienced in utero, and to 12 when I calculate famine intensity experienced in the first and second year. z is equal to 1959, 1960 and 1961.

Table 3.4 reports selected summary statistic of the sample for analysis. Since I restrict the sample to cohorts born between 1954 and 1966, the average ages of men and women were both 39.48 in 2000. Ninety-nine percent of women were married in 2000, but only 94% of men were married in the same period. Women's education level was lower than men's; only 5% of women completed high school, and on average, they finished 6.72 years of schooling. In contrast, 12% of men completed high school, with an average of 8.07 years of schooling. Men also performed better than women in the labor market: 94% of men were employed from October 25 to 31 in 2000, but only 85% of women were employed during this period. Among those employed, men worked 6.05 days and women worked 5.86 days from October 25 to 31 in 2000. However, women's living conditions, which are used to measure wealth, were slightly better: women were living in households with 0.85 average rooms per capita and 25.62 square meters average housing area per capita, while men had 0.82 average rooms per capita and 24.47 square meters average housing area per capita. The last three rows show the famine intensity experienced in utero, in the first year and in the second year for all those experiencing famine during these early years. Table 3.4 shows that, on average, the famine intensities experienced in these three stages are almost the same: 0.27 for in-utero, 0.29 for the first year, and 0.31 for the second year.

3.5 Empirical strategy and results

3.5.1 Empirical strategy

In examining the relationship between famine experienced in early life and adult outcomes, I seek to isolate different effects of the famine on the cohorts experiencing famine in utero, in the first year and in the second year. I also seek to isolate deviation of adult outcomes from the mean in one's birth province, as well as from the mean of the national birth cohort. Because particular provinces in China may be subject to slow-moving changes over long periods of time (reflecting, for example, different rates of economic development), I also try to isolate variations in a person's outcomes that diverge from the long-running trends in her birth province.

In this paper, I estimate the following reduced-form linear relationship between adult outcomes Y_{ijmt} of an adult i born in province j in month m of year t:

$$Y_{ijmt} = \beta_0 + \beta_1 * UFI_{jmt} + \beta_2 * FFI_{jmt} + \beta_3 * SFI_{jmt} + \lambda_{jt} TREND + \mu_j + \delta_t + \varepsilon_{ijmt}$$

(3.1)

The coefficients of interest are β_1, the impact of famine intensity experienced in utero UFI_{jmt} on adult outcomes; β_2, the impact of famine intensity experienced in the first year FFI_{jmt} on adult outcomes; and β_3, the impact of famine intensity experienced in the second year SFI_{jmt} on adult outcomes. $\lambda_{jt}TREND$ is a linear time trend specific to the province, which absorbs the long-run linear trends in the outcomes that may vary depending on the province (TREND is a linear time trend, and the coefficient λ_{jt} allows the time trend to vary across provinces). μ_j is birth province fixed effect, δ_t is birth year fixed effect, and ε_{ijmt} is an individual-specific error term. Serial and spatial correlation is likely to be a problem in this setting, biasing the OLS standard error estimates downward[55]. In particular, the concern is about correlation among the error terms of individuals experiencing the same or similar measured famine intensity. Standard errors allow for an arbitrary variance-covariance structure within birth provinces (standard errors are clustered by birth province).

Figure 3.5 shows different cohorts and the timings of when they were affected by the famine. Group (1) includes cohorts born between January 1954 and January 1957, who were not affected by the famine in utero or in the first two years. Group (2) includes cohorts born between February 1957 and January 1958, who were affected by the famine only in the second year. Group (3) includes cohorts born between February 1958 and January 1959, who were affected by the famine in the first and second years. All individuals in above three groups were conceived before the famine. Group (4) includes cohorts born between February 1959 and June 1959, and group (5) consists of cohorts born between July 1959 and October 1959. Both of these two groups were conceived before the famine, and were affected by the famine in utero, and in the first and second years. The difference between them is that cohorts in group (4) were affected by the famine at least five months after they were conceived, but cohorts in group (5) were affected by the famine in the first five months in utero. Group (6) includes cohorts born between November 1959 and October 1962; these were conceived in the famine and affected by the famine in utero and in the first and second years. Group (7) includes cohorts born between November 1962 and December 1966, who were conceived after the famine when compensatory birth occurred.

First, I estimate equation (3.1) using groups (1), (2), (3) and (4). All the cohorts in groups (1), (2) and (3) were conceived before the famine. Cohorts in these three groups

[55] Moulton, 1986; Bertrand et al.,2004

were not affected by the famine in utero, but cohorts in group (2) and (3) were affected by the famine in the first and second years. Cohorts in group (4) were also conceived before the famine, but affected by the famine at least 5 months after they were conceived, when they could not be aborted by their parents. Therefore, their parents' fertility decisions were not affected by the famine and women's health were also not affected by the famine when those children were conceived. As a result, estimates using these groups are free from bias due to selective fertility. Here, groups (2), (3) and (4) are treatment groups and group (1) is a control group. We should expect to see negative effects estimated using this sample group. However, the existence of positive selection from excess death could not be controlled, which makes the estimates here a lower bound of the famine effects. Figure 3.6 shows the hypothetical results; outcomes of cohorts in groups (2), (3) and (4) are expected to be lower than those of group (1).

In order to examine how positive fertility selection affects the estimates, I estimate equation (3.1) using groups (1), (5), and (6). Although cohorts in group (5) were conceived before the famine, they were affected by the famine in the first five months in utero, when their parents could still choose to abort them. Cohorts in group (6) were conceived during the famine. Therefore, cohorts in these two groups should be affected by their parents' selective fertility decisions; cohorts in group (6) were also affected by selection from their mothers' fertility abilities in the famine. At this time, groups (5) and (6) are treatment groups, and the control group is still group (1). Since selection effects from excess death on both the estimates using the first sample and the estimates using this sample cannot be controlled, differences in these two sets of estimates should mainly be driven by positive fertility selection. If we see weaker estimates of famine effects from this sample compared to effects from the first sample, we have evidence for the existence of positive fertility selection in the famine. Figure 3.6 shows the case that the outcomes of cohorts in groups (5) and (6) exceed the outcomes of cohorts in group (1); this is possible if positive fertility selection is strong enough to offset the negative effects of the famine.

As discussed in Section 3, a larger portion of children born after the famine are expected to have lower initial health and worse outcomes in adulthood. In order to test this negative selection, I add group (7) to the first sample. That is, I estimate equation

(3.1) using groups (1), (2), (3), (4), and (7). This time, cohorts in groups (2), (3) and (4) are treatment groups, but those in groups (1) and (7) are control groups. Figure 3.6 shows that, if negative selection existed, adding group (7) would make the estimated famine effects less strong than those estimated using the first sample.

When discussing the magnitude of the estimated famine effects, I focus on the impact per 0.1 unit change in the famine intensity, meaning I examine the impact per 10% by which the death rate in the famine exceeded the predicted normal level.

3.5.2 Effects of famine controlling for fertility selection

Table 3.5 presents regression results from estimating equation (3.1) for a variety of outcome variables using the first sample described above, i.e. cohorts born between January 1954 and June 1959. This table is divided into upper and bottom panels. In the upper panel, estimates are from regressions using the sample of women. In the bottom panel, the estimates come from regressions using men. For each outcome in each panel, the coefficients on famine intensity experienced in utero, the first year, and the second year are presented separately. Standard errors are presented in parentheses. For brevity, regression coefficients for the constant term, the large numbers of fixed effects and the provincial linear time trends are not shown. This format will be followed in subsequent results in Tables 3.6, 3.7, 3.8, 3.9, 3.10, and 3.12.

Table 3.5 shows in the upper panel that experiencing famine in utero and in the second year did not have any significant effects on women's adult outcomes, but experiencing famine in the first year had significant negative effects on adult outcomes. This shows that nutrition intake in the first year of life has the most important effects[56]. Since only estimates of the first year effects are significant, the following discussion focuses on the first year effects.

Experiencing famine in the first year has negative effects on women's educational attainment. I use two variables to measure educational attainment. One is schooling years,

[56] This might be because children grow fastest in the first year, according to the *Children Growth Chart* published by the US Center for Disease Control and Prevention in 2000. The growth rate of girls' weight is 186% in the first year, but drops to 26.2% in the second year. The girls' height growth rate is 48.7% in the first year, and drops to 17.2% in the second year. This growth pattern is the same for boys. Their weight increases by 188% in the first year, and drops to 21.7% in the second year. Boys' height growth rate is 50.6% in the first year, and only 16% in the second year.

and the other is an indicator of completing high school. Both coefficients are negative and statistically significant. The coefficient of schooling years is significant at the 5% level, and the coefficient of the high school indicator is significant at the 1% level. In addition, the magnitudes of these two coefficients are large. With 0.1 unit increases in the famine intensity experienced in the first year, women exposed to the famine in the first year completed 0.024 fewer schooling years and were 0.54% less likely to complete high school.

The second set of variables is labor market performance. I also use two variables to measure individuals' labor market performance: an employment indicator about whether the individual was employed from Oct. 25 to Oct. 31 in 2000, which has a positive, but not significant coefficient, and work days of those employed in the same period in 2000, which has a negative coefficient and is statistically significant at the 10% level. A 0.1 unit increase in famine intensity in the first year led to 0.014 fewer work days for employed women. The famine might affect women’s employment through two channels. The negative effects of the famine on women directly reduced their probability of employment because of poor health and a subsequent lower education level. However, lower wealth due to negative famine effects, which will be shown soon, might give women incentives to work outside the home in order to increase the family income, which increases the ratio of employed women in the population. The insignificant, but positive coefficient of employment indicator might reflect that the second effect slightly exceeded the first.

In addition to educational attainment and labor market performance, I also investigate the effects of the famine on wealth. I use average rooms per capita and average housing area per capita to measure wealth. The coefficients of both of these variables are negative and statistically significant. The coefficient of average rooms per capita is significant at the 1% level, and the coefficient of average housing area per capita is significant at the 10% level. With a 0.1 unit increase in famine intensity, average rooms per capita decreases by 0.005, and average housing area per capita decreases by 0.199. These two variables represent household level wealth; there are two possible ways for the famine effects on women to affect household level variables; one is in direct way, women affected by the famine had lower education and less work days as shown above, which reduced the income they could earn and contribute to the family; the other might go through an indirect way, women exposed to the famine in early life might marry to men with lower income, which also reduced the household wealth.

The bottom panel in Table 3.5 shows the regression results for men. We can see that most of the first year effects (effects on schooling year, high school indicator, employment indicator, and average housing area per capita) are still negative, but no coefficients are statistically significant. With a 0.1 unit increase in famine intensity, men experiencing the famine in their first year of life completed 0.014 fewer years of schooling, were 0.07% less likely to complete high school and 0.03% less likely to be employed, and lived in a household with 0.1 square meters less housing area per capita. The famine effects on men are much weaker than on women. This is consistent with a gender bias model in which available resources are given to boys in bad times[57].

One might be concerned that, since the selection from excess death could not be controlled for here, the insignificant effects of the famine on men might be driven by stronger selection effects on men if boys were more likely to die than girls during the famine. In order to test this, I regress a female indicator on famine intensity experienced in utero, in the first year, and in the second year after controlling for birth province fixed effects, birth year fixed effects, and provincial time trends. Regression results are shown in Table 3.A in Appendix. Results showed that effects of famine experienced in utero, in the first year and in the second year on the proportion of women in the sample are not significant. This means that there is no systematic difference between survival of boys and girls in different provinces in the famine. Therefore, the insignificant effects of the famine on men should not be driven by the stronger selection effects of the famine on men.

3.5.3 Effects of famine allowing for positive fertility selection

In Table 3.6, I investigate the effects of the famine on women and men, not controlling for positive fertility selection in the famine. We need to compare the results

[57] Dreze and Sen, 1989; Behrman, 1988; Rose, 1999; Alderman and Gertler, 1997; Duflo, 2003, Cameron and Worswick, 2001; Jayachandran, 2005; Maccini and Yang, 2006.

shown in this table with the corresponding results in Table 3.5. The discussion will remain focused on the first year effects.

For women, the coefficients of schooling years and the high school indicator are both negative, but become statistically insignificant. A 0.1 unit increase in famine intensity leads to 0.009 fewer schooling years and a 0.03% decrease in the probability to complete high school. These two coefficients in Table 3.5 are significant. With the same increase in famine intensity, the decreases in schooling years and probability to complete high school are smaller in Table 3.6. Therefore, we can see that allowing for positive fertility selection leads to insignificant famine effects on educational attainment, and the magnitude decreases.

For the labor market performance of women, the coefficient of work days is still negative and statistically significant at the 10% level, but the magnitude decreases compared with the corresponding coefficient in Table 3.5. A 0.1 unit increase in famine intensity leads to 0.01 fewer work days per week, which is smaller than the corresponding number in Table 3.5. The coefficient of employment indicator becomes negative, while it is positive in Table 3.5, and this coefficient in both Tables 3.5 and 3.6 is not statistically significant. The weaker effect of the famine on work days shows some evidence for positive fertility selection. As shown below, positive fertility selection leads to higher household wealth for women affected by the famine, which gives women less incentive to work outside the home, which provides a possible reason for the negative coefficient of the employment indicator in Table 3.6.

For women, the coefficients of average rooms per capita and average housing area per capita are positive and statistically insignificant. A 0.1 unit increase in famine intensity led to an increase of 0.002 average rooms per capita and 0.038 square meters average housing area per capita. Compared with those significantly negative effects shown in Table 3.5, the positive effects in Table 3.6 show that positive fertility selection effects are strong enough to offset the negative effects of the famine on cohorts exposed to it in the first year of life.

All the coefficients estimated using men sample are still insignificant. Although the effects of the famine on schooling year, high school indicator, and average housing area per capita become weaker, showing some evidence for positive fertility selection in the famine, the evidence is not as strong as that for women.

By comparing the results in Tables 3.6 and 3.5, especially those with women, we can see that positive fertility selection in the famine does exist and leads to a downward bias of the estimates of famine effects if we do not control for it.

3.5.4 Effects of famine allowing for negative fertility selection

When investigating the effects of the famine, previous researchers focused only on the effects of the famine on those cohorts who were directly affected. However, the cohorts born after the famine are expected to have lower quality due to negative selection after the famine since more parents with lower endowments decided to have children, and more women with weaker health conceived. This indirect effect of the famine has never been studied.

Table 3.7 shows the estimated results allowing for negative selection. We should also compare the results in Tables 3.7 and 3.5. We can see the results for women first: the coefficients of schooling year and high school indicator are still negative, and the coefficient of high school indicator is statistically significant at the 1% level. A 0.1 unit increase in famine intensity led to 0.017 fewer years of schooling, and 0.48% decrease in the probability of completing high school. However, in Table 3.5, both coefficients are statistically significant, and a 0.1 unit increase in famine intensity led to 0.024 fewer years of schooling and 0.54% decrease in the probability of completing high school. This shows that including cohorts affected by negative selection in the control group makes famine effects on educational attainments weaker, providing evidence for the existence of negative selection after the famine.

Regarding labor market performance of women, the coefficients of both the employment indicator and work days are not significant (Table 3.7). In addition, a 0.1 unit increase in famine intensity led to a 0.02% increase in the probability of being employed and 0.015 fewer working days for those employed. However, in Table 3.5, a 0.1 unit increase in famine intensity led to a 0.07% increase in the probability of being employed and 0.014 fewer work days. Although the coefficient of the employment indicator in both Tables 3.5 and 3.7 is not significant, the change of coefficient magnitude is consistent with our expectation. Since women affected by negative fertility selection might be more likely to work because of lower family wealth, including them in the sample would narrow the gap between the probability of being employed for women affected by the famine and women not affected by the famine in the first year. When I include cohorts affected by negative fertility selection, the coefficient of work days changes from significant in Table 3.5 to insignificant in Table 3.7, although the magnitude increases slightly. The results of labor market performance also show evidence for the existence of negative fertility selection after the famine.

The last two variables are average rooms per capita and average housing area per capita. In Table 3.7, the coefficients of these two variables for women are not significant. With a 0.1 unit increase in famine intensity, the average rooms per capita decreases by 0.003, and average housing area per capita decreases by 0.1. In Table 3.5, both of these coefficients are significant, and with the same increase in famine intensity, the average rooms per capita and average housing area per capita decrease by 0.005 and 0.199, respectively. The magnitude of both coefficients in Table 3.5 are larger than those in Table 3.7, which is consistent with our expectation that negative selection after the famine would make the estimates of famine effects weaker if we include cohorts conceived after the famine as a control group.

In Table 3.7, all the coefficients estimated using men remain insignificant. Although the effects of the famine on probability of being employed and average rooms per capita become weaker, showing some evidence for negative fertility selection in the famine, the evidence is not as strong as that shown for women.

Although I do not estimate exactly how large the negative selections are, the above comparisons, especially for women, provide suggestive evidence that negative selection did exist after the famine.

3.6 Discussion

3.6.1 Robustness

Because of lack of individual income in China's 2000 population census data, I use household wealth (measured by average number of rooms per capita and average housing area per capita) as a proxy. One concern might be that other members in the same family but not affected by the famine could also contribute to the household wealth, leading to downward bias of the estimated famine effects. In China, household heads and their spouses are usually thought of as main contributors of household income and therefore household wealth. In this section, I use a sub-sample, only including household heads and their spouses, to check whether the results found in the above are robust. Table 3.8 shows the results controlling for fertility selection, Table 3.9 the results allowing for positive fertility selection, and Table 3.10 the results allowing for negative fertility selection after the famine.

In Table 3.8, we can see that the effects of the famine on men are not significant. For women, only the effects of the famine experienced in the first year are significant, but the effects of the famine experienced in utero or in the second year are not significant. This pattern is the same as that shown in Table 3.5. Although the first year effects on schooling years appears insignificant, the P-value is actually 10.1%, very close to 10%. The significance of all other coefficients of first year effects is the same as in Table 3.5. With a 0.1 unit increase in famine intensity experienced in the first year, schooling years decreased by 0.017, the probability of completing high school decreased by 0.52%, work days decreased by 0.014, average rooms per capita decreased by 0.005 and average housing area per capita decreased by 0.175. Compared with Table 3.5, the magnitudes of the famine effects on different outcome variables are very similar. This shows that the famine effects controlling for fertility selection are robust to the change in the sample.

Moving from Table 3.8 to Table 3.9, which allows for positive fertility selection, we can see that all significant first year effects shown in Table 3.8 become insignificant in Table 3.9. For women, with a 0.1 unit increase in famine intensity experienced in the first year, schooling years decreased by 0.011, the probability of completing high school decreased by 0.02%, work days decreased by 0.01, average rooms per capita increased by 0.002, and average housing area per capita increased by 0.04. Compared with the results in Table 3.8, the famine effects become weaker if positive fertility selection is not controlled. For men, all coefficients are not significant, which is the same as in Table 3.8. The famine effects on schooling years, high school indicator, working days, and average housing area per capita become weaker. This still provides some evidence for positive fertility selection, but not as strong as for women.

Table 3.10 shows the results allowing for negative fertility selection. For women, except for the coefficient of high school indicator, all coefficients are not significant. A 0.1 unit increase in famine intensity led to 0.005 fewer schooling years, a 0.44% lower probability of completing high school, 0.015 fewer work days, 0.002 fewer average rooms per capita, and 0.069 square meters less average housing area per capita, and also led to a 0.02% increase in the probability of being employed. Compared with the results in Table 3.8, the famine effects become weaker, except for the effect on work days, which is almost the same as in Table 3.8. For men, the effects of negative selection are not as strong as for women. This is consistent with the results using the whole sample.

From the discussion above, we can see that the results found in this paper are robust with regard to the change in sample characteristics.

3.6.2 Effects of possible endogeneity of the 1959-1961 famine

A possible bias in the estimation might come from the possibility that the variation of famine intensity across provinces might be correlated with other time-specific variables, which can affect cohorts affected by the famine in early life through channels other than the famine.

In order to determine whether the variations of famine intensity were correlated with other variables, I perform a regression on the average provincial excess death rates in 1959-1961 and pre-famine provincial economic variables, political variables, and natural disaster intensities. Provincial economic variables include grain output per capita and GDP per capita. Provincial political variables include the ratio of Chinese Communist Party (CCP) members in the population and the ratio of members participating in the dining hall. I use the ratio of lands affected by natural disasters to measure provincial disaster intensity. Table 3.11 shows the regression results. We can see that: (1) the economic variables are not significant, and the joint F tests of economic variables are also not significant. This means that economic factors were not correlated with the variation of famine intensity across provinces. (2) Provincial disaster intensity is also not significant, indicating that natural disaster was not a reason for the occurrence and variation of the famine. (3) The ratio of CCP members in the population is significant at the 10% level in different specifications and the coefficients are negative[58]. However, the joint F tests of these two political variables are not significant, which shows that although variation of the famine intensity might be correlated with the ratio of CCP members in the population, it does not provide strong evidence that the famine was related to provincial political factors.

The correlation between the ratio of CCP members in the population and the variation of famine intensity leads to a bias in estimates in two situations: (1) if the ratio of CCP member in the population before the famine was also correlated with other big shocks, like the Great Proletarian Culture Revolution (1966-1976), which might also affect cohorts exposed to the famine in early life; (2) if the ratio of CCP members in the population before the famine affected adult outcomes of the cohorts affected by the famine in very early life through some other time-specific channels in 1959-1961 other than the famine. However, the famine was occurring nationally, and affected almost all factors in 1959-1961; it is very hard to find other time-specific channels not related to the famine but affecting cohorts exposed to the famine in the early years of life. As far as the Culture Revolution is concerned, if the negative effects of the famine shown in this paper are driven by the Cultural Revolution, then the estimates using urban samples should have the same pattern as those estimated using rural samples, since the Cultural Revolution definitely affected urban regions[59]. I therefore estimate equation (3.1) using an urban sample[60]. Table 3.12 shows the estimated results, which should be compared to those in Table 3.5. We can see women's results first. All the signs of the coefficients of the first year famine intensity are the opposites of those in Table 3.5, and the coefficients of schooling year and high school indicator are statistically significant. For men, except for the two coefficients of labor market performance, the signs of the coefficients of other variables are also the opposites of those in Table 3.5, although none of them are statistically significant. Results shown in Table 3.12 imply that the estimated famine effects shown in Table 3.5 were not driven by the Cultural Revolution; otherwise the

---

[58] Yang (1996) argued that having fewer CCP member represented lower political stands; in order to catch up to other provinces, these provinces would exaggerate grain outputs more than other provinces, so would be allocated more quota of procurements, which might cause higher mortality when the famine hit.

[59] Gregory and Meng, 2002a, 2002b; Giles et al., 2007

[60] Urban sample includes cohorts born between January 1954 and June 1959; individuals in this sample had non-agricultural *hukou* in 2000.

estimates using the urban sample should have the same pattern as those in Table 3.5.[61] Therefore, although the variation of famine intensity was correlated with the ratio of CCP members in the population in 1957, the probability that it affected the estimates is low.

Even though the variations of the famine intensity across provinces were totally random, the estimates could be biased if there were other post-famine programs targeting the specific population affected by the famine. For example, the government might invest more in provinces affected by the more severe famine. However, to the best of my knowledge, there were no such programs in China from the end of the famine to 2000.

3.6.3 Measurement issues

In this paper, I investigate the effects of the famine on individuals exposed to it in early life, especially in utero, in the first year and in the second year. Therefore, it might be better to use the provincial death rates of 0-2 year old children during the famine to construct famine intensity for each province. Because of insufficient data, I generate provincial famine intensities from crude death rates. The following analyzes the difference between the famine intensities created using crude death rates and those created using death rates of 0-2 year old children.

Let's assume that α is the ratio of 0-2 year old children in the population in the famine, and $\overline{\alpha}$ is the ratio of 0-2 year old children in the population in a normal year. dr_{02} is the death rate of children 0-2 years old, and dr_{other} is the death rate of persons in all other age cohorts in the famine. I also assume that $\overline{dr_{02}}$ and $\overline{dr_{other}}$ are death rates for 0-2 year old children and other age cohorts in normal years, respectively. In this paper, the famine intensity is calculated as:

[61] A possible reason for the estimated results in Table 12 could be as follows: since the Cultural Revolution was mainly a political movement, the ratio of CCP members in the population was positively correlated with the intensity of the Cultural Revolution in different provinces. So, the higher the ratio of CCP members in the population, the higher the intensity of the Cultural Revolution, which led to lower outcomes of the cohorts affected by it. As we see in Table 11, the ratios of CCP members in the populations of different provinces were negatively correlated with provincial famine intensity. Therefore, the correlation between famine intensity and the different outcomes in Table 13 turns out to be positive.

$$edr_{1,famine} = \frac{\alpha * dr_{02} + (1-\alpha) * dr_{other}}{\bar{\alpha} * \overline{dr}_{02} + (1-\alpha) * \overline{dr}_{other}} - 1 = \frac{dr_{02} * (\alpha + (1-\alpha) * \frac{dr_{other}}{dr_{02}})}{\overline{dr}_{02} * (\bar{\alpha} + (1-\bar{\alpha}) * \frac{\overline{dr}_{other}}{\overline{dr}_{02}})} - 1$$

If we measure the famine intensity using that of 0-2 year old children, the famine intensity should be :

$$edr_{2,famine} = \frac{dr_{02}}{\overline{dr}_{02}} - 1$$

We can see that whether $edr_{1,famine}$ over- or under-estimates $edr_{2,famine}$ depends on whether $\frac{(\alpha + (1-\alpha) * \frac{dr_{other}}{dr_{02}})}{(\bar{\alpha} + (1-\bar{\alpha}) * \frac{\overline{dr}_{other}}{\overline{dr}_{02}})}$ is larger or smaller than 1, which is indeterminate depending on the population structure and the magnitude of the death rate of children aged 0-2 years old relative to the death rates of other age cohorts, but not systematically correlated with $edr_{2,famine}$.

We can write $edr_{1,famine} = edr_{2,famine} + \varepsilon$, where ε is measurement error. Since there is no systematic difference between $edr_{1,famine}$ and $edr_{2,famine}$, this errors-in-variable problem leads to attenuation bias in OLS, i.e. the negative effects of famine are under-estimated.[62]

In this paper, I link famine intensity to different individuals using their birth provinces. This leads to another problem of measurement, since during the famine, children might migrate with their parents to other provinces where the famine was not as severe as in their birth provinces. Then, the famine intensity experienced by these children should be lighter than that linked to them by birth province. This is another source of downward bias in the estimated famine effects.

3.7 Conclusion

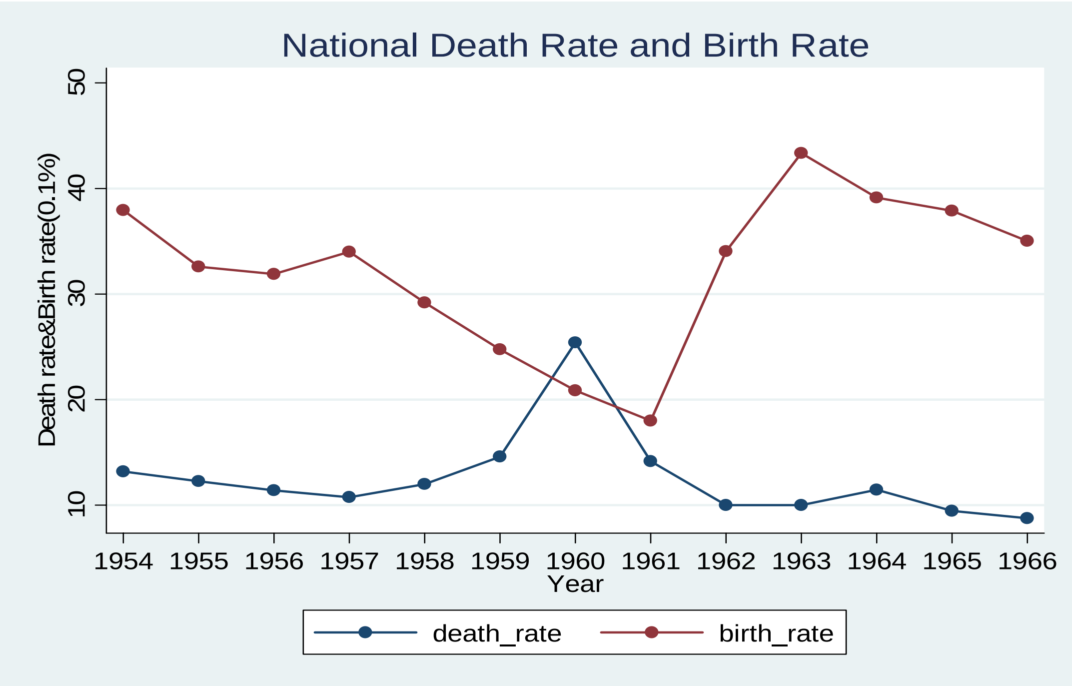
[62] Wooldridge, 2002

Famine in 1959-1961 caused about 30 million excess deaths, but little research exists about the long term effects of the famine. This chapter uses the China 2000 population census data to study this issue.

In this chapter, I propose a conceptual framework to demonstrate how fertility selection causes downward bias in the estimation of famine effects. Using a rural sample from the China 2000 population census data, after controlling for fertility selection, I find that experiencing the famine in the first year of life has negative effects on women's adult outcomes. Women exposed to the famine with 10% higher death rate than normal in the first year of life completed 0.024 fewer years of schooling, had a 0.54% lower probability of completing high school, worked 0.014 fewer days per week, and lived in houses with 0.005 fewer rooms and 0.199 square meters less housing area per capita. I do not find significant effects of being exposed to the famine in the first year for men. I also do not find significant effects of being exposed to the famine in utero or in the second year for either women or men. In this paper, I find that if fertility selection is not controlled for, the estimated effects of the famine become weaker. I also provide some suggestive evidence for negative fertility selection arising from compensatory birth immediately after the famine.

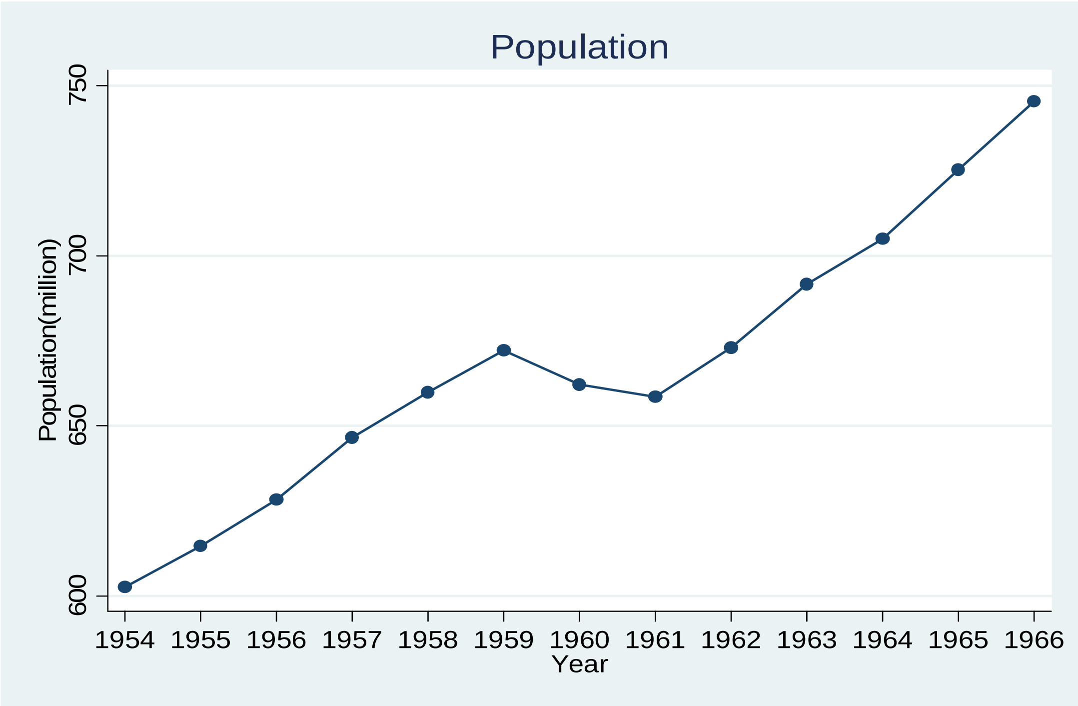
In addition to revealing the long-term effects of China's 1959-1961 famine, these results have important implications for policy. Our findings point to a group (newborn infants), which is particularly vulnerable to fluctuations in exogenous environments. The long-term effects of early-life conditions on schooling and other socioeconomic outcomes decades later should be included in the cost-benefit analyses of programs targeting this subpopulation. As such, our findings provide additional justification for interventions that protect infants from temporary shocks.

Figure 3.1 National death rate and birth rate (unit: 0.1%)



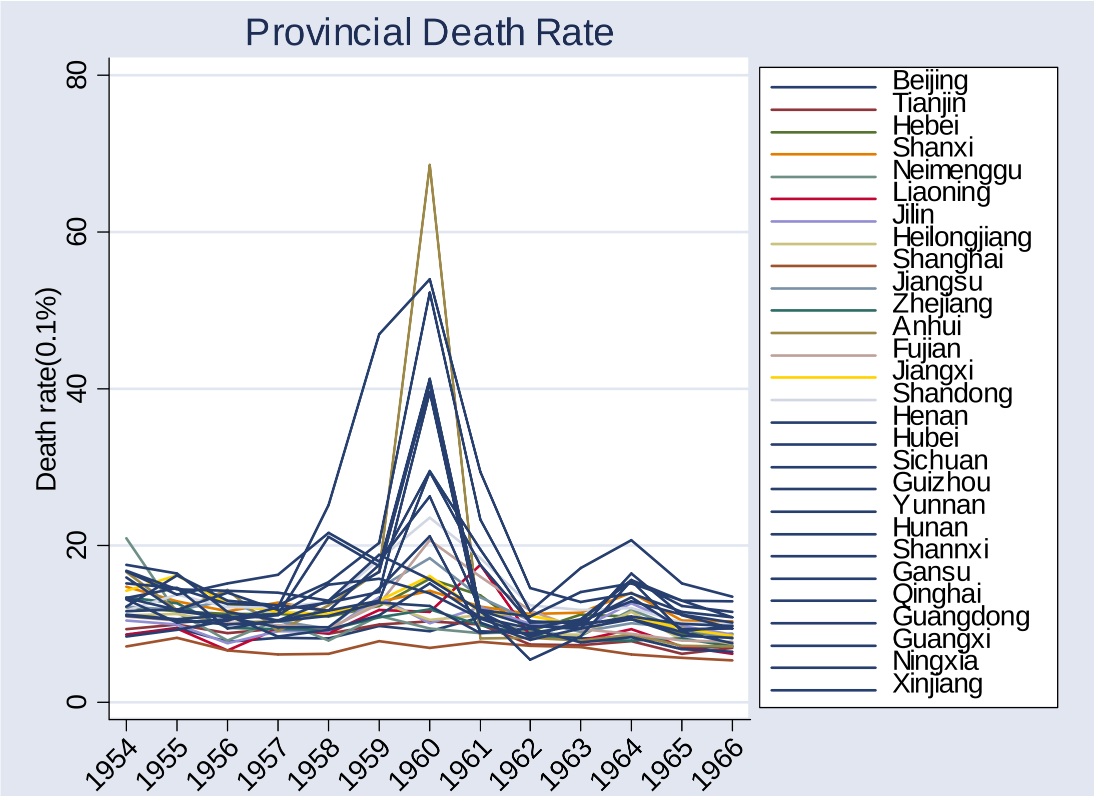
Source: China National Bureau of Statistics

Figure 3.2 Population



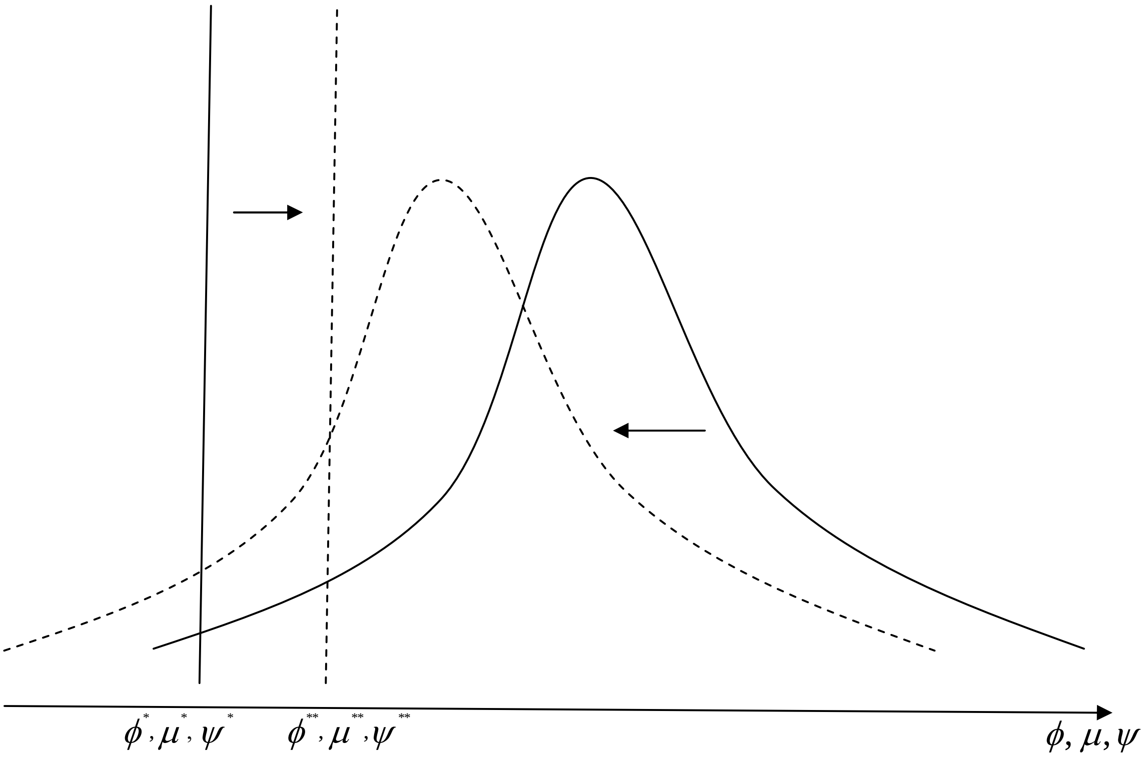
Source: China National Bureau of Statistics

Figure 3.3 Provincial death rates (unit: 0.1%)



Source: 1949-2004 China Statistical Data Compilation

Figure 3.4 Distribution and shift of parents' endowments, or women's health, or children's health.



μ

Figure 3.5 Famine timing

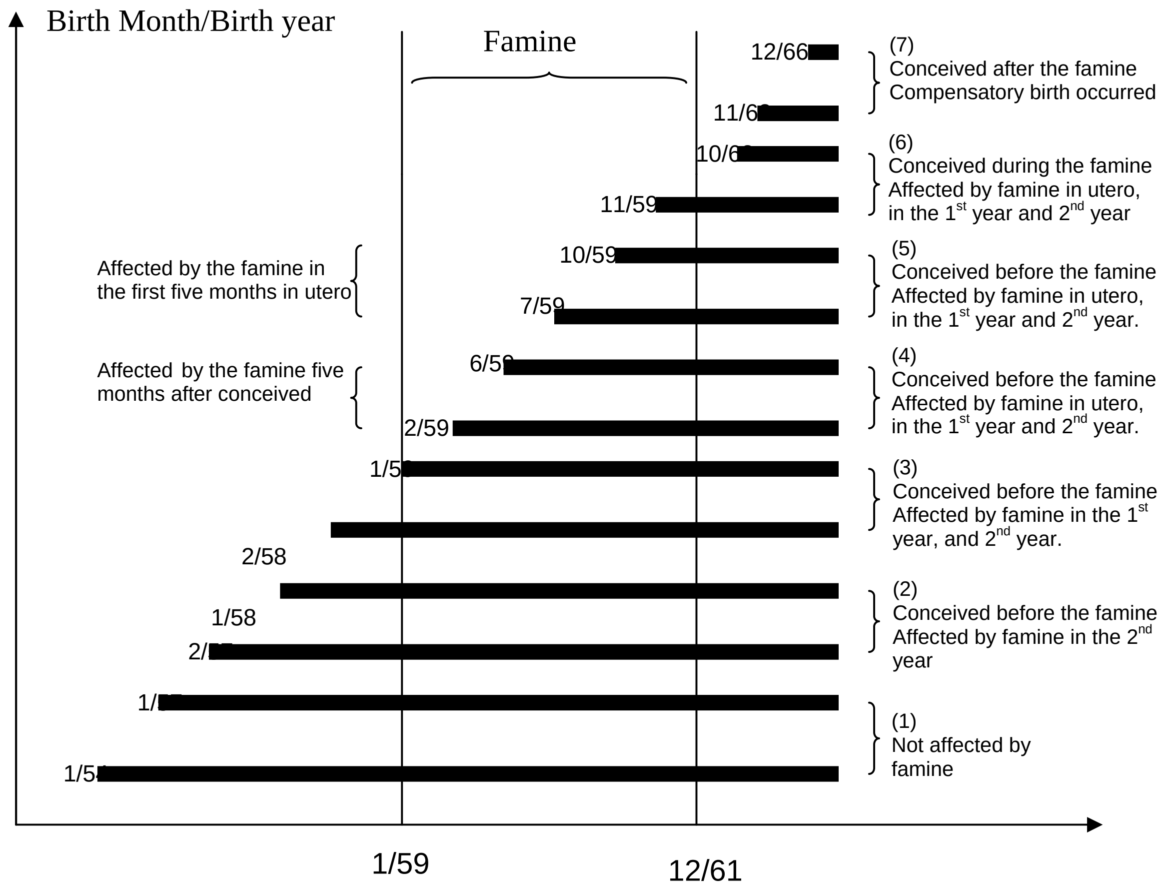
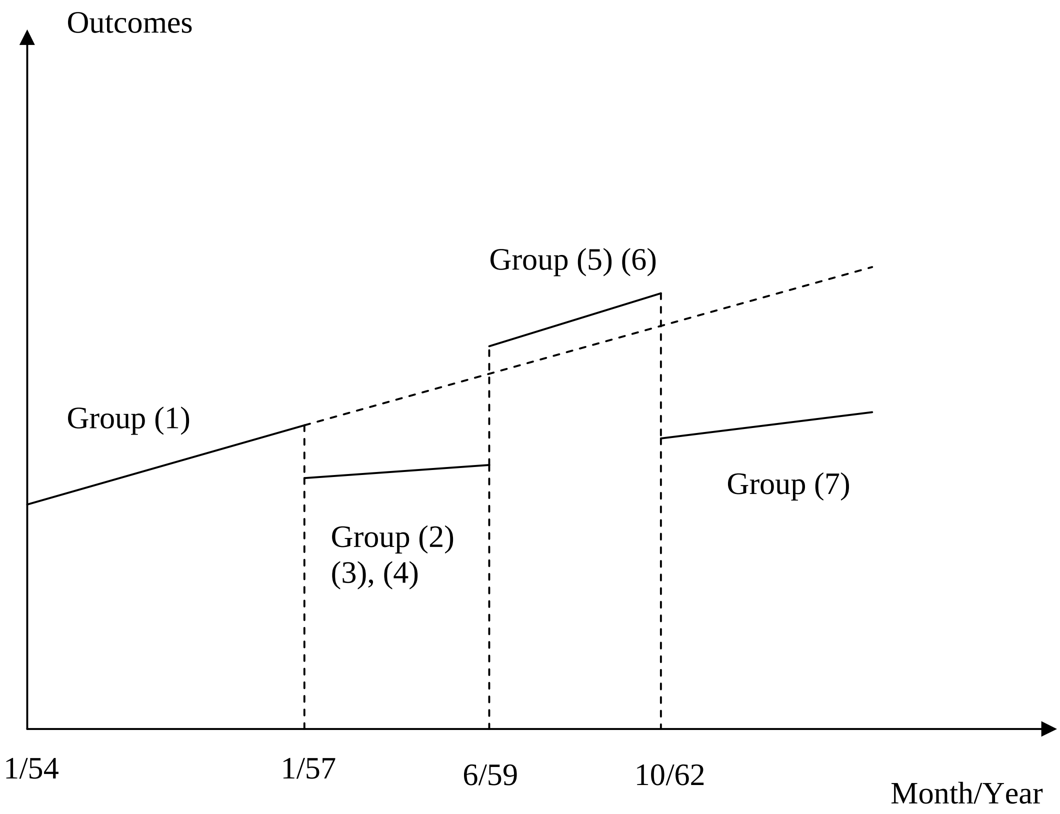


Figure 3.6 Hypothetical results



Note: month/year means birth month/birth year.

Table 3.1 Provincial death rate distribution in different years. (Unit: 0.1%)

Province	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966
Beijing	8.6	9.5	7.7	8.2	8.1	9.7	9.1	10.8	8.8	8.1	8.3	6.8	7.2
Tianjin	9.3	9.9	8.8	9.4	8.7	9.9	10.3	9.9	7.4	7.3	7.8	6.2	6.9
Hebei	12.1	11.6	11.3	11.3	10.9	12.3	15.8	13.6	9.1	11.2	10.9	8.7	8.7
Shanxi	14.7	12.9	11.6	12.7	11.7	12.8	14.2	12.2	11.3	11.4	14	10.4	10.3
Neimenggu	20.9	11.4	7.9	10.5	7.9	11	9.4	8.8	9	8.5	11.8	9.3	8.1
Liaoning	8.6	9.4	6.6	9.4	8.8	11.8	11.5	17.5	8.5	7.9	9.3	7.1	6.2
Jilin	10.4	9.9	7.5	9.1	9.1	13.4	10.1	12.1	10	9.4	12.6	9.7	8.6
Heilongjiang	11.1	11.3	10.1	10.5	9.2	12.8	10.5	11.1	8.6	8.6	11.5	8	7.4
Shanghai	7.1	8.2	6.6	6.1	6.2	7.8	6.9	7.7	7.2	7	6.1	5.7	5.3
Jiangsu	12.2	11.8	13	10.3	9.4	14.6	18.4	13.4	10.4	9	10.1	9.5	8.1
Zhejiang	13.4	12.6	9.5	9.3	9.2	10.8	11.9	9.8	8.6	7.9	7.9	8.1	7.1
Anhui	16.6	11.8	14.3	9.1	12.4	16.7	68.6	8.1	8.2	7.9	8.6	7.2	7.1
Fujian	10.9	10.4	10.2	9.8	9.4	12.5	20.7	16	11.7	9.3	8.7	7.9	7.7
Jiangxi	14.2	16.2	12.5	11.5	11.3	13	16.1	11.5	11	9.8	10.9	9.4	8.5
Shandong	11.7	13.7	12.1	12.1	12.8	18.2	23.6	18.4	12.4	11.8	12	10.2	9.9
Henan	13.3	11.8	14	11.8	12.7	14.1	39.6	10.2	8	9.4	10.6	8.5	8.2
Hubei	15.9	11.6	10.8	9.6	9.6	14.5	21.2	9.1	8.8	9.8	10.9	10	9.7
Hunan	17.5	16.4	11.5	10.4	11.7	13	29.4	17.5	10.2	10.3	12.9	11.2	10.2
Guangdong	11.2	10.6	11.1	8.4	9.2	11.1	15.2	10.8	9.4	7.6	8.3	6.8	6.4
Guangxi	15.2	14.6	12.5	12.4	11.7	17.5	29.5	19.5	10.3	10.1	10.6	9	7.5
Sichuan	8.4	9.2	10.4	12.1	25.2	47	54	29.4	14.6	12.8	13.9	11.5	10.8
Guizhou	12.2	16.2	13	12.4	15.3	20.3	52.3	23.3	11.6	17.1	20.7	15.2	13.5
Yunnan	16.7	13.7	15.2	16.3	21.6	18	26.3	11.8	10.9	14.1	15.2	13	10.8
Shannxi	11	10.5	9.9	10.3	11	12.7	12.3	8.8	9.4	10.6	15.6	13	12.9
Gansu	11.6	11.9	10.8	11.3	21.1	17.4	41.3	11.5	8.3	10.4	15.6	12.3	11.5
Qinghai	13.3	14.6	9.4	10.4	13	16.6	40.7	11.7	5.4	8.4	15.5	9.1	9.7
Ningxia	13.1	10.2	10.6	11.1	15	15.8	13.9	10.7	8.5	10.2	13.4	9.3	9.4
Xinjiang	16.8	14.4	14.2	14	13	18.8	15.7	11.7	9.7	9.4	16.4	11.1	9.4
Nation	12.8	12.0	10.8	10.7	12.0	15.1	23.2	13.1	9.5	9.8	11.8	9.4	8.8

Table 3.2 Distribution of sample used for analysis

Province	# of persons born in this province		
	Men	Women	Total
Beijing	415	432	847
Tianjin	430	422	852
Hebei	5692	5513	11205
Shanxi	2246	2304	4550
Neimenggu	1554	1504	3058
Liaoning	2649	2699	5348
Jilin	1651	1595	3246
Heilongjiang	1821	1862	3683
Shanghai	441	430	871
Jiangsu	5458	5353	10811
Zhejiang	4068	3957	8025
Anhui	4183	4115	8298
Fujian	2265	2296	4561
Jiangxi	2666	2795	5461
Shandong	7671	7849	15520
Henan	7150	7079	14229
Hubei	4060	4079	8139
Hunan	4957	4787	9744
Guangdong	4225	4164	8389
Guangxi	3321	3136	6457
Sichuan	8253	8224	16477
Guizhou	2314	2343	4657
Yunnan	2970	3044	6014
Shannxi	2781	2742	5523
Gansu	1830	1859	3689
Qinghai	280	286	566
Ningxi	297	302	599
Xinjiang	756	736	1492
Total	86,404	85,907	172,311

Note: (1) In this table, numbers for Sichuan include Sichuan Province and Chongqing Municipality, since Chongqing was separated from only after 1997, so this paper put them together.

Table 3.3 Provincial excess death rate in 1959-1961 (Unit:1%)

Province	1959	1960	1961
Beijing	4.89	3.99	5.39
Tianjin	4.90	5.15	4.60
Hebei	4.92	8.24	5.86
Shanxi	3.94	5.12	2.91
Neimenggu	3.48	1.89	1.31
Liaoning	6.93	6.34	12.05
Jilin	7.26	3.49	5.03
Heilongjiang	6.37	3.89	4.31
Shanghai	4.56	3.37	3.88
Jiangsu	7.37	11.04	5.92
Zhejiang	4.46	5.56	3.46
Anhui	9.17	61.30	1.04
Fujian	6.13	14.12	9.21
Jiangxi	4.45	7.59	3.04
Shandong	9.59	14.75	9.32
Henan	6.27	31.79	2.42
Hubei	7.00	13.55	1.30
Hunan	3.82	20.19	8.26
Guangdong	5.29	9.32	4.86
Guangxi	9.03	21.13	11.24
Sichuan	37.65	44.13	19.01
Guizhou	9.23	40.60	10.98
Yunnan	6.30	14.57	0.04
Shannxi	4.95	3.90	-0.25
Gansu	8.30	31.84	1.68
Qinghai	8.89	32.84	3.69
Ningxia	7.97	5.84	2.41
Xinjiang	8.98	5.88	1.88
National	7.58	15.41	5.17

Table 3.4 Summary statistics

	Woman					Man				
	Mean	S.D.	Max	Min	OBS	Mean	S.D.	Max	Min	OBS
Age	39.48	3.94	46	34	85907	39.48	3.93	46	34	86404
Marriage	0.99	0.08	1	0	85907	0.94	0.24	1	0	86404
High school(indicator)	0.05	0.23	1	0	85907	0.12	0.33	1	0	86404
Schooling years	6.72	2.97	19	0	85907	8.07	2.53	19	0	86404
Employment(indicator)	0.85	0.35	1	0	85907	0.94	0.23	1	0	86404
Work days	5.86	1.56	7	1	73387	6.05	1.38	7	1	81604
Average rooms per capita	0.85	0.56	27	0	85339	0.82	0.53	27	0	84604
Average housing areas per capita	25.62	18.88	490	0	85337	24.47	17.38	350	0	84601
Famine intensity in utero	0.27	0.33	1.89	-0.30	18755	0.27	0.33	1.89	-0.30	18595
Famine intensity in the first year	0.29	0.33	1.89	-0.30	17677	0.29	0.34	1.89	-0.30	17756
Famine intensity in the second year	0.31	0.35	1.89	-0.30	19784	0.31	0.34	1.89	-0.30	20440

Table 3.5 Effects of famine controlling for fertility selection

Women	**Dependent Variables:**					
	Education		Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	1.295	0.074	0.176	0.229	0.179	-1.059
	(0.831)	(0.079)	(0.206)	(0.196)	(0.190)	(6.089)
Famine intensity in the first year	-0.240	-0.054	0.007	-0.140	-0.053	-1.989
	(0.090)**	(0.008)***	(0.021)	(0.078)*	(0.015)***	(0.996)*
Famine intensity in the second year	0.161	0.012	0.011	0.077	-0.027	-0.169
	(0.159)	(0.017)	(0.014)	(0.092)	(0.022)	(0.561)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	32755	32755	32733	27839	32640	32639
R-squared	0.11	0.02	0.10	0.06	0.08	0.15

Men	**Dependent Variables:**					
	Education		Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	1.536	0.214	-0.080	0.067	-0.195	-1.112
	(0.984)	(0.133)	(0.060)	(0.205)	(0.129)	(3.537)
Famine intensity in the first year	-0.136	-0.007	-0.003	0.052	0.019	-1.002
	(0.230)	(0.018)	(0.008)	(0.082)	(0.033)	(0.699)
Famine intensity in the second year	0.317	0.022	-0.004	0.031	-0.014	0.826
	(0.200)	(0.025)	(0.006)	(0.053)	(0.015)	(0.979)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES

Observations	33229	33229	33208	31356	32746	32745
R-squared	0.07	0.02	0.09	0.04	0.08	0.14

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and June 1959.

Table 3.6 Effects of famine allowing for positive selection

Women	**Dependent Variables:** Schooling year	Education: High School	Labor Market Outcomes: Employment	Labor Market Outcomes: Work day	Housing: # of Rooms/family member	Housing: Area/family member
Famine intensity in utero	-0.132	-0.014	0.008	-0.016	-0.013	-0.481
	(0.091)	(0.009)	(0.010)	(0.037)	(0.015)	(0.613)
Famine intensity in the first year	-0.089	-0.003	-0.005	-0.109	0.022	0.379
	(0.154)	(0.010)	(0.009)	(0.056)*	(0.025)	(0.837)
Famine intensity in the second year	0.152	0.009	0.002	0.042	-0.020	-1.141
	(0.212)	(0.017)	(0.011)	(0.073)	(0.034)	(1.026)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	36731	36731	36804	31285	36666	36665
R-squared	0.15	0.03	0.10	0.06	0.09	0.16

Men	**Dependent Variables:** Schooling year	Education: High School	Labor Market Outcomes: Employment	Labor Market Outcomes: Work day	Housing: # of Rooms/family member	Housing: Area/family member
Famine intensity in utero	-0.090	-0.013	0.002	-0.046	0.002	-0.443
	(0.094)	(0.013)	(0.004)	(0.047)	(0.015)	(0.574)
Famine intensity in the first year	-0.058	0.009	-0.005	0.040	0.018	0.589
	(0.106)	(0.015)	(0.006)	(0.050)	(0.017)	(0.574)
Famine intensity in the second year	-0.043	-0.005	0.002	-0.018	-0.051	-0.972
	(0.129)	(0.018)	(0.011)	(0.083)	(0.034)	(0.911)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES

Observations	36351	36351	36422	34329	35844	35842
R-squared	0.08	0.03	0.09	0.04	0.09	0.15

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and January 1957, and cohorts born between July 1959 and October 1962.

Table 3.7 Effects of famine allowing for negative selection

Women	**Dependent Variables:**	Education	Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	1.076	0.056	0.162	0.230	0.187	-3.325
	(0.826)	(0.074)	(0.201)	(0.199)	(0.209)	(6.891)
Famine intensity in the first year	-0.169	-0.048	0.002	-0.148	-0.027	-0.997
	(0.122)	(0.011)***	(0.022)	(0.099)	(0.018)	(1.061)
Famine intensity in the second year	-0.025	0.007	-0.001	0.086	-0.016	-0.263
	(0.135)	(0.014)	(0.012)	(0.064)	(0.021)	(0.808)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	68719	68719	68697	58781	68221	68219
R-squared	0.15	0.02	0.10	0.06	0.10	0.16

Men	**Dependent Variables:**	Education	Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	1.287	0.230	-0.060	0.000	-0.161	0.640
	(0.971)	(0.147)	(0.057)	(0.195)	(0.160)	(3.652)
Famine intensity in the first year	-0.196	-0.030	0.001	0.047	0.021	-1.210
	(0.260)	(0.019)	(0.008)	(0.069)	(0.031)	(0.837)
Famine intensity in the second year	0.020	0.012	0.005	-0.036	0.008	0.583
	(0.112)	(0.019)	(0.007)	(0.036)	(0.015)	(0.874)
Provincial time trend	YES	YES	YES	YES	YES	YES

Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	69438	69438	69417	65677	67958	67956
R-squared	0.07	0.02	0.08	0.04	0.10	0.13

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and June 1959, and cohorts born between November 1962 and December 1966.

Table 3.8 Effects of famine controlling for fertility selection (Robustness)

Women	**Dependent Variables:**					
		Education	Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	1.268	0.094	0.176	0.244	0.185	0.052
	(1.209)	(0.112)	(0.183)	(0.199)	(0.200)	(6.682)
Famine intensity in the first year	-0.170	-0.052	0.008	-0.137	-0.046	-1.753
	(0.100)	(0.008)***	(0.020)	(0.076)*	(0.017)**	(0.955)*
Famine intensity in the second year	0.144	0.013	0.010	0.090	-0.027	-0.050
	(0.164)	(0.016)	(0.015)	(0.087)	(0.021)	(0.534)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	31605	31605	31605	26894	31587	31586
R-squared	0.11	0.02	0.10	0.06	0.08	0.15

Men	**Dependent Variables:**					
		Education	Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	2.254	0.274	0.008	0.102	-0.199	-1.690
	(1.497)	(0.177)	(0.052)	(0.164)	(0.123)	(3.447)
Famine intensity in the first year	-0.055	0.001	0.001	0.025	0.038	-0.681
	(0.247)	(0.019)	(0.008)	(0.087)	(0.033)	(0.743)
Famine intensity in the second year	0.242	0.022	-0.005	0.052	-0.021	0.738
	(0.203)	(0.025)	(0.005)	(0.057)	(0.014)	(0.966)
Provincial time trend	YES	YES	YES	YES	YES	YES

Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	31360	31360	31360	29695	31213	31212
R-squared	0.07	0.02	0.10	0.04	0.08	0.14

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and June 1959. They are either household head, or household head's spouse. Households with only household head or household head's spouse are not excluded.

Table 3.9 Effects of famine allowing for positive selection (Robustness)

Women	**Dependent Variables:**	Education	Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	-0.114	-0.015	0.006	-0.027	-0.014	-0.512
	(0.094)	(0.009)	(0.009)	(0.038)	(0.013)	(0.610)
Famine intensity in the first year	-0.110	-0.002	-0.008	-0.099	0.023	0.401
	(0.154)	(0.010)	(0.009)	(0.059)	(0.024)	(0.857)
Famine intensity in the second year	0.108	0.002	0.007	0.076	-0.028	-1.372
	(0.223)	(0.016)	(0.011)	(0.075)	(0.033)	(1.066)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	35282	35282	35368	30083	35341	35340
R-squared	0.14	0.03	0.10	0.06	0.10	0.16

Men	**Dependent Variables:**	Education	Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	-0.075	-0.015	0.002	-0.070	-0.000	-0.626
	(0.097)	(0.013)	(0.004)	(0.054)	(0.018)	(0.558)
Famine intensity in the first year	0.027	0.014	-0.007	0.046	0.009	0.390
	(0.114)	(0.016)	(0.007)	(0.053)	(0.019)	(0.592)
Famine intensity in the second year	-0.124	-0.006	0.002	-0.031	-0.041	-0.753
	(0.135)	(0.018)	(0.010)	(0.089)	(0.038)	(0.955)
Provincial time trend	YES	YES	YES	YES	YES	YES

Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	34075	34075	34168	32289	33975	33973
R-squared	0.09	0.04	0.10	0.04	0.09	0.15

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and January 1957, and cohorts born between July 1959 and October 1962. They are either household head, or household head's spouse. Households with only household head or household head's spouse are not excluded

Table 3.10 Effects of famine allowing for negative selection (Robustness)

	Dependent Variables:					
Women	Education		Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	-0.464	-0.033	0.180	0.237	0.173	-3.004
	(0.461)	(0.038)	(0.182)	(0.201)	(0.216)	(7.452)
Famine intensity in the first year	-0.051	-0.044	0.002	-0.145	-0.019	-0.687
	(0.120)	(0.011)***	(0.021)	(0.094)	(0.018)	(1.018)
Famine intensity in the second year	-0.043	0.007	0.000	0.091	-0.016	-0.161
	(0.146)	(0.014)	(0.013)	(0.060)	(0.021)	(0.849)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	64958	64958	64895	55538	64807	64805
R-squared	0.15	0.02	0.10	0.06	0.10	0.16
	Dependent Variables:					
Men	Education		Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	0.252	-0.071	0.014	0.026	-0.170	-0.095
	(0.487)	(0.115)	(0.053)	(0.154)	(0.159)	(3.845)
Famine intensity in the first year	-0.053	-0.014	0.003	0.022	0.034	-0.983
	(0.258)	(0.019)	(0.008)	(0.069)	(0.033)	(0.883)
Famine intensity in the second year	-0.028	0.011	0.002	-0.023	0.000	0.528
	(0.125)	(0.020)	(0.006)	(0.036)	(0.015)	(0.890)
Provincial time trend	YES	YES	YES	YES	YES	YES

Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	63376	63376	63325	60113	62883	62881
R-squared	0.07	0.03	0.09	0.04	0.10	0.14

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and June 1959, and cohorts born between November 1962 and December 1966. They are either household head, or household head's spouse. Households with only household head or household head's spouse are not excluded

Table 3.11 Factors influencing famine intensity in different provinces

	Avg. EDR(59-61)	Avg. EDR(59-61)	Avg. EDR(59-61)
Grain output per capita in 1958	-10.707		
	(9.865)		
Grain output per capita in 1956-1958		-12.312	
		(11.561)	
Grain output per capita in 1954-1958			-10.506
			(11.106)
GDP per capita in 1958	0.012		
	(1.176)		
GDP per capita in 1956-1958		-0.428	
		(1.399)	
GDP per capita in 1954-1958			-0.555
			(1.544)
CCP member/population in 1957	-2.265	-2.386	-2.348
	(1.153)*	(1.151)*	(1.165)*
Member in Dining Hall/population in 1957	0.004	-0.003	-0.004
	(0.028)	(0.030)	(0.030)
Ratio of cultivated land affected by disaster	2.990	2.859	3.055
	(3.587)	(3.573)	(3.553)
P-value of F test :			
(1) Grain output per capita=GDP per capita=0	0.53	0.49	0.53
(2) CCP member/population= Member in dining hall/population=0	0.16	0.14	0.16
Observations	25	25	25
R-squared	0.34	0.34	0.33

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.12 Urban sample

Women	Dependent Variables: Education		Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	-2.143	-0.437	-0.002	-0.131	-0.072	0.282
	(2.026)	(0.352)	(0.135)	(0.531)	(0.094)	(3.624)
Famine intensity in the first year	1.348	0.189	-0.031	0.156	0.052	0.358
	(0.381)***	(0.061)***	(0.038)	(0.119)	(0.034)	(1.234)
Famine intensity in the second year	-0.031	0.001	0.059	-0.032	0.013	-0.077
	(0.253)	(0.024)	(0.052)	(0.073)	(0.029)	(0.990)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES
Observations	11687	11687	11700	7359	11667	11667
R-squared	0.02	0.05	0.03	0.03	0.05	0.05

Men	Dependent Variables: Education		Labor Market Outcomes		Housing	
	Schooling year	High School	Employment	Work day	# of Rooms/family member	Area/family member
Famine intensity in utero	-0.278	0.052	-0.029	-0.071	0.003	-2.730
	(1.729)	(0.220)	(0.102)	(0.234)	(0.096)	(3.747)
Famine intensity in the first year	0.185	0.002	-0.032	0.073	-0.043	2.769
	(0.388)	(0.071)	(0.022)	(0.084)	(0.078)	(2.179)
Famine intensity in the second year	0.156	0.084	0.037	-0.085	-0.019	-1.065
	(0.279)	(0.032)**	(0.031)	(0.058)	(0.046)	(1.152)
Provincial time trend	YES	YES	YES	YES	YES	YES
Birth province fixed effect	YES	YES	YES	YES	YES	YES
Birth year fixed effect	YES	YES	YES	YES	YES	YES

Observations	12600	12600	12618	10671	12433	12431
R-squared	0.01	0.03	0.01	0.03	0.04	0.05

Robust standard errors in parentheses(cluster in birth province)

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: both women and men samples include cohorts born between January 1954 and June 1959. They had non-agricultural hukou in 2000.

Table 3.A Effects of famine on proportion of women in sample

	Female indicator
Famine intensity in utero	0.150
	(0.113)
Famine intensity in the first year	-0.024
	(0.025)
Famine intensity in the second year	-0.012
	(0.020)
Provincial time trend	YES
Birth province fixed effect	YES
Birth year fixed effect	YES
Observations	65984
R-squared	0.00

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Data Appendix

[1] Average number of rooms per capita: Total number of rooms divided by total number of family members.

[2] Average housing areas per capita: Total household construction areas divided by total number of family members.

[3] Schooling years: The number of years of schooling is created from respondent's self report of the highest level of education. In China, primary school is typically six years, followed by three years of junior high school, followed by three years of senior high school (or three years of professional school), followed by four years of university (*BenKe*) (or three years of college (*DaZhuan*)), followed by three years of graduate school study for MA degree, then followed by three years of graduate school study for Ph.D. degree. The respondents getting degrees from adult education programs are dropped.

[4] High school indicator: Whether the highest level of education of the respondent is high school or above. The respondents getting degrees from adult education programs are dropped.

[5] Employment indicator: Whether the respondent worked at least one hour during Oct. 25 and Oct. 31 in 2000.

[6] Work day: How many days did the employed respondent work during Oct. 25 and Oct. 31 in 2000?

References

[1] Alderman,Harold,Gertler,Paul (1997). *Family resources and gender differences in human capital investments: the demand for children's medical care in Pakistan*. In: Haddad,Lawrence,Hoddinott,John, Aldermand, Harold (Eds), Intrahousehold resource allocation in developing countres: Models, methods, and policy. Baltimore and London: John Hopkins University Press; 1997. p. 231-248 .

[2] Alderman,Harold,Hodinott,John,Kinsey,Bill,(2006). *Long term consequences of early childhood malnutrition*. Oxford Economic Papers 58,450-474.

[3] Almond,Douglas,Chay,Kenneth,(2003). "The long run and intergenerational impact of poor infant health: evidence from cohorts born during civil rights era." New York, NY (Columbia University) and Berkeley, CA (University of California at Berkeley), Working Paper.

[4] Almond, Douglas,(2006). *Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 U.S. population*. Journal of Political Economy 114,672-712 (August).

[5] Almond,Douglas, Edlund,Lena, Li,Hongbin, Zhang,Junsen, (2007). *Long-term effects of the 1959-1961 China famine: mainland China and Hong Kong.* New York, NY (Columbia University), New York, NY (Columbia University), Shatin, HK (Chinese University of Hong Kong) and Shatin, HK (Chinese University of Hong Kong), Working Paper.

[6] Ashton,Basil, Hill,Kenneth, Piazza,Alan, Zeitz,Robin, (1984). *Famine in China,1958-61*. Population Development Review 10(4), 613-645(December).

[7] Barker,D.J.P, (1998). Mothers, babies and health in later life. Edinburh: Churchill Livingstone; 1998.

[8] Barker,D.J.P., Bleker,O.P., Montfrans,G.A.van, Osmond,C., Ravelli,A.C.J., Roseboom,T.J., (2005). *Cardiovascular disease in survivors of the Dutch famine*. In: Hornstra,Gerard, Uauy,Ricardo, Yang,Xiaoguang(Eds), The impact of maternal nutrition on the offspring. Nestle Nutrition workshop series, Pediatric Program, vol. 55.

[9] Behrman,Jere R., Deolalikar,A.B., (1988). *Nutrients-impacts and determinants*. The World Bank Economic Review 2(3), 299-320.

[10] Behrman, Jere R., (1988). *Nutrition, health, birth order and seasonality: intrahousehold allocation among children in rural Inida*. Journal of Development Economics 28(1),43-62.

[11] Behrman,Jere R., Rosenzweig,Mark, (2004). *Returns to birthweight*. Review of Economics and Statistics 86(2), 586-601.

[12] Bertrand,Marianne, Duflo,Esther, Mullainahtanm,Sendhil, (2004). *How much should we trust difference-in-difference estimates?* Quarterly Journal of Economics 119(1), 249-275(February).

[13] Cameron, Lisa A., Worswick, Christopher, (2001). *Education expenditure responses to crop loss in Indonesia: a gender bias*. Economic Development and Cultural Change 49(2), 352-363.

[14] Chen,Yuyu, Zhou,LiAn, (2007). *The long-term health and economic consequences of the 1959-1961 famine in China*. Journal of Health Economics 26, 659-681.

[15] China National Bureau of Statistics, (2005). China compendium of statistics: 1949-2004. China National Bureau of Statistics Press.

[16] Currie,Janet, Hyson,Rosemary, (1999). *Is the impact of health shocks cushioned by socio-economic status? the case of low birthweight*. American Economic Review Papers and Proceedings 89(2), 245-250.
[17] Currie, Janet, Madrian,Brigitte C., (1999). *Health, health insurance and the labor market*. In: Ashenfelter,Orley C., Card, David(Eds), Handbook of Labor Economics, vol.3. North Holland: Amsterdam; 1999. p.3309-3407.
[18] Diamond,Jared, (1991). *Pearl harbor and the emperor's physiologists*. Natural Histroy 12, 2-7.
[19] Dreze,Jean, Sen,Amartya. Hunger and public action. Oxoford: Oxford University Press; 1989.
[20] Duflo, Esther, (2003). *Grandmothers and granddaughters: old age pension and intrahousehold allocation in South Africa*. The World Bank Economic Review 17(1), 1-25.
[21] Ford, Kathleen, Huffman,Sandra L., Chowdhury,A.K.M.A., Becker,Stan, Allen,Hubert, (1989). *Birth-interval dynamics in rural Bangladesh and maternal weight*. Demography 26(3), 425-437(August).
[22] Forster, Robert, Ranum,Orest(Eds.). Biology of man in history. Baltimore and London: The Johns Hopkins University Press; 1975.
[23] Frisch, Rose E., McArthur,J.W., (1974). *Menstrual cycles: fatness as a determinant of minimum weight for height necessary for their maintenance or onset*. Science 184, 949-951.
[24] Frisch, Rose E., (1978). *Population, food intake, and fertility*. Science, New Series 199(4324), 22-30.
[25] Frisch, Rose E.. Female fertility and the body fat connection. Chicago: University of Chicago Press; 2002.
[26] Giles, John, Park,Albert, Wang,Meiyan, (2007). *The Great Proletarian Cultural Revolution, disruptions to education, and returns to schooling in urban China*. Lansing, MI (Michigan State University), Ann Arbor, MI (University of Michigan), Beijing, China (Institute of Population Study), Working Paper.
[27] GØrgens,Tue, Meng,Xin, Vaithianathan,Rhema, (2007). *Stunting and selection effects of famine: a case study of the great Chinese famine*. Discussion Paper 2543, IZA.
[28] Gregory, R.G., Meng, Xin, (2002a). *The impact of interrupted education on subsequent educational attainment: a cost of the Chinese cultural revolution*. Economic Development and Cultural Change 50(4), 935-959.
[29] Gregory, R.G., Meng, Xin, (2002b). *Exploring the impact of interrupted education on earnings: the educational cost of the Chinese cultural revolution*. Canberra, Australia (Australia National University) and Canberra, Australia (Australia National University), Working Paper.
[30] Grossman, Michael, (1972). *On the concept of health capital and demand for health*. Journal of Political Economy 80(2), 223-255.
[31] Jayachandran, Seema, (2005). *Air quality and early-life mortality: evidence from Indonesia's wildfires*. Los Angeles,CA (University of California at Los Angeles), Working Paper.
[32] Jowett, A. John, (1991). *The demographic response to famine: the case of China 1958-1961*. GeoJournal 23(2), 135-146(February).

[33] Langsten, R., (1985). *Determinants of natural fertility in rural Bangaldesh reconsidered*. Population Studies 39(1), 153-161(March).
[34] Lin, Justin Y., Yang, Dennis T., (2000). *Food availability, entitlements and the Chinese famine of 1959-61*. Economic Journal 110, 136-158(January).
[35] Lindeboom, Maarten, Van den Berg,Gerard J., Portrait, France. (2006). *Economic conditions early in life and individual mortality*. American Economic Review 96(1), 290-302(March).
[36] Luo, Zhehui, Mu,Ren, Zhang,Xiaobo, (2006). *Famine and overweight in China*. Review of Agricultural Economics 28(3), 296-304.
[37] Maccini, Sharon, Yang,Dean, (2006). *Under the weather: health, schooling, and socioeconomic consequences of early-life rainfall*. Ann Arbor, MI (University of Michigan) and Ann Arbor, MI (University of Michigan), Working Paper.
[38] Meng, Xin, Qian,Nancy, (2006). *The long run impact of childhood malnutrition: evidence from China's great famine*. Canberra, Australia (Australia National University) and Providence, RI (Brown University), Working Paper.
[39] Moulton, Brent, (1986). *Random group effects and the precision of regression estimates*. Journal of Econometrics 32(3), 385-397(August).
[40] Peng, Xizhe, (1987). *Demographic consequences of the Great Leap Forward in China's provinces*. Population and Development Review, 13(4), 639-70.
[41] Ravallion, Martin. Markets and famines. Oxford: Oxford University Press; 1987.
[42] Rose, Eliana, (1999). *Consumption smoothing and excess female mortality in rural India*. Review of Economics and Statistics 81(1), 41-49(February).
[43] Royer, Heather, (2005). *Separated at birth: estimating the long-run and intergenerational effects of birthweight using twins*. Ann Arbor, MI (University of Michigan), Working Paper.
[44] Sen, Amartya. Poverty and famine: an essay on entitlement and deprivation. Oxford: Oxford University Press; 1981.
[45] SG,Elias, PA, van Noord, PH,Peters, Tonkelaar I,den, R,Kaaks, DE, Grobbee, (2007). Menstruation during and after caloric restriction: the 1944-1945 Dutch famine. Fertil Steril, June.
[46] Stein, Zena, Susser,Mervyn, Saenger,Gerhart, Marolla,Francis. Famine and human development: the Dutch hunger winter of 1944-1945. New York, London and Toronto: Oxford University Press; 1975.
[47] US Center for Disease Control and Prevention, (2000). Children growth chart.
[48] Wooldridge, Jeffrey M.. Econometric Analysis of Cross Section and Panel Data. Massachusetts Institute of Technology Press; 2002.
[49] Yang, Dali L.. Calamity and reform in China. Stanford University Press; 1996.
[50] http://en.wikipedia.org/wiki/Cultural_Revolution

Chapter 4

Does school quality matter?: Evidence from a natural experiment in rural China[63]

4.1 Introduction

One of the most fundamental questions in the economics of education is the extent to which school quality affects educational outcomes. This is an especially critical question in developing countries, where budgetary resources are scarce and governments must trade off improvements in educational quality with expanding access to education. Most empirical studies of the effects of various measures of school quality (e.g., educational spending and teacher quality) are difficult to interpret because students, teachers, and resources are almost never randomly allocated between schools.[64] Therefore, there are almost always unobserved factors correlated with both school quality and student attributes.

A few studies have attempted to address the endogeneity problem using different approaches, such as comparison with matched control groups (Angrist and Lavy, 2001; Rivkin, Hanushek and Kain, 2005; Rockoff, 2004), conducting randomized trials to examine the impact of specific schooling inputs (Glewwe, Kremer and Moulin, 2007; Banerjee, Cole, Duflo, and Linden, 2006; Dee, 2001), and exploiting natural experiments that create plausibly exogenous variation in class size (Angrist and Lavy, 1999; Hoxby, 2000) or in the quality of schools attended, e.g., lotteries (Gould, Lavy and Paserman,

---

63 This chapter is based on a paper co-authored with Albert Park, Chang-tai Hsieh and Xuehui An.

64 Although some research has found no significant link between educational attainment, usually measured by test scores, and school quality, measured by teachers' education level, school resources, equipment, class size, etc. (Altonji, Elder and Taber, 2005; Hoxby, 2000), other research has found a significantly positive relationship (Angrist and Lavy, 1999, 2001; Case and Deaton, 1999; Dee, 2001; Eide and Showalter, 1998; Goldhaber and Brewer, 1997; Gould, Lavy and Paserman, 2004; Greenwald, Hedges and Laine, 1996; Krueger, 1999; Krueger and Whitmore, 1999; Lai, Sadoulet and Janvry, 2006; Lavy, 2002; Newhouse and Beegle, 2005; Rivkin, Hanushek and Kain, 2005; Rockoff, 2004; Urquiola, 2006). School quality also has been found to be correlated with students' subsequent economic outcomes (Altonji and Dunn, 1996; Betts, 1995, 2001; Card and Krueger, 1992a).

2004; Hoxby 2007). However, these studies have weaknesses. For example, studies using matched control groups do not convincingly control for unobservable student differences, and many of the studies focus narrowly on one dimension of schooling quality even though many other dimensions may be more important.

In this paper, we present new evidence from a natural experiment in China. We quantify the impact of school quality on students' educational attainment using a regression discontinuity research design that exploits the fact that in China's rural counties admission to the best high schools are strictly determined by entrance examination scores. In addition to informing general debates over the importance of school quality, because many countries in Asia (e.g., Taiwan, Hong Kong, Singapore, and Korea) employ a similar mechanism for allocating students to schools, our results may help such countries evaluate the benefits of maintaining such a merit-based allocation system at the high school level rather than shifting to other allocation systems, such as neighborhood-based schools.

China has a magnet school system for high schools in rural counties. Typically, a rural county has one magnet school located in the county seat and a number of regular high schools spread across the county, and students resident in a given county attend one of the county's public high schools. Graduating middle school students must take county-wide high school entrance examinations, which determine whether they attend a magnet high school, a regular high school, or no high school at all. In any given county, the magnet school has the best quality as well as the highest entrance examination score cutoff line. This paper uses a regression discontinuity design to compare students with nearly identical entrance exams who attend different quality schools because they are just above or just below the cutoff score for admission to the magnet school. In this paper, we first investigate whether entering a good school has any effects on the probability for students to take college entrance examinations. Then for those students taking college entrance examinations, we examine the effects of entering a good school on their educational attainment, measured by standardized college entrance examination scores and the probability of entering college. Using data from 7 counties in Gansu province in China, we find that entering the magnet high school can decrease the probability to take college entrance exam by 14.9 percent, increase students' standardized college entrance score by 0.256 standard deviation, and increase students' probability of entering college by 22.5 percentage points.

The rest of the paper is organized as follows. Section 4.2 describes the data and variable definitions. Section 4.3 presents the methodology for implementing the regression discontinuity design and describes the empirical specification. Section 4.4 describes students' assignment to different schools in the sample used for analysis. Section 4.5 presents the empirical results, and section 4.6 describes the differences between magnet schools and regular schools that might explain differences in educational attainment. Section 4.7 concludes.

4.2 Data and variables

4.2.1 Data collection

The data used in this paper was collected from high schools in rural counties in Gansu Province in western China during the summer of 2004. Gansu is one of China's poorest provinces, with a GDP per capita of $744 in 2004 which ranked 30th among China's 31 provinces. It had a population of 26 million in 2004, ranking 22nd among Chinese provinces.

The sample used for analysis includes 12035 students in entering classes from 1997 to 2001 of 34 high schools located in 7 rural counties. China's high schools have three grade levels, so all students in the sample had completed high school and taken college entrance examinations by the time of the survey. Initially, the data includes high schools in 19 counties, distributed in 74 county-years[65] (Table 4.1 shows statistics of these 74 county-years). However, a few counties had data only for one school. In addition, county education bureaus have discretion in determining how students are allocated to schools. For example, some schools allow students with poorer scores to pay additional fees in order to attend the best schools, a practice which has become more prevalent over time. It is possible that some counties also allow personal connections to principals or school officials to influence admissions decisions. For these reasons, we exclude 21

[65] 100 county-years, including 20 counties in 5 years, were intended to be surveyed, but 26 of them were lack of either school information or student information. Therefore, we collected information for both students and schools in 74 county-years.

county-years for which data on only one high school is available (the first panel in Table 4.1 shows these county-years). We also drop counties that did not strictly adhere to examination score cutoff lines in determining who was admitted to the best high schools. We restrict the sample by excluding 11 county-years for which the percentage of students entering the magnet high school, where the magnet high school is defined as having the highest average students' entrance score and the highest high school entrance score cutoff line among all schools in the same county in the same year, having lower entrance score than this school's cutoff line is higher than 30 percent. We also exclude 4 county-years where the percentage of regular students having high school entrance score higher than the highest cutoff line is extremely high, actually higher than 70%. Panel 2 in Table 4.1 shows these county-years. According to the above definition of magnet school, we drop 4 county-years for which the school with the highest mean high school entrance score is different from the school with the highest entrance score cutoff line**,** and 12 county-years where the cutoff lines of high school having the highest mean high school entrance score are missing. Panel 3 and panel 4 in Table 4.1 show these two groups of county-years.

Nearly 100 percent of students report a high school entrance exam score, but only 59.01 percent of those with high school entrance exam scores also report a college entrance examination score. Attrition could be due to the following reasons: the student dropped out, transferred to another school, or decided not to sit for the college entrance examination; or the school kept incomplete records. In this paper, besides students' educational attainment, we also investigate the effects of entering magnet school on the students decisions to take college entrance exam, so we drop those county-years for which the schools might not keep the college entrance score of those students not admitted by colleges. Table 4.1 shows the percentage of students having missing college entrance and the percentages of students admitted or not admitted by colleges. Those 6 county-years in panel 5 in Table 4.1 are most likely to keep college entrance scores only for students admitted by colleges either because the percentage of students having missing college entrance score is the same as (or very close to) the percentage of students not entering the colleges or because the percentage of students having college entrance score is very close to the percentage of students entering colleges. We also exclude these 6 county-years. Panel 6 in Table 4.1 shows 2 county-years having more than 45%

students with missing information of their gender; after these students are dropped; only information for one school is remained, so we drop these 2 county-years. We are left with a sample of 12035 students for 14 county-years which are shown in panel 7 in Table 4.1.

This survey collected information on students' gender, birth year, year of high school entrance, high school entrance exam score, and college entrance examination score. The survey also collected data from schools on the high school entrance exam score cutoff line and school characteristics such as teachers' education levels and school facilities.

4.2.2 Variable definitions

Two treatment variables are defined. The variable *treatment* is assigned to equal one if the student attends the best school as defined above. The other variable *eligible* is assigned to equal one if the student's high school entrance score is higher than the high school entrance exam cutoff line of the best school in the county-year. While *treatment* more accurately reflects whether students attended better schools, it is more subject to selection bias. The coefficient on *eligible* should be viewed as capturing a lower bound estimate of the true treatment effect. Our preferred estimate is one in which *eligible* is used as an instrument for *treatment*.

The high school entrance examination scores and college entrance examination scores are the most important variables. We use the following formulas to standardize test scores.

$$\hat{HS}_{ijt} = \frac{HS_{ijt} - \overline{HS}_{jt}}{HSSD_{jt}} \tag{4.1}$$

$$\hat{CS}_{ijpt} = \frac{CS_{ijpt} - \overline{CS}_{jpt}}{CSSD_{jpt}} \tag{4.2}$$

Here, HS_{ijt} is the high school entrance examination score for student *i* in county j who entered high school in year *t*, $\overline{HS}_{jt}$ is the mean high school and college entrance exam score for students in county *j* who entered high school in year *t*, and $HSSD_{jt}$ is the standard deviation of high school entrance exam scores among students in county *j* who began high school in year *t*. The variables in (4.2) are similarly defined, but for college entrance examination scores. The only difference is that there is an additional subscript *p* which is for the college exam type, which in China can be liberal arts, science, physical education, musical education, or arts education. Thus, CS_{ijpt} is the college entrance examination score for student *i* entering high school in year *t* in county *j* taking test *p*.

Another educational attainment variable is an indicator variable for whether the student is qualified to attend college. In the survey, there is a question about whether the student actually enrolled in college but many answers to this question were missed. We therefore compares student's college entrance exam score with the lowest college admission cutoff lines in Gansu Province in the year the student took the college entrance exam in order to determine whether the student qualified for college.[3] This measure is similar to the eligibility variable in that it is not subject to selection biases associated with the student's decision to actually attend college conditional on his or her entrance exam score being above the cutoff line. Such decisions will be influenced by credit constraints, family income and wealth, and parental expectations.

In this paper, student age at the time of entering high school is first calculated by subtracting students' birth year from the year they entered high school. Unfortunately, there are many missing values for birth year. We therefore impute student age by setting missing values to be equal to the average student age in the same school in the same year. In cases where age is missing for all students in the whole school, it is imputed to be equal to the mean student age of all students entering school in the same county in the same year. And if data is missing for all students in the county-year, then age is imputed to equal the average student age for all students in the sample entering high school in the same year.

Table 4.2 gives summary statistics for the variables used in the analysis. The upper panel is for all the students including those having and not having college entrance

---

[3] The cutoff lines from different provinces come from http://www.eol.cn/include/cer.net/gaokao/zhuanti/2006_fenshuxian.shtml#2000. There are different lowest cutoff lines for different types of college entrance exams.

scores. Thirty eight percent of students are female, and the average age is 15.78. Also, 44 percent of students in the sample attend the best school in the county, while 50 percent are eligible to attend the best school in the county. The bottom panel shows the statistics for students whose college entrance scores are not missing. The mean value of their normalized college entrance scores is very close to zero. Forty nine percent of students qualify for college entrance. 35 percent are female, and the average age is 15.80. And 49 percent of students in this sample attend the best school in the county, while 56 percent are eligible to attend the best school in the county. Comparing these two groups of students, we can see that the ratio of female students and the average age are similar, but the percentages of students entering the best high school or eligible to attend the best high school are 6 percent points higher in the smaller sample.

4.3 Methodology

4.3.1 Regression discontinuity design

We employ regression discontinuity design (RDD) to quantify the impact of school quality on educational attainment. First developed by Thistlethwaite and Campbell (1960), in recent years there has been renewed interest in identification issues that arise using RDD (Hahn, Todd and van der Klaauw, 2001; Lee, 2004), estimation issues (Porter, 2003), and RDD applications (Angrist and Lavy, 1999; van der Klaauw, 2002).

The basic regression model we employ is the following:

$$Y_i = \beta + \alpha * T_i + u_i$$

$$T_i = 1\{S_i \geq \bar{S}\}$$

In the first equation, Y_i is outcome variable, and T_i is the treatment variable, which equals one if treated and zero otherwise. The second equation describes the selection rule: those who fall below some distinct cutoff point $\bar{S}$ are placed in the control group (T_i =0), whereas those on or above that point are placed in the treatment group (T_i =1).

In regression discontinuity design, under the assumption that the conditional mean function E[u|S] is continuous at $\bar{S}$, the treatment effect α can be identified as follows:

$$\alpha = \lim_{S \downarrow \bar{S}} E[Y \mid S] - \lim_{S \uparrow \bar{S}} E[Y \mid S]$$

Intuitively, the treatment effects are identified by the sample of individuals within a very small interval around the cutoff point. Since these individuals have essentially the same S_i value, we can expect individuals just below the cutoff line on average to be very similar to individuals just above the cutoff line and thus to have similar average outcomes in the absence of the treatment as well as similar average outcomes when receiving treatment. With those to the right of the cutoff line receiving treatment and those to the left not, a comparison of the average outcomes of both groups should therefore provide a credible estimate of the treatment effect.

In the model, S_i could be related to the outcome Y_i which would cause T_i to be related to Y_i as well. This could cause OLS estimate of α to be inconsistent. One approach proposed by Heckman and Robb (1985) to estimate the treatment effect in this case is to specify and include the conditional mean function E[u| T, S] as a "control function" in the outcome equation. Therefore, in practice, the following equation is estimated:

$$Y_i = \beta + \alpha * T_i + k(S_i) + \omega_i \qquad (4.3)$$

As long as $k(S_i)$ is continuous in S, identification can be guaranteed because of the discontinuity in the function T(S). In this paper, $k(S_i)$ is approximated by a 5 order polynomial function of the standardized high school entrance exam score, that is:

$$k(S_i) \approx \sum_{j=1}^{5} \eta_j * S_i^j$$

Lee (2003) proposes a direct test of the continuity assumption by checking whether there are discontinuities in the relationship between the treatment effect and any predetermined characteristics. That is, the following equation can be estimated as a pre-assumption test:

$$X_i = \phi + \varphi * T_i + \sum_{j=1}^{5} \mu_j * S_i^j + \varepsilon_i \qquad (4.4)$$

If φ is not significant, then the continuity assumption is valid. In this paper, student age and gender are two predetermined characteristics that can be tested.

4.3.2 Fuzzy regression discontinuity design and IV estimation

In the regression discontinuity design discussed above, treatment depends on the selection variable S in a deterministic way. However, in reality, it is more likely for treatment assignment to depend on S in a stochastic manner, which is called as fuzzy regression discontinuity design in the literature. In this case, OLS estimates of equation (4.3) are biased.

In our context, even after we apply an exclusion criteria to ensure that violations of the eligibility rule are less than 30% in the best school and exclude county-years where the percentages of regular students having high school entrance score higher than the cutoff line are high, there are still students in the best schools whose high school entrance exam scores are below the cutoff line and students not in the best schools whose scores are above the cutoff line. In this case, the OLS estimate of α in equation (4.3) using the variable *treatment* is likely to be downward biased. In our used sample, i.e. 12035 students, 19.06% students not in the best schools have high school entrance examination score above the cutoff line, but only 10.31% students in the best schools have high school entrance examination score below the cutoff line.

This is where the second treatment variable *eligible* can avoid the problems associated with bias caused by fuzzy discontinuity design. The variable *eligible* itself does not suffer from fuzziness and so can be used to cleanly estimate a treatment effect. However, the impact of eligibility is not of primary policy interest, one wants to estimate the impact of actually attending better schools. To obtain an unbiased estimate of actually attending better schools, we can use *eligible* as an instrument for *treatment*.

4.4 Student assignments

Students are assigned to magnet schools and regular schools according to their high school entrance examination score. Figure 4.1 shows the distribution of students with different high school entrance scores in magnet school and regular schools in 7 counties used for analysis. In each panel, on the y-axis, a “one” indicates a student that is treated, i.e. enters a magnet school; while a “zero” indicates untreated students. In order to pool data from different years for each county, we create a variable that indicates each student’s score relative to its respective regional cutoff in each year, which is shown in x-axis. Each dot in graphs on the left hand side in each panel represents each student. Each graph on the right hand side is nonparametric predication from a local polynomial smoother with degree equal to zero and bandwidth equal to 0.25. These figures highlight two things; one is that there are sharp changes in the probability to be treated close to the cutoff. We also conduct regressions of *treatment* on *eligible* after controlling for female dummy, age, county fixed effects, year fixed effects, and middle school fixed effects. Column (1) and column (5) in Table 4.3 show the estimated results for the whole sample and the sub-sample of students having college entrance scores respectively. A student eligible for magnet school is 58.8% (in the whole sample) or 57.9% (in the sub-sample) more likely to be treated, i.e. enroll in the magnet school.

However, the second thing we can see from these graphs is that, in practice, the strictness of implementation of cutoff line is different in different counties. From the figures, we know that the cutoff line should be implemented the most strictly in county 16, 20, 26, and 29. Columns 9 and 10 in Table 4.1 show the percentage of magnet students having entrance score lower than cutoff line and the percentage of regular students having entrance score higher than cutoff line in each county-year. We can see large variations in these two percentages. In Panel 7 in Table 4.1 where the county-years are used for analysis in this paper, we can see that there are 28.87% magnet students having high school entrance score lower than cutoff line in county 28 in year 2000; but in county 16 in year 1997, county 16 in 1999, and county 29 in 1998, there are no magnet students having high school entrance score lower than cutoff line. We can also see that 39.47% regular students in county 19 in year 2000 have higher high school entrance examination score than cutoff line; this percentage is the highest among all 14 county-year in Panel 7 in Table 4.1; however, in county 16 in 2001, only 2.64% regular students have high school entrance score higher than the cutoff line. Because of this fuzziness in the implementation of cutoff lines, in the following analysis, our preferred results come from the regressions using *eligible* as an instrument for *treatment*.

4.5 Results

4.5.1 Pre-assumption tests

We first report results for the pre-assumption tests that test whether sex or age jumps in a discontinuous fashion at the high school entrance exam cutoff line. In Figure 4.2, the x-axis measures the difference between the standardized high school entrance exam scores and the standardized cutoff line in the same county in the same year; the y-axis in Panel A measures the proportion of students that are female, while the y-axis in Panel B measures the students' age. The graph on the left hand side in each panel is from the whole sample, but the graph on the right hand side in each panel is from the sub-sample including students with college entrance scores. The curve in each graph is nonparametric prediction from local polynomial smoother with degree 0 and bandwidth 0.25. We can see that for female proportion and student age, whether we use the whole sample or the sub-sample, there is no jump at x=0, where the high school entrance examination score is equal to the cutoff line. These figures provide support for the validity of the regression discontinuity design.

Table 4.4 presents the regression results for the pre-assumption tests. Columns (1)-(4) are for the whole sample, and columns (5)-(8) are for the sub-sample of students with college entrance scores. In the regression specifications, we control for a 5^{th}-order polynomial function of the normalized high school entrance score, county fixed effects, year fixed effects and middle school fixed effects. Columns (1), (3), (5) and (7) present results for OLS regressions of female and age on *eligible*; while columns (2), (4), (6) and (8) present results for regressions of female and age on *treatment* using *eligible* as IV. From the table, we can see that, the coefficients before female and age are not statistically significant, whether they are estimated using OLS or IV regression. These results again show that the students' pre-characteristics are continuous at the cutoff line.

4.5.2 Treatment effects

Figure 4.3 plots three outcome variables, i.e. index for having college entrance scores, normalized college entrance scores, and index for college eligibility as a function of the normalized high school entrance exam score relative to cutoff lines. Each graph is nonparametric prediction from local polynomial smoother with degree 0 and bandwidth 0.5.

Panel A in Figure 4.3 presents the estimates of the index for having college entrance scores. When the students are treated and untreated, i.e. the curve on the left and the curve on the right, the proportion of students having college entrance score is continuous, but there is a gap between these two curves at x=0 by a magnitude of roughly 0.05; and the curve representing untreated students is above the curve representing treated students. It means entering magnet school might reduce the probability to take college entrance examination. Panel B in Figure 4.3 presents a similar graph for normalized college entrance score. The curve of treated students is above the curve of untreated students at x=0; the gap is roughly equal to 0.3, which implies a 0.3 standard deviation increase in college entrance examination score. Panel C also presents a similar graph for the index of college eligibility. From the graph, we can also see that the curve of treated students is above the curve of untreated students at x=0; the gap is roughly equal to 0.2, which implies a 20% increase in qualifying college. These pictures provide rough visual estimates of the size of the treatment effect using RDD.

4.5.2.0 First stage results

Table 4.3 also presents the first stage regression results. Columns (2), (3) and (4) are the results using the whole sample; columns (6), (7) and (8) are the results using the sub-sample of students having college entrance examination scores. In all the specifications, county fixed effects, year fixed effects and 5 order polynomial function of normalized high school entrance score are controlled for. In columns (2) and (6), we do not control for other variables; in columns (3) and (7), we control for sex and age; in column (4) and (8), we control for sex, age and middle school fixed effects. The last row of Table 4.3 presents F-value of the null hypothesis that the coefficient of *eligible* is equal to zero.

For the whole sample, all three coefficients of *eligible* are statistically significant at 1% level, with the magnitude of 0.468, 0.469 and 0.381. The F-value is equal to 1344.91, 1346.00 and 962.89. It shows that for the whole sample, *eligible* can be used as a valid IV for *treatment*. For the sub-sample, from Table 4.3, we can also see that all three coefficients of *eligible* are statistically significant at 1% level, with the magnitude of 0.492, 0.493 and 0.409. The corresponding F-values are 887.61, 889.56 and 643.21. These results also show that for the sub-sample, *eligible* is also a valid IV for *treatment*.

4.5.2.1 Effects on students' decision to take college entrance examination

Upper panel in Table 4.5 presents the estimates of effects of entering magnet school on students' decision to take college entrance examination. Three types of specifications are estimated. The first one uses *treatment* as the independent variable of interest. The second one uses *eligible* as the key independent variable. The third uses *eligible* as an instrument for *treatment*, which is our preferred specification for the reasons described earlier. In the upper panel, we control for county fixed effects and year fixed effects in all specifications. In columns 1 to 3, we control only for the 5-th order polynomial function of standardized high school entrance exam scores, and for the time-being do not control for gender, student age, or middle school fixed effects. All three coefficients for the treatment variables are negative and statistically significantly different from zero, with magnitudes of -0.046, -0.067, and -0.143. As expected, the IV estimate is larger than both OLS estimates. The estimates using *treatment* as the variable of interest produce the smallest treatment effect magnitudes. Columns 4 to 6 present estimates from specifications that control for sex and student age. There are just trivial changes in the coefficients, changing to -0.046, -0.066 and -0.141. Columns 7 to 9 present results adding controls for which middle school the student attended. In this case, the impact of *treatment* changes to -0.056 (column 7), the impact of *eligible* falls to -0.057 (column 8); but the IV estimate of the treatment effect does not change much, becoming -0.149 (column 9). One interpretation for the negative treatment effects is because the RDD is comparing students just above the cutoff line and students just below the cutoff line. For the students with score just above the cutoff line and entering the magnet school, they could be the worst students in the best school; but for those students with score just below the cutoff line and entering the regular school, they could be the best students in the regular school. Therefore, it is very possible for the worst students in the best school to give up taking college entrance examination when they think their chance to be admitted is low. But those best students in the regular schools would decide to take college entrance examination because they are the best in their schools. In the bottom panel, we conduct a regression of index for taking college entrance examination on treatment, normalized high school entrance score, the interaction of treatment and normalized high school entrance score. As the upper panel, we use three different specifications, i.e. not controlling for sex, age, or middle school fixed effects, controlling for sex, age, and controlling for sex, age and middle school fixed effects. We can see the IV estimates, which are preferred. All the coefficients before *treatment* are statistically negative, all the coefficients before the interaction of *treatment* and normalized high school entrance score are also statistically negative, but the coefficients of normalized high school entrance score are statistically positive. It shows that given the save high school entrance score, the probability to take college entrance examination is higher for students entering regular schools, which explain why the estimates from RDD are negative, which identify the probability difference between treated students and untreated students at the cutoff line.

4.5.2.2 Effects on students' college entrance score and college eligibility

Table 4.6 presents the estimates of effects of entering magnet school on students' college entrance examination score and students' probability of qualifying to enter college. Three types of specifications are estimated. The first uses *treatment* as the independent variable of interest. The second one uses *eligible* as the key independent variable. The third uses *eligible* as an instrument for *treatment*, which is our preferred specification for the reasons described earlier. In all the specifications, we control for county fixed effects and year fixed effects. One caveat we need to bear in mind is that, since only students taking college exam are used for analysis, therefore, all the results in Table 4.6 should be interpreted as effects of school quality conditional on students' decision to take college entrance exam. In columns 1 to 3, we control only for the 5-th order polynomial function of standardized high school entrance exam scores, and for the time-being do not control for gender, student age, or middle school fixed effects. All three coefficients for the treatment variables are positive and statistically significantly different from zero, with magnitudes of 0.110, 0.124, and 0.252. As expected, the IV estimate is larger than both OLS estimates. The estimates using *treatment* as the variable of interest produce the smallest treatment effect magnitudes. Columns 4 to 6 present estimates from specifications that control for sex and student age. There are just trivial changes in the coefficients, changing to 0.108, 0.126 and 0.255. Columns 7 to 9 present results adding controls for which middle school the student attended. In this case, the impact of *treatment* decreases to 0.089 (column 7), the impact of *eligible* falls to 0.105 (column 8); but the IV estimate of the treatment effect is 0.256(column 9), slightly higher than the previous two specifications. The results show that entering magnet school can increase the students' college entrance examination score by 0.256 standard deviation.

The bottom panel of Table 4.6 shows similar results on the impact of school quality on students' probability of qualifying to enter college. In these regressions, county fixed effects and year fixed effects are controlled for. In the simplest specification (columns 1 to 3), the effect of *treatment* is 0.058, the effect of *eligible* is 0.104, and the IV estimate is 0.212. Again, as expected the largest estimate is the IV estimate. Controlling for gender dummy or student ages slightly changes estimation results (columns 4 to 6); the effect of *treatment* decreases to 0.057, the effect of *eligible* increases to 0.105, and the IV estimate increases to 0.214. After adding middle school fixed effects, just as for college entrance exam scores, the effect of *treatment* decreases (to 0.045), the effect of *eligible* decreases (to 0.092), and the IV treatment effect estimate increases (to 0.225). Thus, attending a better high school increases the probability of qualifying for college entrance by 22.5 percent.

In practice, the unconditional effects of school quality on the students' probability to be admitted by colleges might be more of interest since students not taking college entrance exam would usually thought of as not being admitted by colleges and therefore the unconditional school's effects are more relevant with policy analysis. In Table 4.7, we define all students not taking college entrance exam as not being admitted by colleges, and then re-estimate the impacts of school quality on students' probability of qualifying to enter college using the same specifications as in Table 4.6. Columns 1 to 3 show the results from the simplest specification in which we only control for county fixed effects, year fixed effects and 5^{th} order polynomial terms. The effect of *treatment* is 0.015 (column 1), the effect of *eligible* is 0.057 (column 2), and the IV estimate is 0.114. Controlling for gender dummy or student ages almost does not change estimation results (columns 4 to 6); the effect of *treatment* decreases to 0.014, the effect of *eligible* increases to 0.058, and the IV estimate increases to 0.116. After adding middle school fixed effects, we get the more reliable estimates. The effect of *treatment* decreases -0.001, the effect of *eligible* decreases to 0.048, and the IV estimate increases to 0.120. As discussed in the above, entering a better high school can (unconditionally) increase the probability of qualifying for college entrance by 12 percent.

4.5.2.3 Effects of school quality within narrow bands of the selection threshold

Table 4.8 shows the results using students with high school entrance examination scores within different neighborhoods around the cutoff lines in their counties in the year they entered high school. There are three panels in Table 4.8; the upper panel shows results for the index of having college entrance examination scores, the middle panel shows results for the normalized college entrance examination scores, and the bottom panel shows the results for the probability of qualifying to enter colleges. Since we only use the sample within the neighborhood around the cutoff lines, we do not include 5-th order polynomial function of standardized high school entrance examination scores. As in Table 4.6, we control for county fixed effects and year fixed effects for all the specifications. For each specification, we use three different bands of the selection thresholds: 1.5, 1 and 0.5 standard deviations of high school entrance examination scores. Specifically, we estimate the effects of school quality using students with high school entrance examination scores within neighborhoods with different bands around the cutoff lines in each county-year. Here, we just present the preferred IV estimates. The first thing we know from Table 4.8 is that all the estimates are statistically significant from zero, and most of them are at 1% level.

In column (1)-(3), we do not control for other variables. In the upper panel, we can see that with the band of neighborhood becoming from 1.5 standard deviations to 0.5 standard deviation, the coefficient is changing from 0.099 to 0.047, and then to -0.071. In column (4)-(6), we control for sex and age. Compared with the first three columns, the magnitude of the coefficient is almost the same, and the coefficient changes from 0.096 to 0.045, then to -0.072. If sex, age and middle school fixed effects are controlled (column (7)-(9)), we can see that the coefficient becomes smaller; with the band becoming narrower, the coefficient changes from 0.085 to 0.038, then to -0.090. One possible reason for the decrease in the treatment effects might be: with the neighborhood becoming narrower and narrower, the qualities of students in the magnet schools become worse and worse but the qualities of students in the regular schools become better and better relative to other students in the same school, therefore, the probability of students in magnet school to take college entrance examination becomes from higher than the probability of students in regular schools to lower.

The middle panel presents results for normalized college entrance scores. In column (1)-(3) where we only control for county fixed effects and year fixed effects, the treatment effect decreases from 1.028 to 0.835, then to 0.629 with the band of neighborhood changing from 1.5 standard deviations to 1 standard deviation, then to 0.5 deviation. After sex and age are added, there are only slight changes to the results; when the band is 1.5 standard deviations (column (4)), the treatment effect is 1.026; when the band is 1 standard deviation (column(5)), the treatment effect is 0.833; and when the band is 0.5 standard deviation (column(6)), the treatment effect is 0.628. If we add sex, age and middle school fixed effects, the treatment effects increase a little bit; they are 1.101 with band of 1.5 standard deviations (column(7)), 0.916 with band of 1 standard deviation (column(8)) and 0.752 with band of 0.5 standard deviation (column(9)). From graph in Panel B in Figure 4.3, we can see that the college entrance examination score is a positive function of high school entrance score; therefore, with the band of neighborhood around the cutoff line becoming narrower, the difference between the mean value of treated students' college entrance score becomes smaller.

The bottom panel presents results for the probability of qualifying to enter college. In column (1)-(3) where we only control for county fixed effects and year fixed effects, the treatment effect decreases from 0.517 to 0.438, then to 0.335 with the band of neighborhood changing from 1.5 standard deviations to 1 standard deviation, then to 0.5 deviation. After sex and age are added, there are only slight changes to the results; when the band is 1.5 standard deviations (column (4)), the treatment effect is 0.516; when the band is 1 standard deviation (column(5)), the treatment effect is 0.437; and when the band is 0.5 standard deviation (column(6)), the treatment effect is 0.335. If we add sex, age and middle school fixed effects, the treatment effects increase a little bit; they are 0.555 with band of 1.5 standard deviations (column(7)), 0.483 with band of 1 standard deviation (column(8)) and 0.404 with band of 0.5 standard deviation (column(9)). The reason for the decrease of treatment effects with the decrease in the band of neighborhood around the cutoff line is the same as that discussed in the last paragraph.

4.6 Magnet schools versus regular schools

A remaining question in interpreting our results of a positive effect of better quality schools on learning is, in what dimensions do magnet schools actually differ from regular schools? The school questionnaire provides annual information on a variety of school characteristics that can inform the answer to this question. We measure school quality for an entering class by the 4-year average values of school indicators that span the years that they attended the school. For example, for students starting high school in September 1997 and graduating in June 2000, we take mean values for the years 1997 to 2000.

Table4.9 presents summary statistics on 3 groups of school variables for both magnet and regular schools, and for each of the 7 counties investigated. The upper panel shows the share of teachers with different teacher rankings. These rankings reflect systematic, multidimensional annual evaluations of teacher performance by school principals, fellow teachers, and students. The ratios of teachers having advanced and first class titles in magnet schools are much higher than those in regular schools, while the ratios of those with second and third class titles is much lower. The middle panel shows the ratios of teachers with different education levels in each county. In magnet schools, the ratios of teachers with four-year college education are much higher than regular schools while the ratio of teachers with three-year college education level in regular schools is much higher than in magnet schools, with the exception of one county. In both types of schools, the ratio of teachers with secondary specialized schooling and high school education are very similar. Finally, the bottom panel describes the basic endowments of schools. Magnet schools have larger total area, larger building areas, better equipment, and more library books.

Table 4.10 presents simple regression-based comparisons between magnet and regular schools. From the upper panel, we can see that the share of teachers with advanced titles in magnet schools is 0.151 higher than in regular schools, the share of teachers with first class titles in good schools is 0.092 higher than that in bad schools. However, the shares of teachers with second and third class titles are 0.093 and 0.049 greater in regular schools. All of these differences are statistically significant. In the middle panel, we see that the share of teachers with four-year college education is 0.402 higher in magnet schools than in regular schools, while the share of teachers with three-year college is higher in regular schools by 0.399 (both statistically significant at the 1% level). However, there are not significant differences in the share of teachers with secondary specialized schooling or who are high school graduates. The bottom panel examines school endowments. Magnet schools are larger than regular schools by 57.63 thousand square meters, their building area is larger by 13.65 thousand square meters**,** the number of library books is 72,770 greater, and the probability for equipments to achieve national criteria is 0.451 greater. All differences in the bottom panel are significantly different from zero at the 1% level except for equipment quality, which is significant at the 5% significance level.

Because we only are able to estimate treatment effects for 7 counties, we are unable to credibly identify the relative contribution of different school characteristics in a multivariate framework. It does appear that magnet schools are clearly advantaged with respect to teacher quality, physical infrastructure, and equipment. Unfortunately, further analysis of what school inputs really matter for learning is beyond the scope of this paper.

4.7 Conclusion

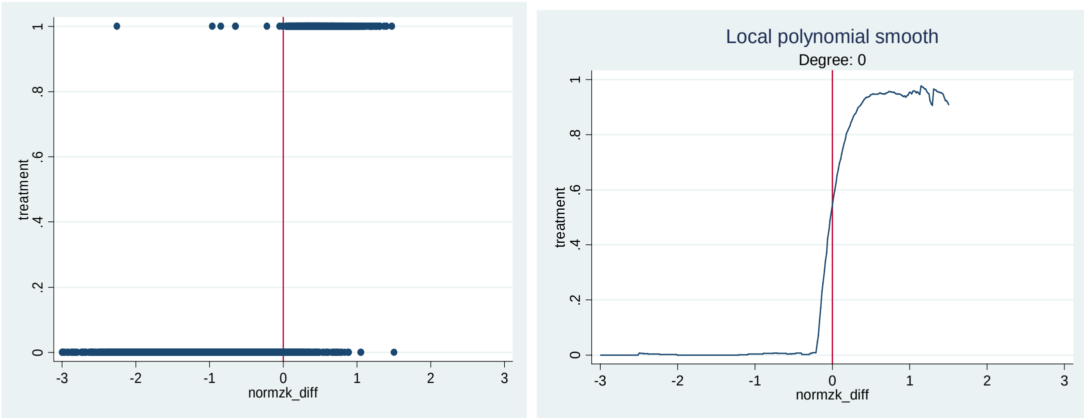
Whether higher school quality can improve students' education achievements is a very important and highly debated question in the economics of education. This paper uses China's magnet school system and a regression discontinuity design approach to credibly estimate the effects of higher school quality on students' education achievements.

Using data from seven counties in Gansu province in China, according to our preferred estimates which control for sex, age and middle school attended, we find that attending a magnet high school can decrease students' probability to take college entrance examination by 14.9 percent, but increase students' college entrance score by

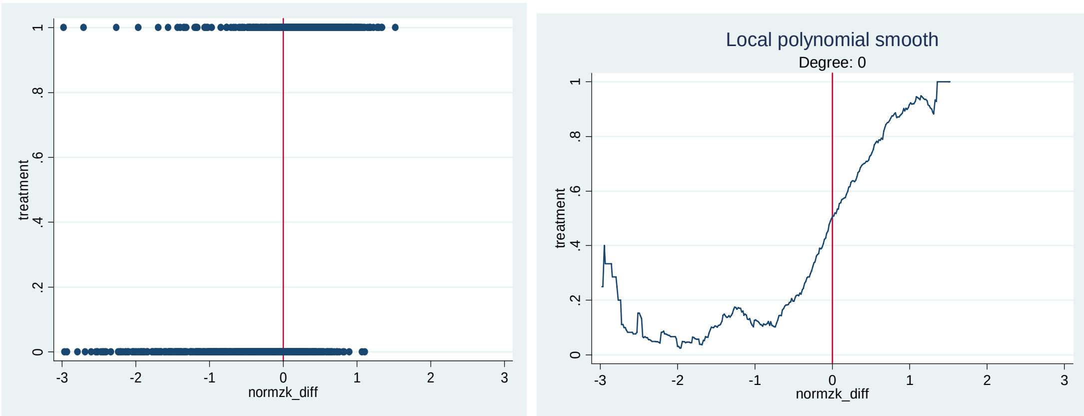
0.256 standard deviation and increase students’ probability of qualifying for college by 22.5 percent.

Figure 4.1 Student assignment

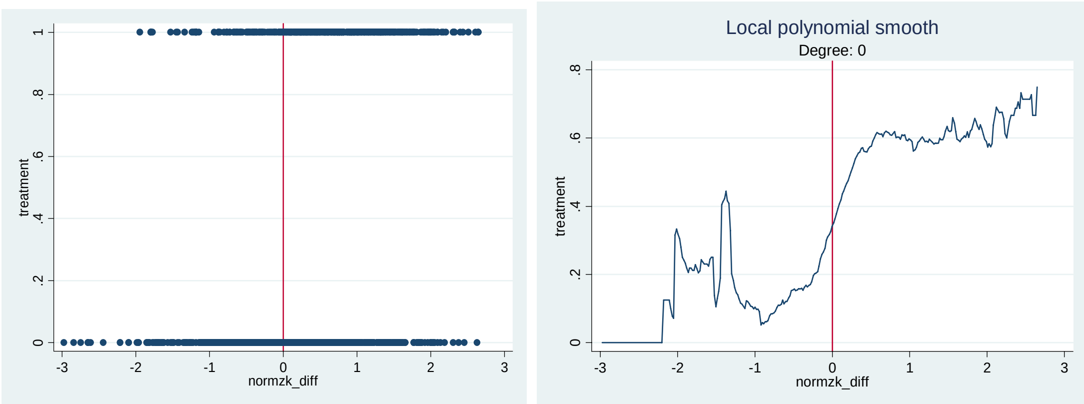
Panel A. County16



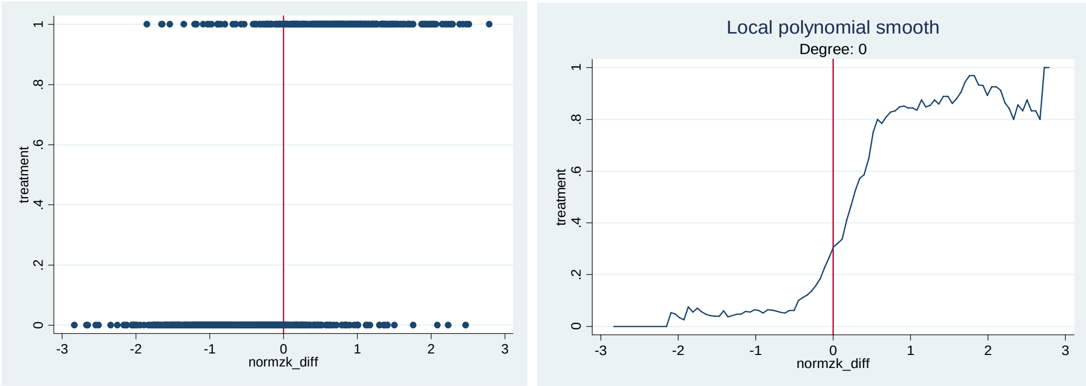
Panel B. County 18



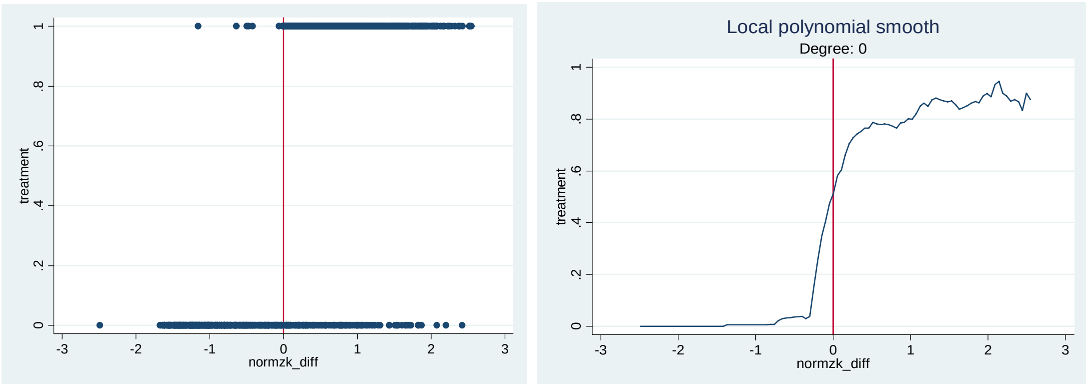
Panel C. County 19



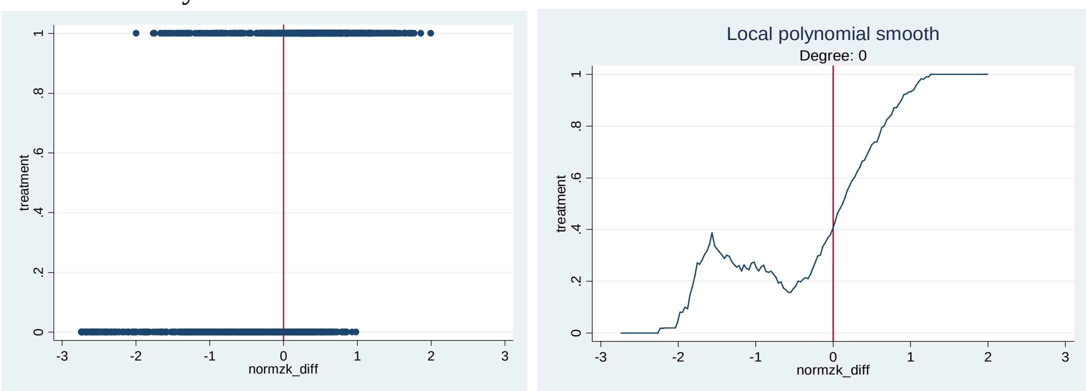
Panel D. County 20



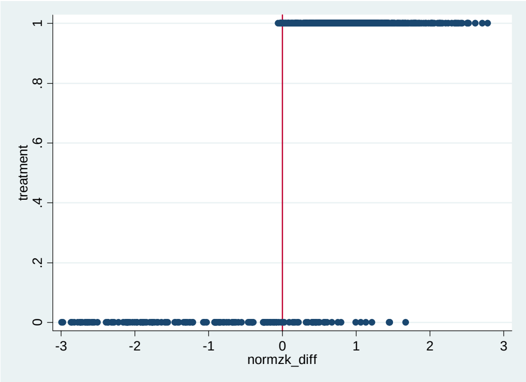
Panel E. County 26



Panel F. County 28



Panel G. County 29



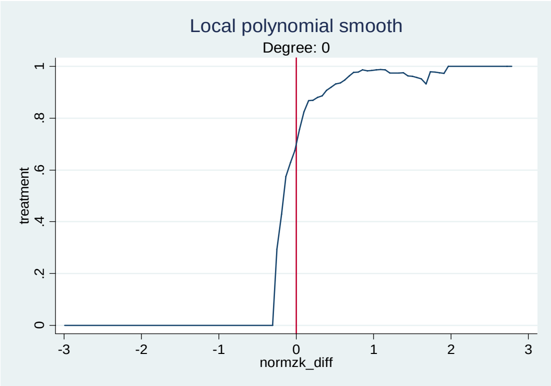
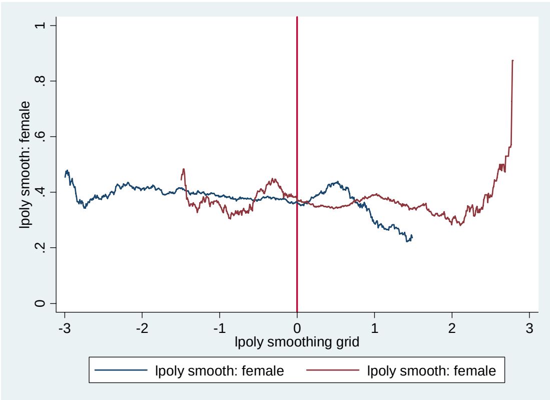


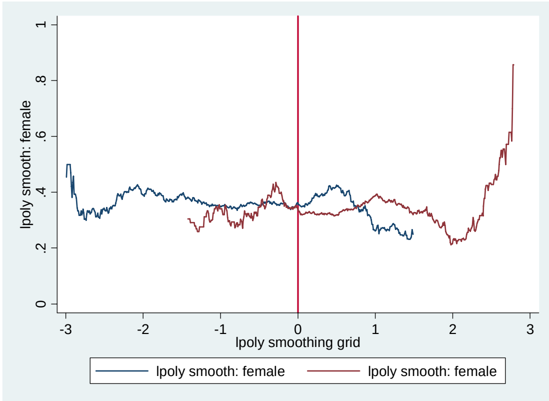
Figure 4.2 Pre-assumption test

Panel A Female

Whole sample

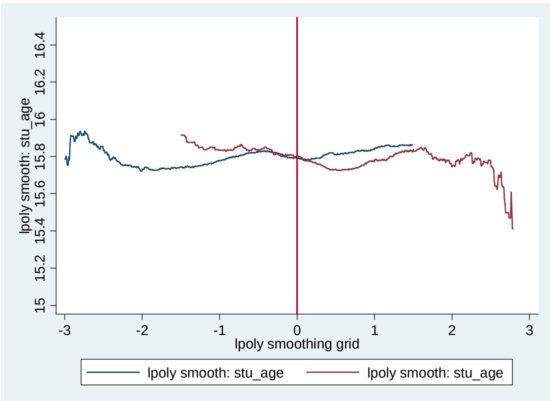


Sub-sample



Panel B Student age

Whole sample



Sub-sample

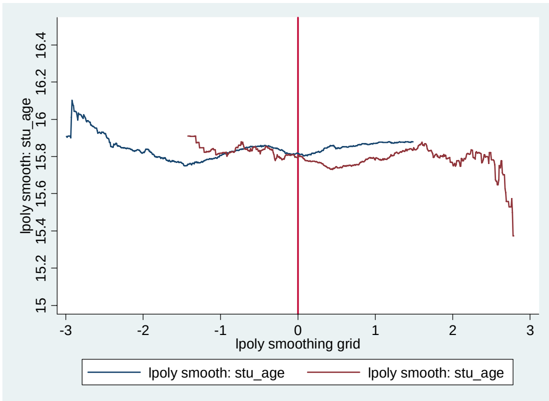
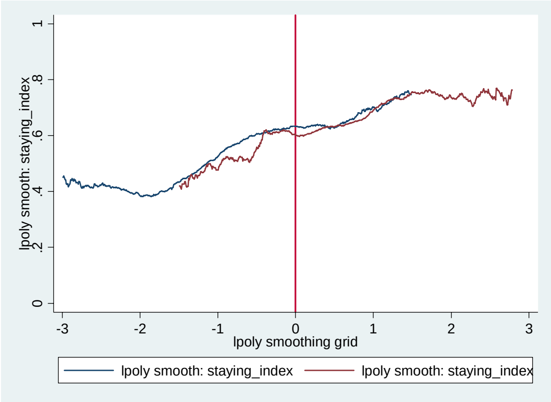
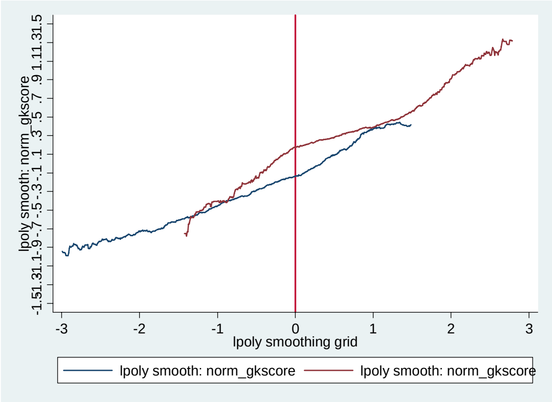


Figure 4.3 Effects of school quality

Panel A Having college entrance score

Panel B Normalized college entrance score





Panel C College eligibility

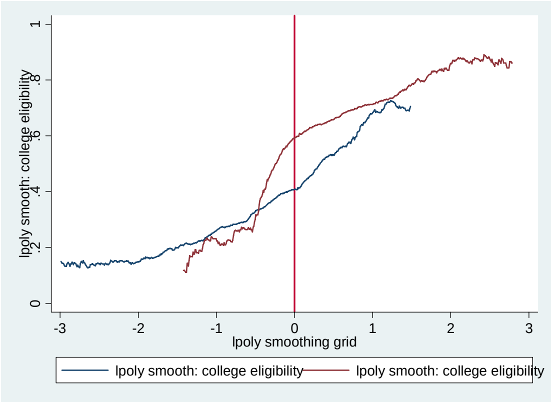


Table 4.1 Sample description

County	Enter year	Num. of schools	Num. of students	Highest cutoffline	Magnet cutoff line	Magnet average entering score	Magnet students (%)	Magnet students with high school entrance score lower than cutoff line(%)	Regular students with high school entrance score higher than cutoff line(%)	Students with missing college entrance score (%)	Students not admitted by college (%)	Students admitted by college (%)	Students with missing college admission information (%)	Students with missing gender (%)
Panel 1														
11	1999	1	486	NA	NA	619.08	100.00%	NA	NA	46.50%	0.00%	0.00%	100.00%	0.00%
11	2001	1	313	NA	NA	597.27	100.00%	NA	NA	22.04%	0.00%	0.00%	100.00%	0.00%
12	1997	1	29	NA	NA	370.66	100.00%	NA	NA	10.34%	0.00%	0.00%	100.00%	0.00%
12	1998	1	29	NA	NA	381.52	100.00%	NA	NA	17.24%	0.00%	0.00%	100.00%	0.00%
12	2001	1	426	587	587	568.49	100.00%	54.69%	NA	19.01%	0.00%	0.00%	100.00%	0.00%
13	1999	1	623	520	520	597.80	100.00%	0.16%	NA	54.41%	0.00%	0.00%	100.00%	0.00%
13	2000	1	943	580	580	591.28	100.00%	38.60%	NA	48.78%	0.00%	0.00%	100.00%	0.00%
13	2001	1	1107	560	560	597.85	100.00%	13.73%	NA	43.72%	0.00%	0.00%	100.00%	0.00%
20	1999	1	196	444	444	486.08	100.00%	11.73%	NA	30.10%	0.00%	0.00%	100.00%	100.00%
23	1997	1	350	575	575	568.72	100.00%	54.86%	NA	63.14%	0.00%	36.86%	63.14%	0.00%
23	1998	1	349	574	574	583.74	100.00%	35.53%	NA	51.29%	0.00%	48.71%	51.29%	0.00%
23	1999	1	352	580	580	592.85	100.00%	27.84%	NA	53.69%	0.00%	46.31%	53.69%	0.00%
23	2001	1	363	589	589	590.90	100.00%	42.15%	NA	62.81%	0.00%	37.19%	62.81%	0.00%
24	1999	1	384	498.5	498.5	531.96	100.00%	20.57%	NA	40.89%	14.32%	44.79%	40.89%	0.00%
24	2000	1	456	509	509	547.66	100.00%	21.49%	NA	22.15%	0.00%	0.00%	100.00%	0.00%
27	1997	1	22	578	578	551.32	100.00%	77.27%	NA	77.27%	0.00%	0.00%	100.00%	0.00%
27	1998	1	24	490	490	538.58	100.00%	0.00%	NA	62.50%	0.00%	0.00%	100.00%	0.00%
28	1999	1	584	566.5	566.5	557.35	100.00%	41.27%	NA	6.34%	25.51%	68.15%	6.34%	0.00%
30	1997	1	443	NA	NA	510.30	100.00%	NA	NA	34.76%	42.89%	22.35%	34.76%	0.00%
30	1999	1	621	NA	NA	539.86	100.00%	NA	NA	8.86%	28.50%	62.64%	8.86%	0.00%

30	2000	1	502	NA	NA	591.94	100.00%	NA	NA	85.26%	0.00%	14.74%	85.26%	0.00%
Panel 2														
17	1997	4	488	545	545	432.87	41.19%	92.04%	3.48%	59.63%	25.82%	5.33%	68.85%	0.00%
17	1998	4	352	540	540	462.84	28.13%	79.80%	8.70%	69.60%	0.28%	1.42%	98.30%	0.00%
17	1999	4	512	558	558	500.20	37.30%	80.10%	3.12%	56.25%	21.68%	22.07%	56.25%	0.00%
17	2000	4	544	564	564	490.39	38.42%	87.56%	2.69%	48.71%	22.79%	28.49%	48.71%	0.00%
17	2001	5	627	560	560	519.96	31.74%	71.86%	3.74%	35.25%	47.21%	17.54%	35.25%	0.00%
18	1998	10	485	576.5	576.5	574.15	41.44%	44.78%	1.76%	0.00%	0.00%	0.00%	100.00%	0.00%
18	2000	10	784	565	565	565.70	44.52%	40.69%	9.43%	0.00%	0.00%	0.00%	100.00%	0.00%
19	2001	5	2279	545	545	548.38	41.55%	39.28%	36.64%	17.11%	0.00%	0.00%	100.00%	0.04%
20	2001	5	1630	562	562	574.11	31.04%	32.41%	14.59%	41.23%	0.00%	0.00%	100.00%	75.09%
22	2001	2	315	400	400	442.58	85.40%	30.48%	6.52%	82.86%	82.86%	17.14%	0.00%	0.00%
28	2001	2	1479	561.5	561.5	533.74	50.71%	71.33%	2.06%	41.31%	0.95%	14.54%	84.52%	0.00%
24	2001	2	513	504	504	559.68	63.35%	0.31%	100.00%	22.42%	10.72%	37.04%	52.24%	0.00%
25	2001	2	581	406	406	475.23	22.55%	6.11%	100.00%	58.00%	0.00%	0.00%	100.00%	0.00%
26	1998	2	408	488	488	553.43	76.96%	3.50%	79.79%	53.68%	0.00%	0.00%	100.00%	0.00%
27	2000	2	104	430	430	488.41	25.96%	0.00%	76.62%	41.35%	0.00%	0.00%	100.00%	0.00%
Panel 3														
19	1997	5	1252	510	431	499.91	13.74%	0.00%	65.93%	26.36%	0.00%	0.00%	100.00%	0.00%
19	1998	5	1291	510	480	557.76	13.56%	3.43%	68.10%	21.84%	0.00%	0.00%	100.00%	0.00%
19	1999	5	1460	530	500	551.04	15.48%	0.00%	59.00%	19.73%	0.00%	0.00%	100.00%	47.88%
20	1997	5	1092	490	485	506.93	36.36%	28.46%	14.68%	50.09%	0.00%	0.00%	100.00%	0.00%
Panel 4														
15	1998	4	1352	432	NA	571.64	34.76%	NA	NA	40.83%	0.00%	0.00%	100.00%	0.00%
15	1999	4	1274	408	NA	583.30	33.36%	NA	NA	26.14%	0.00%	0.00%	100.00%	0.00%
15	2000	4	1628	417	NA	605.39	29.79%	NA	NA	30.22%	24.57%	45.21%	30.22%	0.00%
15	2001	4	1741	517	NA	594.90	40.61%	NA	NA	40.84%	0.00%	59.16%	40.84%	0.00%

27	2001	2	474	431	NA	520.08	88.82%	NA	NA	46.84%	0.00%	0.00%	100.00%	0.00%
28	1997	4	644	551	NA	544.43	20.81%	NA	NA	31.52%	24.53%	12.27%	63.20%	0.00%
28	1998	2	737	588.5	NA	587.24	26.87%	NA	NA	8.82%	44.78%	21.57%	33.65%	0.00%
29	1999	2	425	380	NA	537.35	80.24%	NA	NA	23.29%	43.53%	32.00%	24.47%	0.00%
29	2000	3	775	480	NA	531.89	65.03%	NA	NA	6.84%	6.71%	0.65%	92.65%	0.00%
29	2001	3	1159	490	NA	514.36	63.93%	NA	NA	13.11%	32.01%	28.82%	39.17%	0.00%
30	1998	2	720	NA	NA	590.60	49.86%	NA	NA	68.61%	8.47%	22.92%	68.61%	0.00%
30	2001	2	1359	NA	NA	572.82	50.18%	NA	NA	24.87%	0.00%	25.75%	74.25%	0.00%
Panel 5														
21	1999	2	507	456	456	495.54	75.35%	25.92%	24.80%	55.42%	55.23%	44.58%	0.20%	0.00%
21	2000	3	855	478.5	478.5	519.37	63.63%	8.46%	12.54%	45.96%	45.85%	54.04%	0.12%	0.00%
21	2001	2	875	412	412	492.86	77.14%	3.56%	61.00%	41.83%	41.14%	58.17%	0.69%	0.00%
22	1999	2	159	375	375	445.50	86.79%	17.39%	47.62%	47.17%	47.17%	52.83%	0.00%	0.00%
22	2000	2	179	390	390	435.07	86.59%	27.10%	41.67%	53.63%	53.63%	46.37%	0.00%	0.00%
23	2000	2	508	583	583	601.29	66.14%	19.94%	36.63%	62.01%	6.50%	31.50%	62.01%	0.00%
Panel 6														
20	2000	5	1270	516	516	568.97	41.18%	5.93%	49.40%	19.37%	0.00%	0.00%	100.00%	74.25%
25	2000	2	577	490	490	572.59	54.94%	0.00%	25.00%	49.05%	0.00%	0.00%	100.00%	45.06%
Panel 7														
16	1997	3	531	560	560	597.35	51.32%	0.00%	3.86%	68.42%	0.00%	0.00%	100.00%	0.00%
16	1998	5	1195	564	564	613.75	27.28%	1.53%	8.52%	64.69%	0.00%	0.00%	100.00%	0.00%
16	1999	4	762	554.5	554.5	609.19	42.52%	0.00%	12.56%	56.96%	0.00%	0.00%	100.00%	0.00%
16	2000	3	784	560	560	599.59	34.69%	1.10%	8.98%	44.64%	0.00%	0.00%	100.00%	0.00%
16	2001	3	1032	568	568	597.85	37.75%	0.00%	2.64%	34.85%	0.00%	0.00%	100.00%	0.00%
18	1997	9	463	530	530	547.49	46.44%	26.51%	13.71%	0.00%	0.00%	0.00%	100.00%	0.00%
18	1999	9	593	540	540	557.81	43.34%	22.96%	30.65%	100.00%	0.00%	0.00%	100.00%	0.00%
19	2000	5	1805	540	540	566.20	39.26%	18.90%	39.47%	18.72%	0.00%	0.00%	100.00%	0.00%

20	1998	5	1137	518.5	518.5	558.14	34.92%	14.61%	16.62%	51.36%	0.00%	0.00%	100.00%	0.00%
26	1999	3	391	543	543	579.95	55.87%	0.91%	34.68%	46.68%	0.00%	0.00%	100.00%	0.00%
26	2001	3	812	532	532	575.38	47.60%	1.03%	22.77%	46.62%	0.00%	0.00%	100.00%	0.00%
28	2000	4	1572	553	553	563.09	48.47%	28.87%	23.70%	25.45%	11.64%	17.94%	70.42%	0.00%
29	1997	2	386	438	438	501.42	81.09%	2.56%	35.62%	25.91%	40.67%	33.16%	25.91%	0.00%
29	1998	2	572	430	430	495.91	86.21%	0.00%	10.13%	13.09%	69.28%	17.45%	13.09%	0.00%

Table 4.2 Summary statistics

Variable	Normalized GK score	Eligible for college	Female	Age	Treatment	Eligible
All students						
Mean	NA	NA	0.38	15.78	0.44	0.50
S.D.	NA	NA	0.48	0.36	0.50	0.50
Observation	NA	NA	12035	12035	12035	12035
Students having college entrance score						
Mean	0.00	0.49	0.35	15.80	0.49	0.56
S.D.	1.00	0.50	0.48	0.39	0.50	0.50
Observation	7102	7102	7102	7102	7102	7102

(1) Treatment =1 if the student enters good high school, which is defined as the school whose students' average entering grade is the highest in this county and year

Eligibility =1 if the student's entering grade is higher than the cutoff line of the good high school defined above

Table 4.3 School assignments and the first stage regressions

Dependent variable: treatment	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eligible	0.588	0.468	0.469	0.381	0.579	0.492	0.493	0.409
	(0.007)***	(0.013)***	(0.013)***	(0.012)***	(0.009)***	(0.017)***	(0.017)***	(0.016)***
Female	YES	NO	YES	YES	YES	NO	YES	YES
Age	YES	NO	YES	YES	YES	NO	YES	YES
County fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Middle school fixed effects	YES	NO	NO	YES	YES	NO	NO	YES
5 order of polynomial terms of normalized high school entrance score	NO	YES	YES	YES	NO	YES	YES	YES
Observations	12035	12035	12035	12035	7102	7102	7102	7102
R-squared	0.58	0.53	0.53	0.60	0.57	0.53	0.53	0.59
F value (Eligible=0)	7189.60	1344.91	1346.00	962.89	3920.22	887.61	889.56	643.21

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.4 Pre-assumption tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female	Female	Age	Age	Female	Female	Age	Age
Eligible	0.012		0.018		0.026		-0.002	
	(0.019)		(0.013)		(0.023)		(0.018)	
Treatment(Eligible as IV)		0.031		0.048		0.063		-0.006
		(0.049)		(0.034)		(0.057)		(0.044)
5 order of polynomial terms of normalized high school entrance score	YES	YES	YES	YES	YES	YES	YES	YES
County fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Middle school fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	12035	12035	12035	12035	7102	7102	7102	7102
R-squared	0.03	0.03	0.17	0.17	0.05	0.05	0.13	0.13

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.5 Effects of school quality on students' decision to take college entrance examination

	Dependent Variable : Taking college entrance examination=1								
Treatment	-0.046			-0.046			-0.056		
	(0.011)***			(0.011)***			(0.012)***		
Eligible		-0.067			-0.066			-0.057	
		(0.017)***			(0.017)***			(0.017)***	
Treatment (Eligible as IV)			-0.143			-0.141			-0.149
			(0.036)***			(0.036)***			(0.045)***
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
5 order of polynomial terms of normalized high school entrance score	YES	YES	YES	YES	YES	YES	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	12035	12035	12035	12035	12035	12035	12035	12035	12035
R-squared	0.15	0.15	0.15	0.16	0.16	0.15	0.20	0.20	0.20

	Dependent Variable : Taking college entrance examination=1								
Treatment	-0.044			-0.045			-0.054		
	(0.011)***			(0.011)***			(0.012)***		
Treatment*Normalized high school entrance score	0.011			0.010			0.034		
	(0.011)			(0.011)			(0.011)***		
Eligible		-0.025			-0.025			-0.024	
		(0.015)			(0.015)			(0.015)	
Eligible*Normalized high school entrance score		-0.049			-0.051			-0.016	
		(0.014)***			(0.014)***			(0.015)	
Treatment (Eligible as IV)			-0.059			-0.060			-0.058
			(0.027)**			(0.027)**			(0.031)*
Treatment*Normalized high school entrance score (Eligible*Normalized high school entrance score as IV)			-0.051			-0.054			-0.008
			(0.021)**			(0.021)***			(0.022)
Normalized high school entrance score	0.081	0.103	0.112	0.081	0.103	0.112	0.055	0.072	0.074
	(0.006)***	(0.008)***	(0.009)***	(0.006)***	(0.008)***	(0.009)***	(0.007)***	(0.008)***	(0.010)***
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES

Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES		YES	YES
Observations	12035	12035	12035	12035	12035	12035	12035	12035	12035
R-squared	0.15	0.15	0.15	0.16	0.16	0.15	0.20	0.20	0.20

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Note:

(1) Treatment =1 if the student enters good high school, which is defined as the school whose students' average entering grade is the highest in this county and year
Eligible=1 if the student's entering grade is higher than the cutoff line of the good high school defined above

Table 4.6 Effects of school quality on students' college entrance score and college eligibility

	Dependent Variable : Normalized Gaokao Grade								
Treatment	0.110 (0.029)***			0.108 (0.029)***			0.089 (0.031)***		
Eligible		0.124 (0.042)***			0.126 (0.042)***			0.105 (0.043)**	
Treatment (Eligible as IV)			0.252 (0.086)***			0.255 (0.086)***			0.256 (0.106)**
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
5 order of polynomial terms of normalized high school entrance score	YES	YES	YES	YES	YES	YES	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7102	7102	7102	7102	7102	7102	7102	7102	7102
R-squared	0.23	0.23	0.23	0.23	0.23	0.23	0.26	0.26	0.26

	Dependent Variable : College entrance score is higher than lowest cutoff line=1								
Treatment	0.058 (0.014)***			0.057 (0.014)***			0.045 (0.016)***		
Eligible		0.104 (0.021)***			0.105 (0.021)***			0.092 (0.022)***	
Treatment (Eligible as IV)			0.212 (0.043)***			0.214 (0.043)***			0.225 (0.054)***
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school	NO	NO	NO	NO	NO	NO	YES	YES	YES

fixed effect									
5 order of polynomial terms of normalized high school entrance score	YES	YES	YES	YES	YES	YES	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effect	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7102	7102	7102	7102	7102	7102	7102	7102	7102
R-squared	0.22	0.23	0.21	0.23	0.23	0.21	0.25	0.25	0.23

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Note:

(1) Treatment =1 if the student enters good high school, which is defined as the school whose students' average entering grade is the highest in this county and year

Eligible=1 if the student's entering grade is higher than the cutoff line of the good high school defined above

Table 4.7 Unconditional effects of entering good school on the probability to enter college

	Dependent Variable : College entrance score is higher than lowest cutoff line=1								
Treatment	0.015			0.014			-0.001		
	(0.011)			(0.011)			(0.012)		
Eligible		0.057			0.058			0.048	
		(0.016)***			(0.016)***			(0.017)***	
Treatment (Eligible as IV)			0.114			0.116			0.120
			(0.033)***			(0.033)***			(0.042)***
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
County fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES
5 order polynomial function of normalized high school entrance score	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	12035	12035	12035	12035	12035	12035	12035	12035	12035
R-squared	0.17	0.17	0.16	0.17	0.17	0.16	0.20	0.20	0.19

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

(1) Treatment =1 if the student enters good high school, which is defined as the school whose students' average entering grade is the highest in this county and year

Eligible=1 if the student's entering grade is higher than the cutoff line of the good high school defined above

(2) Students with missing college entrance score are assumed not to enter college

Table 4.8 Effects of school quality on students' performance in different neighborhood

Dependent Variable : Having college entrance score=1

	1.5 S.D.	1 S.D.	0.5 S.D.	1.5 S.D.	1 S.D.	0.5 S.D.	1.5 S.D.	1 S.D.	0.5 S.D.
Treatment (Eligible as IV)	0.099	0.047	-0.071	0.096	0.045	-0.072	0.085	0.038	-0.090
	(0.014)***	(0.017)***	(0.031)**	(0.014)***	(0.017)***	(0.031)**	(0.018)***	(0.021)*	(0.039)**
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
Observations	10107	7885	4277	10107	7885	4277	10107	7885	4277
R-squared	0.13	0.14	0.13	0.13	0.14	0.13	0.18	0.20	0.20

Dependent Variable : Normalized college entrance score

	1.5 S.D.	1 S.D.	0.5 S.D.	1.5 S.D.	1 S.D.	0.5 S.D.	1.5 S.D.	1 S.D.	0.5 S.D.
Treatment (Eligible as IV)	1.028	0.835	0.629	1.026	0.833	0.628	1.101	0.916	0.752
	(0.039)***	(0.045)***	(0.075)***	(0.039)***	(0.045)***	(0.075)***	(0.049)***	(0.056)***	(0.101)***
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
Observations	6188	4881	2642	6188	4881	2642	6188	4881	2642
R-squared	0.03	0.03	0.04	0.03	0.03	0.04	0.07	0.08	0.09

Dependent Variable : College entrance score is higher than lowest cutoff line=1

	1.5 S.D.	1 S.D.	0.5 S.D.	1.5 S.D.	1 S.D.	0.5 S.D.	1.5 S.D.	1 S.D.	0.5 S.D.
Treatment (Eligible as IV)	0.517	0.438	0.335	0.516	0.437	0.335	0.555	0.483	0.404
	(0.020)***	(0.024)***	(0.041)***	(0.020)***	(0.024)***	(0.041)***	(0.025)***	(0.030)***	(0.056)***
Female	NO	NO	NO	YES	YES	YES	YES	YES	YES
Imputed age	NO	NO	NO	YES	YES	YES	YES	YES	YES
Middle school fixed effect	NO	NO	NO	NO	NO	NO	YES	YES	YES
Observations	6188	4881	2642	6188	4881	2642	6188	4881	2642
R-squared	0.07	0.06	0.09	0.07	0.06	0.09	0.09	0.09	0.13

Standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%

(1) Treatment =1 if the student enters good high school, which is defined as the school whose students' average entering grade is the highest in this county and year

Eligible=1 if the student's entering grade is higher than the cutoff line of the good high school defined above

(2) Each regression above includes county fixed effects and year fixed effects.

(3) S.D. is the standard deviation of high school entrance examination score

Table 4.9 Comparison between good schools and bad schools

	Ratio of teachers with different titles							
	Advanced title		First class title		Second class title		Third class title	
County	Good	Bad	Good	Bad	Good	Bad	Good	Bad
16	0.26	0.03	0.40	0.27	0.30	0.48	0.01	0.02
18	NA	NA	NA	NA	NA	NA	NA	NA
19	0.17	0.03	0.38	0.38	0.36	0.50	0.00	0.07
20	0.17	0.07	0.42	0.43	0.34	0.37	0.09	0.13
26	0.12	0.04	0.40	0.36	0.45	0.39	0.03	0.18
28	0.20	0.11	0.46	0.33	0.26	0.35	0.00	0.00
29	0.04	NA	0.49	NA	0.43	NA	0.02	NA

	Ratio of teachers with different education level							
	Four year college		Three year college		Secondary specialized school		High school	
County	Good	Bad	Good	Bad	Good	Bad	Good	Bad
16	0.86	0.25	0.08	0.70	0.00	0.04	0.00	0.02
18	NA	NA	NA	NA	NA	NA	NA	NA
19	0.63	0.32	0.37	0.68	0.00	0.00	0.00	0.00
20	0.75	0.28	0.25	0.68	NA	0.03	NA	0.02
26	0.48	0.32	0.51	0.64	0.01	0.02	NA	0.03
28	0.74	0.20	0.25	0.72	0.01	0.00	0.00	0.08
29	0.31	0.38	0.58	0.63	0.08	NA	0.09	NA

	School characteristics							
	School area (Square meters)		Building area (Square meters)		Books		Does equipment satisfy criteria	
County	Good	Bad	Good	Bad	Good	Bad	Good	Bad
16	122866	24822.48	25425.70	4698.19	127000	3566.34	1.00	0.50
18	NA	NA	NA	NA	NA	NA	NA	NA
19	55360	18749.75	15486	7062.50	66250	28314.83	1.00	1.00
20	53020	11534.25	6520	3719.31	9500	10336.13	0.00	0.25
26	57276	26329	13500	6668.44	137000	5800	1.00	0.00
28	47923	52000	28500	6009	22200	13494.63	NA	0.00
29	67000	63000	9000	4060	8000	7775	1.00	1.00

Table 4.10 Different characteristics between magnet schools and regular schools

	Share of teachers with different titles			
	Advanced title	First class title	Second class title	Third class title
Magnet school	0.151	0.092	-0.093	-0.049
	(0.019)***	(0.050)*	(0.040)**	(0.019)**
County fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Observations	37	39	39	38
R-squared	0.72	0.23	0.28	0.51

	Share of teachers with different education level			
	Four year college	Three year college	Secondary specialized school	High school
Magnet school	0.402	-0.399	-0.020	-0.007
	(0.071)***	(0.067)***	(0.015)	(0.010)
County fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Observations	39	39	35	33
R-squared	0.54	0.56	0.31	0.46

	School characteristics			
	School area (10 thousand square meters)	Building area (10 thousand square meters)	Books (10 thousand)	Does equipment satisfy criteria
Magnet school	5.763	1.365	7.277	0.451
	(0.754)***	(0.154)***	(1.128)***	(0.177)**
County fixed effects	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
Observations	39	39	38	34
R-squared	0.71	0.76	0.63	0.42

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

References

[1] Ahn, Hyungtaik, and James L. Powell (1993). *Semiparametric estimation of censored selection models with a nonparametric selection mechanism,* Journal of Econometrics 58 3-29.

[2] Altonji, Joseph G. and Thomas A. Dunn (1996). *Using siblings to estimate the effect of school quality on wages*, The Review of Economics and Statistics, vol. 78, No. 4, 665-671.

[3] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005). *Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools,* Journal of Political Economy, vol. 113, No. 1, 151-184.

[4] Angrist, Joshua D. and Victor Lavy (1999). *Using Maimonides' rule to estimate the effect of class size on scholastic achievement,* The Quarterly Journal of Economics, 533-575.

[5] Angrist, Joshua D. and Victor Lavy (2001). *Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools,* Journal of Labor Economics, vol. 19, No. 2, 343-369.

[6] Benerjee, A., S. Cole, E. Duflo, and L. Linden (2006) Remedying education: evidence from two randomized experiments in India. MIT working paper.

[7] Betts, Julian R. (1995). *Does school quality matter? Evidence from the national longitudinal surveys of youth* The Review of Economics and Statistics, vol. 77, No. 2, 231-250.

[8] Betts, Julian R. (2001). *The impact of school resources on women's earnings and educational attainment: Findings from the national longitudinal survey of young women* Journal of Labor Economics, vol. 19, No. 3.

[9] Card, David, and Alan B. Krueger (1992a). *Does school quality matter? Returns to education and the characteristics of public schools in the United States,* Journal of Political Economy, Vol. 100, No. 1. 1-40.

[10] Card, David and Alan B. Krueger (1992b), *School quality and black-white relative earnings: A direct assessment,* The Quarterly Journal of Economics, vol. 107, No. 1, 151-200.

[11] Case, Anne and Angus Deaton (1999). *School inputs and educational outcomes in South Aftica,* The Quarterly Journal of Economics, vol. 114, No. 3. 1047-1084.

[12] Dee, Thomas S. (2001). *Teachers, race and student achievement in a randomized experiment,* NBER Working Paper 8432.

[13] Eide, Eric and Mark H. Showalter (1998). *The effect of school quality on student performance: a quantile regression approach,* Economics Letters, 58, 345-350.

[14] Greenwald, Rob, Larry V. Hedges, and Richard D. Laine (1996). *The effect of school resources on student achievement,* Review of Educational Research, vol. 66, No. 3, 361-396.

[15] Goldhaber, D. Dan and Dominic J. Brewer (1997). *Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity,* The Journal of Human Resources, vol. 32, No. 3, 505-523.

[16] Gould, Eric D., Victor Lavy, and M. Daniele Paserman (2004). *Immigrating to opportunity: Estimating the effect of school quality using a natural experiment on Ethiopians in Israel,* The Quarterly Journal of Economics, 489-526.

[17] Hahn, Jingyong, Petra Todd, and Wilbert van der Klaauw (2001). *Identification and estimation of treatment effects with a regression-discontinuity design,* Econometrica, 69, 201-209.
[18] Heckman, James J. (2005). *China's human capital investment,* China Economic Review 16 50-70.
[19] Heckman, James J. and R. Robb (1985). *Alternative methods for evaluating the impact of interventions* Longitudinal Analysis of Labor Market Data, edited by J. Heckman and B. Singer, Cambridge University Press.
[20] Hoxby, Caroline M. (2000). *The effects of class size on student achievement: new evidence from population variation,* The Quarterly Journal of Economics, November 2000, 1239-1285.
[21] Hoxby, Caroline M. (2007). *The effects of New York City's charter schools on student achievement,* Working paper.
[22] Krueger, Alan B. (1999). *Experimental estimates of education production functions,* The Quarterly Journal of Economics, 497-532.
[23] Krueger, Alan B. and Diane M. Whitmore (1999). *The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR,* Working Paper # 427, Industrial Relations Section, Princeton University.
[24] Lai, Fang, Elisabeth Sadoulet, and Alain de Janvry (2006). *The effect of school and teacher quality on student performance: Using a natural experiment from the middle school reforms in Beijing,* Working Paper.
[25] Lavy, Victor (2002). *Evaluating the effect of teachers' group performance incentives on pupil achievement,* Journal of Political Economy, vol. 110, No. 6, 1286-1317.
[26] Lee, David S. (2003). *Randomized experiments from non-random selection in U.S. House Elections,* NBER working paper # 8441.
[27] Lee, David S., Enrico Moretti, and Matthew J. Butler (2004). *Do voters affect or elect policies? Evidence from the U.S. House,* Quarterly Journal of Economics, 119, 807-860.
[28] Newhouse, David and Kethleen Beegle (2005). *The effect of school type on academic achievement: Evidence from Indonesia* World Bank Policy Research Working Paper 3604.
[29] Porter, Jack (2003). *Estimation in the regression discontinuity model*, Harvard University manuscript
[30] Rivkin, Steven G., Eric A. Hanushek, and John F. Kain (2005). *Teachers, schools, and academic achievement,* Econometrica, vol. 73, No. 2, 417-458.
[31] Rockoff, Jonah E (2004). *The impact of individual teachers on student achievement: Evidence from panel data*, American Economic Review, 94, 2, 247-252.
[32] Thistlethwaite, D. and D. Campbell (1960). *Regression-Discontinuity analysis: an alternative to the ex post facto experiment,* Journal of Educational Psychology 51, 309-17.
[33] Urquiola, Miguel (2006). *Identifying class size effects in developing countries: evidence from rural Bolivia* , The Review of Economics and Statistics, 88(1), 171-177.
[34] Van der Klaauw, Wilbert (2002). *Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach,* International Economic Review, Vol43 (4), November.

Chapter 5
Conclusions

My dissertation investigates human capital issues, including education and health, in China. In the first chapter, I test for evidence of an intra-household flypaper effect by evaluating the impact of an educational fee reduction reform in rural China on different categories of household expenditures. Using data from Gansu Province in China, I find that educational fee reductions were matched by increased voluntary educational spending on the same children receiving fee reductions, providing strong evidence of an intra-household flypaper effect. The second chapter investigates the long-term effects of China’s 1959-1961 famine. Using China’s 2000 population census data, I find that women affected by the famine in the first year of life were living in smaller houses, achieved lower level of education, and provided less labor in their adulthood. But there are no long term effects on men affected by the famine in their early years of life. In the third chapter, I investigate the impact of school quality on students’ educational attainment using a regression discontinuity research design that compares students just above and below entrance examination score thresholds that strictly determine admission to the best high schools in China’s rural counties. Using data from Gansu Province in China, I find that attending the best high school in one’s county of residence decreases the probability to take college entrance examination; increases college entrance scores and the probability of entering college.