Fertility Policies and Human Capital Investment in China

by

Tianran Dai
B. S, B. A., Peking University, 2004
M. A., Peking University, 2007
M. A., Brown University, 2008

Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in the
Department of Economics at Brown University

Providence, Rhode Island
May 2015

This dissertation by Tianran Dai is accepted in its present form by
the Department of Economics as satisfying the dissertation requirement
for the degree of Doctor of Philosophy.

Date ______________ ______________________________
Andrew Foster, Director

Recommended to the Graduate Council

Date ______________ ______________________________
Vernon Henderson, Reader

Date ______________ ______________________________
Louis Putterman, Reader

Approved by the Graduate Council

Date ______________ ______________________________
Peter M. Weber
Dean of the Graduate School

Vita

The author was born in Tongling, Anhui on July 8, 1982. She graduated from No. 1 Middle School of Tongling in 2000 and then attended Peking University, graduating with a B.S. degree in Geophysics and a B.A. degree in Economics in June of 2004. After graduation she attended National School of Development (NSD) at Peking University and graduated with a Master's degree in Economics in June of 2007. Three years of training at NSD made her decide to pursue economics as her career and apply for an advanced study in the United States of America. She was admitted to the Ph.D. program in the Department of Economics at Brown University and then moved to Providence, Rhode Island. Graduate classes at Brown University inspired her to make Development Economics, Labor Economics, and Urban Economics her Ph.D. fields with a focus on Development Economics. Under the guidance of Professor Kenneth Chay, Professor Andrew Foster, Professor Vernon Henderson, and Professor Louis Putterman, the author wrote this thesis. This dissertation represents the culmination of her efforts to complete the requirements for the Degree of Doctor of Philosophy in the Department of Economics at Brown University.

Acknowledgements

I thank Andrew Foster for inspiring me to become a Development Economist, investing a tremendous amount of time in helping me to acquire structural and analytical skills, putting up with endless discussions of my model details, and providing all possible research directions when I was lost in my unclear results. I thank Vernon Henderson and Louis Putterman for helping me walk through the economic model and the institutional background of my work, and encouraging me to be as meticulous and precise as possible. I thank Kenneth Chay for guiding me through the hazy parts of my work and showing me how to take an interesting but vague idea and turn it into a well-defined research question. I also thank Sophocles Mavroeidis and Anna Aizer for giving me the opportunities to be their research assistants when I was uncertain about future research topics and used the research experience as a way to explore my economic interests. I must also thank my fellow graduate students in the basement of the Economics Department, Nathan Schiff and Ang Sun, for all our conversations that helped to shape this thesis and future work. Finally, I thank my parents for all our conversations that were derived from their own experience and knowledge as well as for their continual and consistent support.

Contents

List of Tables

List of Figures

0.1 Thesis Introduction

After the founding of the People's Republic of China, improved medical care substantially lowered the mortality rate and helped people live longer. Concerned about a resource shortage, China's central government started to implement fertility policies in the mid-1950s. In a speech introducing the policies, Mao stated that unmanaged human reproduction was equivalent to anarchy with no government, no organization, and no rules (Mao, MacFarquhar, Cheek, Wu, Goldman and Schwartz 1989, p. 159). From then on, the government aspirations to limit births were stamped on every couple's practice of family planning.

The central government's strong role in encouraging the adoption of family planning technology and generating guidelines for fertility control led to large changes in Chinese women's fertility in just 20 years. This quick and far-reaching transformation in household behavior had a significant impact on human capital formation, with repercussions for the country that continue today. On one hand, the demographic transition in China is similar to what many developing countries had experienced. It has played an important role in the economic development. On the other hand, it is unique in its impact in stimulating human capital growth in China. Here in my thesis I will demonstrate how the fertility transition affects the accumulation of human capital for both the young and the old generations. To my knowledge, this is the first study of the impact of the whole history of fertility policies, ranging from an initial experiment to the one-child policy, on the human capital formation in China.

What is even less documented in the literature is the difference in the timing of the implementation of fertility policies across provinces. The initial framework of population control in the early 50s was inspired by the Malthus theory. Although the central government failed in the first trial in certain areas, the urgency to control the growth rate of the national population in order to sustain food and resources had been agreed on among the top leaders. From 1964 onward, after the innovation of new oral contraceptive, several provinces began to spread the knowledge and technology of birth control. In the 70s, the central government promoted the program called ”later longer and fewer” among majority provinces, which was the beginning to regulate the family size using incentive and disincentive measures. Approaching the end of 1970s, the government further tightened the rules and launched the one-child policy among all provinces. I will utilize this variation in both geography and timing to study the fertility change and the education attainment.

My first paper focuses on the effect of the fertility policies on the number of children in a household and children's education outcomes driven by the change in the family size. Using the geographical variation in the timing of policy implementation, I find that the fertility policies have a great impact on the number of children of women. Fertility policies also increase the college attainment of children if they are from a household with fewer children. This paper provides empirical evidence for the quantity-quality trade-off theory.

My second paper studies the effect of the reduction in the young population resulting from fertility policies on the educational upgrade among older cohorts born before the policies. This paper is based on the observation that more adults returned to school and acquired a college education in recent history. I show that fertility polices reduce the size of the young cohort and increase their educational levels. Through imperfect substitutability of workers with the same education but different ages, the incentives for older adults to acquire more education increase. This paper shows the link between the fertility policies and the educational upgrading of the labor force and explains the reason behind the interesting phenomenon of adult continuing studies. This paper has two policy implications. First, it is important to provide educational institutions or educational channels that are adapted to meet the existing demand for skilled labor. Second, because of the complementarity and/or the substitutability between skilled labor, it is crucial to further enhance the free mobility of labor across locations, and it will benefit future economic growth.

Both of my papers are broadly related to the field of human capital investment. The first paper provides evidence from household decisions at the micro level. The second paper is about the production function at a macro level. Of course, over a longer period of time, such as twenty years that were studied here, other economic factors will change as well. The impacts from fertility policies on human capital investment should be viewed as an aggregate impact from all other possible economic channels. For example, the savings rate may increase as a consequence of fertility change, which would yield an increase in public expenditure on education and facilitate an individual's will to invest more in their education. Future studies should give more attention to the supply of education resources in response to the fertility change.

Chapter 1

Quantity-Quality Trade-off: Evidence from High School and College Enrollment of Young Adolescents in China from 1982 to 2000

1.1 Introduction

A long history of fertility policies existed in China. After the establishment of the country in 1949, the population growth rate increased tremendously with the improvement of health and medical conditions. To cope with rapid population growth and alleviate its pressure on resources, the central government launched a series of fertility policies that were intended to control population growth within a limited scope. However, while it has been more than six decades since the policies were first defined, there have been few studies looking at the impact of the fertility policies on schooling outcomes of children. Understanding this impact is crucial for policymakers who will use certain measures such as investment in family planning services to boost the economic growth. This paper tries to speak to two specific questions. First, how did fertility policies intended to control population affect womens fertility rate in China? Secondly, did the change in the fertility level subsequently affect the human capital of children?

This paper uses empirical evidence from China and speaks directly to the question of the impact of the fertility policies on the demand for children and the human capital investment in children in the household. Chinese fertility policies differ from other family planning programs in several aspects. First, local governments provided birth control methods, including free access to abortion, all at one time. Therefore, the substitution between different methods and its associated direction of the effect on the fertility rate are not a primary concern here. The influence on the fertility level of a woman is viewed as the gross effect of these methods. Second, political campaigns mobilized couples to comply with the policies. The central government combined the provision of birth control methods with the idea of having a small offspring size. Large and extensive media education existed to influence young peoples decision making. Persuasion by the community health workers and the labor union was entrenched in everyday life. Local governments usually resorted to incentives and disincentives to achieve a certain population goal. All these measures could generate large effects on the fertility level.

Previous papers have shown that the access to family planning technology increases womens schooling and the rate of labor force participation (Bailey 2006, Angeles, Guilkey and Mroz 2005). In some contexts, it is shown that family planning, such as pills and access to abortion, improves the education and labor market outcomes of children (Pop-Eleches 2006, Joshi and Schultz 2007, Roy and Foster 1996). Nevertheless, some researchers have not found any effect of family planning on human capital investment in children (Valente 2014, Mitrut and Wolff 2011).

Some papers have looked at the effect of family planning on the educational attainment of children rather than focusing on infant health (Pop-Eleches 2006, Roy and Foster 1996). It is possibly because the quantity quality trade-off might not be an important channel when children are small in age. A lack of proper longitudinal data prevents researchers from further exploring this issue. Here this paper uses census data over a long period to study the impact of family planning on college enrollment of older children who live at home. The current schooling outcomes of children are linked to the timeline of fertility policies that were implemented about two decades prior. There are several advantages of my study. First, I can observe the schooling of children, such as the college attainment, as long as they live with their parents. By selecting them at a relatively young adult age, I avoid the problem of selection bias for the main findings. Second, I take advantage of the time and geographic variation in the policy implementation to deal with the endogeneity issue of fertility and schooling. Also, the data structure allows me to control for province fixed effects that may determine the placement of the policies. Third, to ensure that the impact of the fertility policies is not confounded by other socioeconomic factors, I do a robustness test by examining the policy impact on minority Chinese.

My results show that the fertility policies in China have significant impacts on the fertility of women in the long run. The reduction in the family size due to policies increases the college enrollment by about 1.6 to 1.9 percentage points, which attributes to more than 50 percent of the increase in the college enrollment in 1990-2000. My research can be broadly linked to the investment in human capital in developing countries (Joshi and Schultz 2007). This paper highlights the importance of family planning in promoting the human capital accumulation.

1.2 Background

The initial attempt to curb the fast growing population happened around 1956, after the country had made great medical improvement and that increased quality of life. With the mortality rate decreasing and the marriage and fertility rate increasing, the central government recognized the need to control population growth and initiated a new policy. This first policy failed to notably change the fertility rate, possibly due to disagreement between top leaders about whether reducing population growth can lead to better economic growth

(Greenhalgh 1994, Chen and Kols 1982, Peng 1997, Jiang 1996). The second campaign to control the population started from 1963 or 1964. Provinces, including Beijing, Tianjin, Shanghai and Liaoning, initiated the research of birth control methods and exerted efforts to distribute many kinds of birth control equipment to the couples. Family planning committees were established in these provinces. The committee served as the earliest organization responsible for the family planning activities in each province and used personnel to educate and persuade couples to have fewer children. Some incentives were also created during this period. For example, provinces provided free birth control methods and surgeries, extended maternal leave, food rations, and/or wage bonuses (Peng 1997, Zhang 2002) to couples who planned to have fewer children. However, the influence of this campaign was limited and the campaign stopped a couple of years later. It was only after 1970 that a large promotion for slow population growth was launched by the central government.

From 1971 to 1975, all provinces except some remote provinces launched the later longer and fewer fertility program (Peng 1997). Each province set up their own organizations and facilities and specified fertility rules following the general instruction of the nation. Basically, the national guideline imposed a minimum age for marriage, which was older than the social customs of the time. It also suggested a longer birth spacing of three to four years for childbearing and encouraged couples to have fewer children, usually two for urbanites and three for rural residents (Peng 1997, Zhang 2002). Each province set their own agenda for implementing these policies with some provinces enforcing these rules earlier than others. For example, Beijing, Tianjin, Shanghai, Liaoning, Fujian, Gansu, etc., started to implement the policy around 1971-1973, while Anhui, Henan, Hubei, Jiangxi, Jilin, Sichuan and Xinjiang started during 1974-1975 (Peng 1997). To enforce the policy, provinces resorted to more media channels, both formal and informal ways for purposes of education and persuasion. They provided free birth control methods, increased the registration difficulties for marriage and delivery, and used incentives such as cash subsidies, food subsidies, or better housing (Chen and Kols 1982, Peng 1997). The types of enforcement and its intensity tended to vary across provinces[1]. A number of reasons, such as the economic environment, social norms, or political power of the government may explain the differential enforcement (Attane 2002). To some extent, the Chinese government viewed the fertility policy as a strategic action to control the production of the countrys population, so the strength of government may play an important role in enforcement. However, this kind of strength depended on a few top leaders in the central government and the provincial government. Based on evidence from media and interviews (Chen and Kols 1982), I argue that the strong or weak leadership in the provincial government will greatly determine whether the province implemented the policy earlier or later.

During the 1970s, the later longer and fewer program reduced the fertility rate considerably (Coale 1984, Chen and Kols 1982, Yang and Chen 2004). Nevertheless, the result from the program was not sufficient to achieve the governments fertility policy goal. The central government tightened its rules and launched the one-child policy in 1979 (Greenhalgh 1994, Peng 1997). The central government suggested couples have only one child. In some rural areas where there were strong preferences for boys, this quota was relaxed to allow two children given the first child was a girl (Short and Zhai 1998, Chen and Kols 1982). In addition to the methods and incentives that had been used in the later longer and fewer program, the one-child policy used

[1]The variation might also exist at the community level, but it could be endogenous to the fertility choice of local people (Gu, Wang, Guo and Zhang 2007). I only investigate the total provincial difference in whether a policy was implemented or not regardless of the rural-urban differential within a province.

some new incentives and disincentives. It imposed a large amount of fines on couples that did not comply with the policy. Extra children were not qualified for health benefits and education subsidies. Couples with only children received rewards from the government including cash subsidies, better pensions, better housing, and/or job promotion opportunities (Short and Zhai 1998). Similar to the later longer and fewer program, the amount of incentives and disincentives tended to vary across provinces and within province as well (Attane 2002, Gu et al. 2007, Short and Zhai 1998). The different feature is that the one-child policy was implemented at the same time by all provinces, with different rules of exemption constructed by individual province (Gu et al. 2007, Peng 1997). The provinces that joined the national one-child policy from 1979 include Guangxi, Yunnan, Qinghai, Ningxia and Guizhou. This group of provinces carried out family planning work gradually from 1980s, and most of them have relaxing policies for minority Chinese. The policy rules were also less strict for Han Chinese than the rest of the country (Peng 1997).

1.3 Theoretical Framework

In this framework, I use the Becker model of the quantity-quality trade-off (Becker 1981, Becker and Lewis 1973, Becker and Tomes 1976) to explain the relationship between the number of children and the schooling of children. An important feature of their model is that the price of schooling is proportional to the number of children in the household. Holding everything else fixed, the demand for schooling will be lower when parents have more children. The key underlying assumptions are that children are the same and parents cannot borrow. The model also predicts parents will choose the number of children and the childrens schooling based on the parents taste for schooling. Parents who have a strong taste for schooling will choose high levels of schooling and fewer children.

A specific feature of my model is the role of the fertility policies implemented by the central government. I allow the birth control technology to affect the number of children as an exogenous factor in the model. The adoption of birth control technology allows couples to be better at ensuring the desired number of children. More importantly, couples now face a cost of having an extra child under the fertility polices of later longer fewer and the one-child policy. The cost is associated with economic rewards or punishment implemented by the central government. Both the access to birth control technology and the implementation of fertility policies are independent of the parent taste for schooling and the endowments of children.

Another feature of the model is that I divide the children at home into two groups, older children and young children based on age. I separate children by age for two reasons. First, I examine the educational outcomes of children in a household at a high school level or a college level and only children above an age limit can possess such a level. I allow the effect of the number of older children on these children's education to be different from the effect of the number of young children. Additionally, while young children were likely to be born after fertility policies were in place, older children could be born before or after policies depending on which provinces they were born[2].

[2]Although children may migrate with their mother to a different province than where they were born, my model does not take this into account. In the other paper of my dissertation, I show that the pattern of migration of adults above age 30 is less likely to be correlated with the timing of fertility polices.

The model depends on whether the schooling decision of young children is taken into account at the moment when older children begin to enroll in high school or college. To the extent that parents are myopic and do not care about the education of young children, the impact of the number of young children on the schooling of older children will be zero[3].

Suppose that the utility function is $ln(S_0)+a_1ln(n_0)+a_2ln(n_y)+a_3ln(X)$, where the variable S_0 is the average schooling of children, the variables n_0 and n_y indicate the number of older and young children in a household, and the variable X is the consumption level of adults. Parameters a_1, a_2, and a_3 are all positive. The average schooling cost is a convex function of the average schooling τS_0^2, where τ is positive. The budget constraint is $P_X X+(n_0+n_y)\tau S_0^2+n_0\theta_F+n_y\theta_G=y$, where y is household income, the parameter P_X is the price of consumption, and the parameters θ_F and θ_G represent the costs of older children and young children separately. The two parameters may be different because older children and young children may be born when different fertility policies were in place.

Maximize the utility function subject to the budget constraint. It yields:

$$
\begin{aligned}
S_0 &= y^{1/2}(2\tau\alpha_1+2\tau\alpha_3)^{-1/2}(n_0+n_y)^{-1/2} \\
n_0 &= n_0(\theta_F,\theta_G,y) \\
n_y &= n_y(\theta_F,\theta_G,y)
\end{aligned}
$$

The model predicts that the impact of n_0 and n_y on the schooling of older children S_0 is negative. The policies when the older children were born and when the young children were born will jointly determine n_0 and n_y. If parents do not care about the education of young children, the budget constraint is $P_X X+n_0\tau S_0^2+n_0\theta_G=y$. It yields:

$$
\begin{aligned}
S_0 &= y^{1/2}(2\tau\alpha_1+2\tau\alpha_2+2\tau\alpha_3)^{-1/2}n_0^{-1/2} \\
n_0 &= n_0(\theta_F,y) \\
n_y &= n_y(\theta_G,y)
\end{aligned}
$$

The prediction is the impact of n_0 on S_0 is negative and the impact of n_y on S_0 is zero. The number of older and young children n_0 and n_y only depends of the policies specific to them.

Here I use a linear logarithmic utility function. It is difficult to decide whether the effect of n_0 and S_0 is positive and negative without imposing such an assumption. When the price of n_0 increases, for example, when fertility policies were implemented, the substitution effect is positive which induces more schooling expenditure for older children. On the other hand, the negative income effect when the price of n_0 increases induces less schooling expenditure. Moreover, if n_0 decreases when the price increases, due to the quantity-quality trade-off mechanism, the price of schooling expenditure will decrease and parents will spend more on the older children due to a positive substitution effect and a positive income effect. In the end, the movement of the schooling expenditure on older children hinges upon the magnitudes of the substitution effect, the income effect when the price of n_0 increases, and the effect of quantity-quality trade-off when n_0 decreases.

[3] An alternative hypothesis is that parents only invest in the education of older children. Older children subsequently invest in the education of young siblings.

A key component that demonstrates the quantity and quality trade-off is the price of children. When the price of children increases due to the disincentives caused by fertility policies, parents choose to have more education per child rather than more children. It is worth mentioning that fertility policies may also affect the schooling level of each child through the aggregate demand for schooling. The price of schooling indicated by τ can reflect this aggregate impact in the schooling decision within a household. For example, given the same supply of college seats and an upward sloping supply curve, when the aggregate number of students who take the entrance exam increases (decreases)[4], each child will have a lower (higher) chance of being admitted to college, i.e., the schooling price τ becomes greater (smaller) for each household. In sum, when the aggregate demand for schooling increases, the average schooling of children within a household may decrease if the supply curve of college seats has an upward slope. Since my identification strategy uses the timing of fertility policies implemented in a province or an area, the coefficient represents the combined effect of the quantity-quality trade-off within a household and the aggregate demand for schooling within a province or an area. However I will provide some evidence that shows how the aggregate demand for high school or college education changes with respect to the timing of fertility policies.

An alternative theory for how the fertility polices affect the education of children is a change in the bargaining power of women associated with the usage of birth control technology. Fertility policies may give women more weight in deciding the number of children in the household (e.g, women can always abort the unwanted children). Through the bargaining power, women will have more allocated resources in the family, which leads to more expenditure on children and higher education.

1.4 Data

The data I use are China census samples in 1982, 1990 and 2000. To look at the policy effect on the fertility rate of women, I need the sample of all women in their reproductive age. Censuses provide information on the fertility history of women if they are between 15 and 50 years old. I restrict the sample to women aged 40-49 who are Han or Hui. Most of the ethnic minority groups concentrate in a few provinces, which does not allow me to do the comparison across groups of provinces. Hui is the only minority group that is widely distributed in China. Hui comprises about 12 percent of the total population of minorities, and less then 1.2 percent of the national population.

I further look at the effect of the number of siblings on the education of children in the household. The number of sibling is captured by two variables: the number of children aged 16 or above and the number of children aged 15 or below. I select the sample of children living in the parent household and aged between 16 and 21, which is the relevant age to acquire a high school degree or enroll in a college. There are some data limitations that make it hard to study every child ever born to a woman. First, the fertility history includes the number of living children a woman has but it does not have the information on the education of each child. Second, the only way that the survey identifies a woman and her child in the sample is through the relationship in the household. So I select my sample based on the following criteria: (1) there is only one

[4] Although the applicants taking the entrance exam to college are not necessarily of the same age, they might be born in fairly close birth cohorts. Therefore, fertility policies can directly affect the aggregate amount of students applying for higher education.

woman living in the household who is Han or Hui and between 40 and 49 years old; (2) there is at least one child living in the household who is related to the woman as in the mother-child relationship; (3) the number of living children the woman has given birth to is greater than the number of children who are living in the household in that year. This is to ensure that neither of the independent variables is negative, since the number of children aged above 16 is calculated as the number of living children minus the number of children aged below 15 living in the household.

Household characteristics are controlled in both the fertility equation and the equation for childrens education. The census data has information on mothers characteristics such as her age, birth year, education, employment status, marital status, and whether the husband is in the household.

In addition, I control for the province-year level macro economic variables, which can contribute to the changes in the fertility and the educational attainment of children. They are the real GDP per capita, the real fixed asset investment per capita, population, population density, and the number of tertiary teachers per capital.

Table 1.1 shows the sample statistics of my data. Table 1.10 shows that provinces in different groups are not statistically different from each other in terms of the initial states of GDP level, density, schooling resources, etc, except those in the first region. The first group of provinces includes three municipalities that are directly under the administration of the central government and therefore may differ from other provinces. To the extent the policies were generated by these macroeconomic conditions, the relationship of the number and the quality of children can be biased when I include the first group in the sample. I will use the specification with and without the first group of provinces as a robustness test later.

1.5 Empirical Models

The effect of the fertility policies on the number of children, either young or older defined by the age of children, can be characterized by the following equation:

$$n_{jpt} = \theta_F + \alpha X_j + \rho W_{pt} + \varepsilon_{jpt} \tag{1.5.0.1}$$

where n_{jpt} is the number of children, either young or older, that a woman j from province p has in year t. The vector X_j represents the characteristics of the woman including her education levels, her employment status and her marital status. It also includes the dummy variables for age of the woman and birth cohort of the woman. The variable W_{pt} is the macro level control variables that vary at province and year: the logarithm value of the real GDP per capita, the logarithm value of the real fixed assets per capita, population, population density, and the number of tertiary teachers per capita. The error term ε_{jpt} is the unobservable term that determines the number of older or young children a specific woman has.

The effect of the fertility policies on the number of children is the main interest, which captures the effect of the initiation of fertility policies in the period from 1963 to 1979. Since the policies are phased into regions over a relatively long period, I break down the effect into two periods interacted with four regions.

Table 1.1: Sample Statistics

	1982	1990	2000
Women of age 40-49			
Having at lease one child aged 16-21	0.81 [0.39]	0.76 [0.43]	0.5 [0.5]
Children aged 16-21 of Women 40-49			
Educ HS+	0.22 [0.41]	0.15 [0.36]	0.31 [0.46]
region-I	0.38	0.29	0.57
region-II	0.22	0.16	0.31
region-III	0.19	0.13	0.28
region-IV	0.17	0.12	0.22
Educ College+	0.009 [0.097]	0.01 [0.1]	0.03 [0.17]
region-I	0.016	0.02	0.084
region-II	0.008	0.01	0.027
region-III	0.009	0.009	0.028
region-IV	0.013	0.012	0.009
# of older siblings 16+	2.92 [1.25]	2.94 [1.22]	1.98 [0.93]
region-I	3.01	2.49	1.35
region-II	3.01	2.9	1.99
region-III	2.82	3.03	1.99
region-IV	2.73	3.12	2.47
# of young siblings 0-15	1.89 [1.24]	0.77 [0.89]	0.46 [0.74]
region-I	1.06	0.45	0.17
region-II	1.72	0.7	0.47
region-III	2.17	0.81	0.47
region-IV	2.57	1.41	0.64

I select the sample of households with at least one child aged 16-21 living at home for the regressions of women's fertility level and children's education outcomes. I list the national average of the education of children aged 16-21, which is whether they enroll in school or graduate at the level of high school and college. I also list the average education of children by region and the average number of older and young siblings of children aged 16-21.

The following equations are used to estimate the impact of the policies:

$$n_{jpt} = \sum_r \sigma_r year_t \times region_r + aX_j + \rho W_{pt} + \varepsilon_{jpt} \quad (1.5.0.2)$$

$$n_{jpt} = \sum_r \sigma_r post_t \times region_r + aX_j + \rho W_{pt} + \varepsilon_{jpt} \quad (1.5.0.3)$$

In the above equations $year_t$ is the year dummy indicating the second year in each period, and $region_r$ is the policy region dummy. The estimate of σ_r from equation 1.5.0.2 is the regional trend of the number of children net of family characteristics and the macro level covariates. I also add the equation 1.5.0.3 as a robustness check. I use the post-policy time dummy interacted with the regional dummies to indicate the policy impact. The interpretation of these interactions is the effect of the initiation of a policy and/or the effect of a later compared to an early implementation of the policy. In the robustness check, I also drop the first region that is statistically different than other regions in the initial state of each of the two periods.

It is worth noting that the estimates of σ_r or τ_r may capture other aggregate impacts that coincide with the fertility policies, for example, the changes in the infant mortality rate across regions, or maternal health improvements across regions. Unfortunately, I do not have direct measures for the changes, but by controlling some macro variables in W_{pt} I will reduce the part in the estimates of σ_r or τ_r that is due to other socioeconomic changes. Since I am not aware of any event that occurs differentially across regions, I argue that changes in these socioeconomic conditions do not differ significantly across the regions. Therefore it is less of a concern that the policy impact that I estimate will be confounded by other aggregate impacts. To justify my argument, I also examine the fertility change for the minority women as a robustness test. Assuming that minorities do not comply with any fertility policies, the estimates of σ_r or τ_r should be zero in equations 1.5.0.2 and 1.5.0.3 if they do not capture other aggregate impacts.

After examining the impact of the fertility policies on the number of children, I turn to examining the quantity-quality trade-off of children within a household. I mainly focus on the children aged 16 to 21 because their education levels vary substantially above middle school for their age. The education level of a child is determined by the following equation:

$$Edu_{ijpt} = \beta_1 NOLDER_{jp,t-16} + \beta_2 NYOUNG_{jpt} + \gamma X_j + \eta W_{pt} + \omega_{ijpt} \quad (1.5.0.4)$$

where Edu_{ijpt} is the education level of child i in household j from province p in year t. It is defined as whether they have a high school degree and whether they enroll in a college or attain a college degree. It depends on the number of children aged 16 and above in the household $NOLDER_{jp,t-16}$, the number of children aged 15 or below $NYOUNG_{jpt}$, mother's characteristics X_j, and macro level variables W_{pt}. The unobserved term, ω_{ijpt}, can be the endowment of the child, the taste for schooling or factors that contribute to the education at the provincial and year level. To be more precise, $NOLDER_{jp.t-16}$ is calculated as the total number of children a woman has given birth to minus the number of children aged 15 or below living in the household. For the schooling choice, the variables $NOLDER_{jp,t-16}$ and $NYOUNG_{jpt}$ is affected by fertility policies, mother's characteristics, macro variables, and other unobserved variables as specified in equations 1.5.0.2 and 1.5.0.3[5].

[5]The number of older children aged 16 or above observed in year t is determined by the fertility polices in year $t-16$. This is why one of the subscript indexes of $NOLDER_{jp,t-16}$ is $t-16$. Likewise, the number of young children aged 15 or below is determined by fertility policies from year $t-15$ to year t.

One concern about the policy effect is that some provinces may be more likely to implement the fertility policies than others because of some time invariant components. For example, densely populated provinces tend to have stricter fertility policies, or provinces with better health infrastructure are easier to carry out family planning activities. Therefore, in the regression I control for the province fixed effects to not only deal with the fixed component of the policy placement but also the fixed preferences for the number of children and education investment as well. Additionally, the province fixed effects can soak up some differences in economic growth across regions.

I use both OLS and 2SLS methods to estimate β_1 and β_2: the effects of the number of older and young children on the education of the 16-21 year olds at the margin of high school and college education. The estimate of β_1 and β_2 from OLS will be inconsistent if the variables $NOLDER_{jp,t-16}$ and $NYOUNG_{jpt}$ are correlated with the error term ω_{ijpt} in 1.5.0.4, which is likely to be the case because ω_{ijpt} contains the unobserved taste for schooling. To get the consistent estimators of β's, I use the 2SLS method to estimate equation 1.5.0.4 with the first stage from 1.5.0.2 or 1.5.0.3. I provide results of GMM estimators with standard errors clustered at the level of province interacted with year. The educational investment of children in the same province and the same year could be correlated due to some unobserved characteristics. However, the regression results do not change much if I use the clusters at the regional year level.

The exclusion restriction for the instruments relies on the correlation between the timing of the implementation of fertility policies $post_{t-16} \times region_r$ and the unobserved component of schooling ω_{ijpt}. The timing of the policy implementation should be uncorrelated with individual child and household characteristics *a priori*. However, it might be correlated with the provincial-year level shocks to schooling, which yields inconsistent estimate of β's. Here I argue that the decision to implement fertility policies within a province is pre-determined by a provincial leader based on her belief of the importance of such policies, so it is independent of the schooling shocks that occur 16 years later. I provide some evidence for this argument from the sample of children of minority women.

1.6 Results

1.6.1 Results on fertility of women

The first step is to establish the causal relationship between the fertility policies and the number of children in the household. Table 1.2 shows the first stage results, i.e., the estimated policy effect on the fertility level of all Han women aged 40-49, which is measured by the total number of living children, separated by age 16[6]. The result is allowed to be different for the period 1982-1990 and 1990-2000.

Column (1) of Table 1.2 shows that in the first period, the number of older children(aged 16 and above) born to a Han woman decreased in the provinces of the first two groups, including Beijing, Fujian, etc., relative to the baseline group by about 0.3 to 0.4. However, the number of older children was not different in the third group relative to the baseline group. This is consistent with the timing of the policies. Women in provinces of the first two groups gained access to birth control methods and were exposed to the LLF policy

[6]The sample is the universe of households with at lease one child of age between 16 and 21.

Table 1.2: The effect of policies on # of older/younger children: Han Ethnicity

VARIABLES	[1] n_O 1982-1990	[2] n_y 1982-1990	[3] n_O 1990-2000	[4] n_y 1990-2000
BeijingXyear	-0.412***	*0.226***	-0.519***	*0.489****
	(0.082)	*(0.101)*	(0.107)	*(0.079)*
FujianXyear	-0.293***	*0.041*	-0.472***	*0.604****
	(0.059)	*(0.071)*	(0.069)	*(0.071)*
AnhuiXyear	-0.103	*-0.204****	-0.569***	*0.421****
	(0.065)	*(0.070)*	(0.094)	*(0.066)*
Mom variables	Y	Y	Y	Y
Macro variables	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Prov FE	Y	Y	Y	Y
Children Age FE	Y	Y	Y	Y
Mom Age FE	Y	Y	Y	Y
Mom Birth year FE	Y	Y	n/a	n/a
Observations	1,068,495	1,068,495	613,177	613,177
R-squared	0.328	0.356	0.411	0.184

The variables BeijingXyear, FujianXyear and AnhuiXyear are the interaction of a dummy indicating a region and a dummy indicating the end year of each period. Beijing indicates the first region that includes Beijing, Tianjin, Shanghai and Liaoning. Fujian indicates the second region that includes Fujian, Gansu, etc. Anhui indicates the third region that includes Anhui, Henan, etc. The omitted region, which is also the baseline region, includes Guangxi, Yunnan, etc. See Figure 2.2 in Chapter 2 for the location of these regions.

during this period. Meanwhile, the average number of older children decreased in the provinces of the second group by 0.5 relative to the baseline group[7]. Therefore, about 60 percent of the drop in the fertility level may be due to the fertility policies.

Column (2) of Table 1.2 indicates the policy effect on the number of young children aged 15 or below in the period 1982-1990. To interpret the differential effect for provinces, we have to back out the years when the young children were born, which is 1967-1982 and 1975-1990 for the respective year. The LLF policy and the one-child policy, combined with access to birth control methods had been in place in every province by the end of each period. The policy effect on the number of young children should be intuitively the same across provinces, but it may differ because the policy diffusion may take different time in different provinces. More importantly, the number of older children would affect the number of young children that were born later, and the effect through this channel shows in the differential policy effect across provinces.

Now we look at the policy effect in the second period 1990-2000 in column (3) of Table 1.2. The comparison is of the difference in fertility policies till 1974 and 1984 across provinces. From 1974 to 1984 the LLF policy was extended further to provinces of the third group, and all provinces implemented the one-child policy with the baseline provinces having a more relaxed policy. The result shows that the decrease in the number of older children is greatest in the third group, but the difference between the first, second and third group is not as large as in the previous period 1982-1990, possibly because the one-child policy effect dominates in the second period and all of them have a large decline in fertility from the one-child policy. The difference between these groups and the baseline group is about a 0.5-0.6 reduction in the number of older children because the baseline provinces have a comparatively relaxed policy. In fact, the observed average number of older children decreased by 0.4-0.5 in these groups relative to the baseline group[8], so the entire drop of the fertility may be due to a relaxed policy in the baseline group of provinces.

Column (4) of Table 1.2 indicates the policy effect on the number of children aged 15 or below in 1990-2000. We compare the fertility policies during two periods: 1975-1990 and 1985-2000. The change in the number of young children in the baseline provinces is smaller than the change in other provinces by 0.4 to 0.6. It is consistent with the fact the largest reduction in the number of young children is due to the diffusion of the birth control methods and the one-child policy in the baseline provinces. The baseline provinces, with the largest family size in the country, implement the policy at the latest time and experience the largest drop in the number of young children in the period 1990-2000. It might also be true that the number of older children plays a role in determining the number of young children at home.

The coefficients on other variables in the regression of Table 1.2, not shown here, indicate that mother with higher education have lower fertility. The effect of maternal education is larger at the margin of high school education than college education. Married women tend to have fewer older children and more young children. Being unemployed is associated with more older children in the first period and fewer young children in the second period. Other macro level variables tend to not have a consistent relationship with the

[7] See Table 1.1. From 1982 to 1990, the average number of older children increased by 0.39 in the fourth region and decreased by 0.11 in the second region. Therefore, the relative change in the number of older children was a drop of 0.5 in the second region compared to the baseline region (i.e., the fourth region).

[8] See Table 1.1. From 1990 to 2000, the average number of older children decreased by 1.14 in the first region, 0.91 in the second region, 1.04 in the third region and 0.65 in the fourth region. The relative change in the number of older children was a drop of 0.3-0.5 in the three regions compared to the baseline region.

Table 1.3: The effect of policies on # of older/younger children: Ethnic Minority Hui

VARIABLES	[1] n_O 1982-1990	[2] n_y 1982-1990	[3] n_O 1990-2000	[4] n_y 1990-2000
BeijingXyear				
FujianXyear	0.039	-0.152	-0.162	0.910***
	(0.079)	(0.136)	(0.127)	(0.099)
AnhuiXyear	0.195*	-0.384**	-0.193	0.842***
	(0.113)	(0.144)	(0.279)	(0.180)
Mom variables	Y	Y	Y	Y
Macro variables	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Prov FE	Y	Y	Y	Y
Children Age FE	Y	Y	Y	Y
Mom Age FE	Y	Y	Y	Y
Mom Birth year FE	Y	Y	n/a	n/a
Observations	7,901	7,901	4,938	4,938
R-squared	0.286	0.373	0.363	0.274

The variables BeijingXyear, FujianXyear and AnhuiXyear are the interaction of a dummy indicating a region and a dummy indicating the end year of each period. Beijing indicates the first region that includes Beijing, Tianjin, Shanghai and Liaoning. Fujian indicates the second region that includes Fujian, Gansu, etc. Anhui indicates the third region that includes Anhui, Henan, etc. The omitted region, which is also the baseline region, includes Guangxi, Yunnan, etc. I drop Hui in the first region because they are very different from Hui in other regions.

fertility over two periods.

To confirm that the policy effects are not confounded by other aggregate impacts, Table 1.3 presents the policy effect for Hui minority women. Most of the coefficients show that the fertility policies did not affect the fertility level differentially across the four groups of provinces. The exception is the fourth column where the dependent variable is the number of young children aged 15 or below. I compare children that were born during 1975-1990 versus 1985-2000, so the result may be driven by the provincial difference in the fertility policies pertaining to Hui minority.

I also examine whether the results are robust to different specifications. I use the post-policy time dummy interacted with the dummy for a group of provinces. This combines the first and the second group into the same group in the second period. In another robustness check, I drop the provinces in the first group that is statistically different than others as shown in Table 1.10. The results lead to the same conclusions as in Table 1.2 and Table 1.3.

1.6.2 Results on education of children

In this part I look at how the number of children, both young and old, affects the education of children of age 16 to 21 living at home. The dependent variables are whether they have a high school education and whether they enroll in a college.

Table 1.4 shows the effect of the number of children on the education for the first period 1982-1990, while Table 1.5 shows the result for the second period 1990-2000. The top part of Table 1.4 lists the 2SLS estimation with errors clustered at the province and year level. I find the effect of the number of children, both young and old, on high school graduation and the enrollment in college is insignificantly different from zero in the first period 1982-1990. The result remains the same for both specifications in column (1) and column (2) and when I drop the provinces in the first group. The bottom of Table 1.4 lists the OLS results. The OLS results overestimate the effect of the number of children, suggesting parents with a taste for better schooling would choose a smaller family size.

The results for 1990-2000 (see Table 1.5) show that one additional old child decreases the chance of entering college by 0.03-0.04 percentage points. However, one additional old child does not affect the chance of graduating from high school. The effect of the number of young children remains insignificant from zero for education. The magnitude implies more than 50 percent of the increase in the share of college students in the second period can be attributed to the decrease in the number of older children in the household[9].

Other variables that significantly contribute to the education levels of children include mother's education levels. When examining high school education, the real GDP level per capita and the real fixed asset investment per capita are positive and significant in two periods. At the level of college education, only the population density seems to be positive and significant in both periods.

In Table 1.6, as a robustness check, I examine whether the fertility policies have an impact on the education of children for the Hui sample. Recalling the results in Table 1.3, I found an insignificant result of the fertility policies on the number of older children in both periods, and a significant result of the fertility policies on the number of young children in the second period. The results of Table 1.6 show that the reduced form effect of the policy on education is not significantly different from zero. Combining the results from Table 1.3 and Table 1.6, I conclude that the instrumental variables capture the fertility policies for Han Chinese women and are unlikely to represent other channels that might affect education.

In this part I consider the potential selection bias in the estimation of the effect of the number of children on education. If the fertility policies yield more education investment in children, and children with higher education are more likely to live outside parents' household, then I may under-estimate the effect of the number of children. In Table 1.7, I examine the likelihood of living in a dorm for children with high school education or who enroll in a college. The estimates show the region-I had a different likelihood for children to live in a dorm relative to other regions for both period and both education levels. Since the difference is induced by fertility policies, this means including region-I in the sample may cause a potential bias to the effect on education. The coefficients in column (1) show that the probability of living in a dorm was

[9]The difference between region-II, region-III and the baseline region in the increase in college enrollment rate is about 0.024-0.029 over 1990-2000. The reduced form estimation of the impact of fertility policies on college enrollment rate shows the increase in the college rate is about 0.016-0.019 higher in region-II and region-III relative to the baseline region.

Table 1.4: The effect of # of older/younger children on high school and college education: 1982-1990

2SLS Results

VARIABLES	[1] **hs**	[2] **hs**	[3] **hs** No region-I	[4] **coll**	[5] **coll**	[6] **coll** No region-I
N_Older kids	0.093	0.013	0.013	-0.005	-0.006	-0.008
	(0.096)	(0.088)	(0.092)	(0.010)	(0.011)	(0.012)
N_Young kids	0.056	0.026	0.012	-0.003	-0.009	-0.008
	(0.087)	(0.077)	(0.076)	(0.009)	(0.009)	(0.009)
M_hs	0.590***	0.526***	0.539***	0.059***	0.059***	0.056***
	(0.108)	(0.095)	(0.101)	(0.012)	(0.012)	(0.014)
M_co	0.234***	0.186***	0.214***	0.087***	0.087***	0.086***
	(0.050)	(0.047)	(0.055)	(0.008)	(0.008)	(0.010)
Marrital_woman	0.132***	0.117***	0.125***	0.007***	0.007***	0.007***
	(0.017)	(0.017)	(0.018)	(0.002)	(0.002)	(0.002)
Presence of husband	-0.107***	-0.087***	-0.093***	-0.003	-0.003	-0.002
	(0.031)	(0.027)	(0.028)	(0.003)	(0.003)	(0.003)
Unemp_woman	-0.057**	-0.033	-0.025	-0.001	-0.001	-0.000
	(0.028)	(0.025)	(0.023)	(0.003)	(0.003)	(0.003)
Year FE	-0.031	-0.069	-0.092	-0.005	-0.017*	-0.014
	(0.090)	(0.095)	(0.098)	(0.009)	(0.009)	(0.010)
Macro	N	Y	Y	N	Y	Y
Observations	1,068,495	1,068,495	996,652	1,068,495	1,068,495	996,652
R-squared	-0.067	0.047	0.057	0.024	0.023	0.020

Control for Prov FE, children age FE, Mom age FE, and Mom birth year.

OLS

	[1]	[2]	[3]	[4]	[5]	[6]
N_Older kids	-0.040***	-0.040***	-0.038***	-0.003***	-0.003***	-0.003***
	(0.003)	(0.003)	(0.003)	(0.000)	(0.000)	(0.000)
N_Young kids	-0.059***	-0.059***	-0.057***	-0.004***	-0.004***	-0.004***
	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)

The endogenous variables are N_Older kids and N_Young kids: the number of older children and young children in a household. I control for whether the mother's education is above the high school level or the college level. I also control for whether the mother is married, whether the husband is living in the household, and whether the mother is employed. The variable year FE is an indicator for year 1990. In addition, I control for the socioeconomic variables at the province and year level. The instrumental variables for N_Older kids and N_Young kids are region dummies interacted with year. The first stage is listed in column (1) and (2) of Table 1.2. Here in column (3) and (6) I exclude the first region as a robustness check.

Table 1.5: The effect of # of older/younger children on high school and college education: 1990-2000

VARIABLES	[1] **hs**	[2] **hs**	[3] **hs** No region-I	[4] **coll**	[5] **coll**	[6] **coll** No region-I
N_Older kids	-0.033	-0.084*	-0.044	-0.042**	-0.044***	-0.031**
	(0.054)	(0.051)	(0.043)	(0.020)	(0.015)	(0.013)
N_Young kids	0.084	-0.020	0.032	0.014	-0.012	0.004
	(0.051)	(0.069)	(0.070)	(0.014)	(0.011)	(0.009)
M_hs	0.478***	0.413***	0.462***	0.052***	0.042***	0.056***
	(0.044)	(0.044)	(0.043)	(0.016)	(0.012)	(0.011)
M_co	0.248***	0.192***	0.251***	0.110***	0.103***	0.115***
	(0.033)	(0.034)	(0.031)	(0.019)	(0.018)	(0.020)
Marrital_woman	0.079***	0.070***	0.081***	-0.006	-0.004	-0.003
	(0.013)	(0.012)	(0.012)	(0.004)	(0.004)	(0.003)
Presence of husband	-0.054***	-0.039***	-0.049***	0.005	0.006	0.003
	(0.013)	(0.012)	(0.011)	(0.004)	(0.004)	(0.003)
Unemp_woman	0.100***	0.094***	0.095***	0.013***	0.011***	0.012***
	(0.017)	(0.016)	(0.018)	(0.002)	(0.002)	(0.002)
Year FE	0.118*	-0.080	-0.021	-0.021	-0.055**	-0.025
	(0.068)	(0.100)	(0.082)	(0.022)	(0.022)	(0.016)
Macro	N	Y	Y	N	Y	Y
Observations	613,177	613,177	575,141	613,177	613,177	575,141
R-squared	0.098	0.166	0.137	-0.020	0.001	0.014

Control for Prov FE, children age FE, and Mom age FE.

OLS						
N_Older kids	-0.080***	-0.080***	-0.076***	-0.011***	-0.011***	-0.010***
	(0.006)	(0.006)	(0.006)	(0.001)	(0.001)	(0.001)
N_Young kids	-0.067***	-0.068***	-0.065***	-0.006***	-0.006***	-0.006***
	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)

The endogenous variables are N_Older kids and N_Young kids: the number of older children and young children in a household. I control for whether the mother's education is above the high school level or the college level. I also control for whether the mother is married, whether the husband is living in the household, and whether the mother is employed. The variable year FE is an indicator for year 2000. In addition, I control for the socioeconomic variables at the province and year level. The instrumental variables for N_Older kids and N_Young kids are region dummies interacted with year. The first stage is listed in column (3) and (4) of Table 1.2. Here in column (3) and (6) I exclude the first region as a robustness check.

Table 1.6: The effect of policies on high school and college education: Ethnic Minority Hui, no region-I

VARIABLES	hs 1982-1990	coll 1982-1990	hs 1990-2000	coll 1990-2000
BeijingXyear				
FujianXyear	-0.053**	-0.015***	0.016	-0.006
	(0.020)	(0.006)	(0.032)	(0.014)
AnhuiXyear	0.021	-0.005	0.021	0.037
	(0.035)	(0.006)	(0.085)	(0.032)
Mom variables	Y	Y	Y	Y
Macro variables	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Prov FE	Y	Y	Y	Y
Children Age FE	Y	Y	Y	Y
Mom Age FE	Y	Y	Y	Y
Mom Birth year FE	Y	Y	n/a	n/a
Observations	7,901	7,901	4,938	4,938
R-squared	0.167	0.032	0.269	0.113

The variables BeijingXyear, FujianXyear and AnhuiXyear are the interaction of a dummy indicating a region and a dummy indicating the end year of each period. Beijing indicates the first region that includes Beijing, Tianjin, Shanghai and Liaoning. Fujian indicates the second region that includes Fujian, Gansu, etc. Anhui indicates the third region that includes Anhui, Henan, etc. The omitted region, which is also the baseline region, includes Guangxi, Yunnan, etc. I drop Hui in the first region because they are very different from Hui in other regions.

differentially related to fertility policies during 1982-1990. Fertility policies induced children with high school education to be more likely to live in a dorm in provinces of group-II relative to the baseline group, which leads to an underestimate of the effect of fertility policies on high school education and therefore the effect of the number of children on high school education in 1982-1990.

Next, in Table 1.8, I check whether the effect of the number of children, both old and young, has different impacts on the education achievement for girls and boys aged 16 to 21. I find that the impact of the number of older children on the schooling level, if any, is greater for girls than for boys. This could be because the schooling level of girls is lower than that of boys when the family size is large. When the family size reduces, the increase in the schooling level is greater for girls. This is also consistent with the aforementioned hypothesis that an increase in women's bargaining power resulting from the birth control technology will increase the human capital investment in children, especially in female children.

As mentioned in the theoretical framework, the identified effect of the number of older children on the average schooling of children involves two separate mechanisms. The first one is the quantity-quality trade-off, which operates through the price of children. The second one is a change in the aggregate demand for schooling resulting from a small birth cohort due to fertility policies; this second effect can affect the general price of schooling for each household. In this part, I examine how the aggregate demand for schooling changes with respect to the implementation of fertility policies. I use the total number of students that enroll in a secondary school or a college as a proxy for the aggregate demand for education. The result of Table 1.9 shows that the aggregate demand for secondary education does not change differentially between the groups of provinces for both periods. However, the aggregate demand for college increases for province in the first and second groups relative to the baseline provinces in the period 1982 to 1990. This relative increase corresponds to the implementation of fertility policies in these two groups. If the supply curve of college seats is upward sloping, the increase in aggregate demand will raise the price of college education and therefore the average schooling will decrease within a household in the first period. The coefficients for college in Table 1.4 are likely to overestimate the effect from the quantity-quality trade-off since there is a positive effect through the aggregate demand channel implied by Table 1.9, meaning the true effect of the quantity-quality trade-off is negative and has a bigger magnitude than what is estimated in Table 1.4. Table 1.9 also shows that the demand for college education does not differ between groups of provinces in the period 1990 to 2000. Therefore, the estimated negative effect of the number of older children on the average schooling of children in Table 1.5 is less likely to be biased than the estimated effect from Table 1.4.

Table 1.7: The effect of policies on selection bias
Dependent variable is the share of children aged 16 to 21 living in a collective unit

VARIABLES	hs 1982-1990	coll 1982-1990	hs 1990-2000	coll 1990-2000
BeijingXyear	0.211***	0.799***	-0.167**	-0.827***
	(0.055)	(0.235)	(0.075)	(0.269)
FujianXyear	0.092***	0.164	0.061	-0.131
	(0.031)	(0.147)	(0.052)	(0.211)
AnhuiXyear	0.090***	0.100	0.026	-0.135
	(0.033)	(0.159)	(0.043)	(0.194)
Macro variabl	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Prov FE	Y	Y	Y	Y
Children Age FI	Y	Y	Y	Y
Observations	336	320	336	314
R-squared	0.684	0.654	0.567	0.390

The variables BeijingXyear, FujianXyear and AnhuiXyear are the interaction of a dummy indicating a region and a dummy indicating the end year of each period. Beijing indicates the first region that includes Beijing, Tianjin, Shanghai and Liaoning. Fujian indicates the second region that includes Fujian, Gansu, etc. Anhui indicates the third region that includes Anhui, Henan, etc. The omitted region, which is also the baseline region, includes Guangxi, Yunnan, etc. I control for the socioeconomic variables at the province and year level, the indicator for the end year of each period, province fixed effects, and age fixed effects for children.

Table 1.8: The effect of # of older/younger children on high school and college education, male and female no region-I

1982-1990

	hs	hs	coll	coll
	82-90	82-90	82-90	82-90
VARIABLES	male	female	male	female
N_Oy	0.028	-0.002	-0.001	-0.016
	(0.095)	(0.089)	(0.013)	(0.010)
N_yy	0.022	0.002	-0.004	-0.012
	(0.077)	(0.076)	(0.010)	(0.009)
Observations	512,863	483,789	512,863	483,789
R-squared	0.039	0.082	0.019	0.005

1990-2000

	hs	hs	coll	coll
	90-00	90-00	90-00	90-00
VARIABLES	male	female	male	female
N_Oy	-0.020	-0.072	-0.025**	-0.041**
	(0.040)	(0.057)	(0.011)	(0.018)
N_yy	0.103	-0.050	0.005	-0.001
	(0.066)	(0.080)	(0.010)	(0.013)
Observations	300,207	274,934	300,207	274,934
R-squared	0.068	0.200	0.014	0.008

The variable N_Oy is the number of older children in a household and the variable N_yy is the number of young children in a household. I drop the first group of provinces, which includes Beijing, Tianjin, Shanghai, and Liaoning.

Table 1.9: The effect of fertility policies on aggregate school enrollments

	1982-1990		**1990-2000**	
	[1]	[2]	[3]	[4]
VARIABLES	**hs+middle**	**college**	**hs+middle**	**college**
BeijingXyear	-10.459	3.238***	74.743	6.634
	(13.878)	(1.021)	(53.156)	(8.611)
FujianXyear	-18.100	1.309**	49.280	6.725
	(11.396)	(0.561)	(41.495)	(4.640)
AnhuiXyear	-15.765	0.818	31.437	4.468
	(19.727)	(0.992)	(42.796)	(5.297)
Year FE	14.151	-0.706	-112.103	-7.656
	(25.113)	(1.563)	(83.579)	(7.244)
Prov FE	Y	Y	Y	Y
Macro	Y	Y	Y	Y
Observations	55	56	56	56
R-squared	0.991	0.970	0.995	0.987

The variables BeijingXyear, FujianXyear and AnhuiXyear are the interaction of a dummy indicating a region and a dummy indicating the end year of each period. Beijing indicates the first region that includes Beijing, Tianjin, Shanghai and Liaoning. Fujian indicates the second region that includes Fujian, Gansu, etc. Anhui indicates the third region that includes Anhui, Henan, etc. The omitted region, which is also the baseline region, includes Guangxi, Yunnan, etc. The dependent variable in column [1] and [3] is the total enrollment in high school and middle school together of each province in each year. The dependent variable in column [2] and [4] is the total enrollment in college of each province in each year. High school includes 10th, 11th, and 12th grades. Middle school includes 7th, 8th, and 9th grades. Year FE is a dummy indicating the end year in each period.

1.7 Conclusion

This paper examines how fertility policies implemented in the early 1960s and 1970s affected the fertility level of Han Chinese women and the subsequent human capital investment in children resulting from fewer children. Using the variation in the timing of policy implementation across provinces, I find fertility policies combined with access to birth control methods substantially reduced the number of children born to a woman. Further, I find the reduction in the number of old siblings increases the college enrollment rate of children 16 to 21 years of age over the period 1990-2000. The magnitude of the impact indicates that fertility policies explain more than 50 percent of the difference in the increase in the college enrollment rate across provinces.

While the impact of fertility policies on the human capital investment seems to be large, we should be cautious in interpreting the result as direct evidence for the quantity-quality trade-off. We know that the provincial governments imposed large penalties on additional children outside quota. Nevertheless, the media continually reports that parents who did not pay for the penalty for their unplanned children cannot enroll them in public school[10]. This points directly to a second channel through which education is affected: an increase in the price of education. Moreover, this channel is particularly relevant after the one-child policy was implemented. The consequence of fertility policies on education in the period 1990-2000 should be viewed as influenced by both the quantity-quality trade-off and an increase in the price of education.

One interesting finding is that the difference in the impact of fertility policies on the college enrollment is insignificant between two periods. The impact is large and significant in the period 1990-2000, but it is insignificantly different from zero although negative in the first period 1982-1990. One possible explanation is that the supply of college education was inelastic in the first period. Even if the demand for college education increased in the household, with a fixed amount of seats in college, the chance of getting into a college would not increase much especially when the size of the pool of applicants did not change much. An interesting topic for future research is to look at the impact of fertility policies on the supply of public or even private education in China.

1.8 Appendix

1.8.1 Simple model of fertility and schooling

This is a simple model that explains how the number of children and the average schooling of children are determined in the household after the fertility policies have been implemented.

The utility function of the household depends on the number of older children, the number of young children and the education of older children, and the consumption of parents in the household. The utility is $U(S_0, n_0, n_y, X)$ where S_0 is the average years of education of older children; n_0 and n_y are the number of older and young children; X is the consumption level of parents in the household. The education of young children is not added to the model. Young children aged 15 or below only have compulsory schooling. I assume parents do not derive utility from their education. I also ignore the consumption of children for simplicity and it is not a crucial ingredient to derive the prediction of the model.

[10]For example, see the article Fighting for identity in The Economist, May 17, 2014.

The household maximizes the utility of the household subject to the constraint:

$$p_X X + n_0 f(S_0) + c(n_0, n_y, \theta_F, \theta_G) = y \tag{1.8.1.5}$$

In the above equation y is the household total disposable income, p_X is the price of consumption goods, $f(S_0)$ is the average cost of education at the average level S_0, and $n_0 f(S_0)$ is the total cost of education for all children in the household, given that the young children have free education or their education costs less money. The last term $c(n_0, n_y, \theta_F, \theta_G)$ is the cost of having more children than what the policies promote. The terms θ_F and θ_G indicate the implementation of the fertility policies. The total cost of more children $c(n_0, n_y, \theta_F, \theta_G)$ can represent the social support payment for extra children in the family under the one-child policy or income benefits that parents are disqualified from under the later longer fewer policy. I use θ_F and θ_G to distinguish the policy impact on older children and young children. Young children may be born under a different policy and the cost of excess birth can be different. For simplicity, I assume a linear form of the cost function, i.e., $c(n_0, n_y, \theta_F, \theta_G) = n_0\theta_F + n_y\theta_G$. From maximizing the utility subject to the budget constraint, it yields the first order conditions:

$$\frac{\partial U}{\partial X} = \lambda P_X \tag{1.8.1.6}$$

$$\frac{\partial U}{\partial S_0} = \lambda P_{S_0} = \lambda n_0 \frac{\partial f(S_0)}{\partial S_0} \tag{1.8.1.7}$$

$$\frac{\partial U}{\partial n_0} = \lambda P_{n_0} = \lambda \left(f(S_0) + \frac{\partial c}{\partial n_0} \right) = \lambda f(S_0) + \lambda \theta_F \tag{1.8.1.8}$$

$$\frac{\partial U}{\partial n_y} = \lambda P_{ny} = \lambda \frac{\partial c}{\partial n_y} = \lambda \theta_G \tag{1.8.1.9}$$

In this set of conditions λ is the marginal utility of income, $(f(S_0) + \partial c/\partial n_0)$ is the shadow price of older children, $\partial c/\partial n_y$ is the shadow price of young children, and $n_0 \times \partial f(S_0)/\partial S_0$ is the shadow price of the quality of children. Substituting the first order conditions into the budget constraint yields[11]:

$$S_0 = S_0(n_0, n_y, P_X, y) \tag{1.8.1.10}$$

$$n_0 = n_0(\theta_F, S_0, n_y, P_X, y) \tag{1.8.1.11}$$

$$n_y = n_y(\theta_G, n_0, P_X, y) \tag{1.8.1.12}$$

$$X = X(S_0, n_0, n_y, P_X) \tag{1.8.1.13}$$

These structural functions predict how the number of older and young children and the schooling of older children, n_0, n_y, and S_0, are generated in the household according to the parent taste for schooling and the budget constraint. I will estimate 1.8.1.10 to test the theory of quantity and quality trade-off. I will use the timing of the implementation of fertility policies θ_F and θ_G as instrumental variables for n_0 and n_y. The timing is assumed to be independent of other variables that affect n_0, n_y, and S_0.

[11]From the equations 1.8.1.6 and 1.8.1.7, we can derive X and λ as a function of other variables: $X = X(S_0, n_0, n_y, P_X)$, the function 1.8.1.13, and $\lambda = \lambda(S_0, n_0, n_y, P_X)$. Using the derived functions of λ and X, and the equations 1.8.1.8 and 1.8.1.9, we can derive θ_F as a function $\theta_F(S_0, n_0, n_y, P_X)$ and θ_G as a function $\theta_G(S_0, n_0, n_y, P_X)$. Plug these functions into the budget constraint. It yields the function 1.8.1.10. Substituting 1.8.1.10, X and λ into 1.8.1.9, it yields the function 1.8.1.12. Substituting 1.8.1.12, X and λ into 1.8.1.8, it yields the function 1.8.1.11.

Generally, it is different to tell the sign of the effect of n_0 or n_y on S_0 without any assumptions. Therefore it remains an empirical exercise whether the fertility policies improve the schooling of children. Suppose the utility function is separable, i.e., $U(S_0,n_0,n_y,X) = U_1(S_0)+U_2(n_0)+U_3(n_y)+U_4(X)$. Using the implicit function theorem, I can derive the derivatives of S_0 with respect to n_0 and n_y. However, the sign of the derivatives is ambiguous only with this assumption.

I further assume that the utility function is $ln(S_0)+a_1ln(n_0)+a_2ln(n_y)+a_3ln(X)$ and the cost of schooling $f(S_0)$ is a convex function, τS_0^2. Then the derivative of S_0 with respect to n_0 is negative; the derivative of S_0 with respect to n_y is zero. If I assume the utility function is linear in the number of children $ln(S_0)+a_1n_0+a_2n_y+a_3ln(X)$ and the cost of schooling is τS_0^2 then both derivatives are negative.

1.8.2 Additional tables

Table 1.10: Sample Means by Family Planning Program Region in 1982, 1990 and 2000

	Family Planning Program Regions				Test of Equal Means (F-Stat)	
	Region-I	Region-II	Region-III	Region-IV	Regions exclusive of region-I	All regions*
1982						
Log GDP Per Capita (1982RMB)	7.0[0.6]	5.9[0.3]	5.8[0.2]	5.6[0.5]	2.21	12.8
Log Fixed Asset Investment Per Capita (1982RMB)	5.6[0.6]	4.4[0.3]	4.2[0.4]	4.1[1.0]	0.66	6.7
Population density (100 persons per sqaure kilometer)	2.0[1.0]	1.0[0.5]	1.1[0.4]	1.0[0.5]	0.08	3.56
No. Tertiary Teachers/100,000 people	164[116]	25.5[9.2]	26.4[12.0]	19.1[7.5]	0.95	12.64
1990						
Log GDP Per Capita (1982RMB)	7.3[0.4]	6.4[0.3]	6.3[0.3]	6.1[0.6]	1.79	8.63
Log Fixed Asset Investment Per Capita (1982RMB)	6.1[0.5]	5.0[0.4]	4.7[0.5]	4.5[0.8]	1.15	7.13
Population density (100 persons per sqaure kilometer)	2.4[1.2]	1.1[0.6]	1.2[0.4]	1.2[0.5]	0.05	3.9
No. Tertiary Teachers/100,000 people	184[124]	32.0[11.1]	33.6[16.8]	24.3[9.0]	0.87	13.1
2000						
Log GDP Per Capita (1982RMB)	8.0[0.5]	7.2[0.5]	7.1[0.3]	6.6[0.5]	3.69	7.21
Log Fixed Asset Investment Per Capita (1982RMB)	7.0[0.7]	6.1[0.5]	5.9[0.4]	5.7[0.8]	1.07	4.62
Population density (100 persons per sqaure kilometer)	2.6[1.4]	1.2[0.8]	1.0[0.4]	0.7[0.2]	1.11	5.39
No. Tertiary Teachers/100,000 people	161[108]	35.8[9.5]	36.7[17.5]	28.0[10.7]	0.86	11.64

Note: F-statistics for all-region comparison show significance at 5% level at least.

Chapter 2

The Impact of Young Cohort Size on Adult Educational Upgrading: Evidence from Family Planning Policies in China

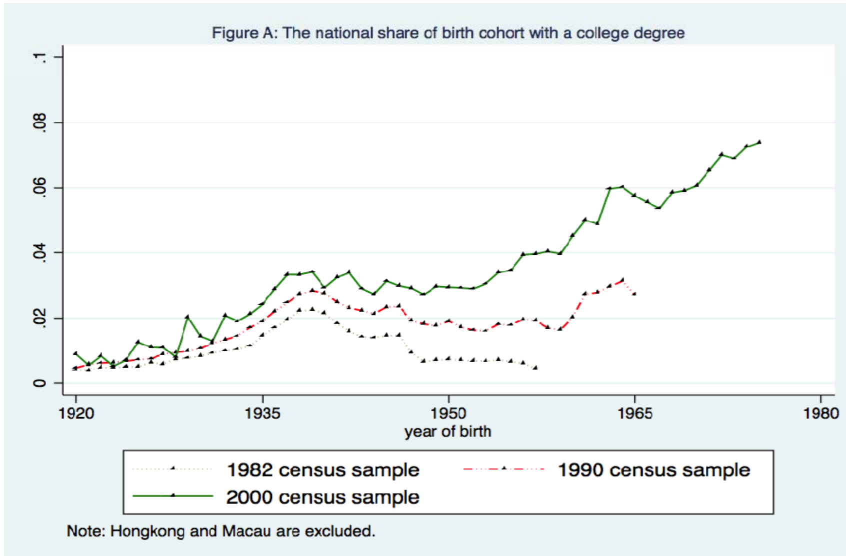
2.1 Introduction

The number of college educated workers in China as of 2000 has more than tripled compared with the level of two decades earlier. A number of empirical studies have documented increasing wage returns to college education over time (Zhang, Zhao, Park and Song 2005, Maurer-Fazio 1999). It has been found that there is an increase in the number of workers who went back to college to acquire a college degree[1]. In Figure 2.1 the national population share with a college degree by birth cohort indicates that the college attainment essentially increased for every cohort aged 45 or less in 1982 over eighteen yeas. Birth cohorts 25 to 34 years old (born from 1948 to 1957) achieved a nearly triple increase in the college completion rate in this period.

Standard models of human capital investment predict low returns to investment during one's adulthood (Carneiro and Heckman 2003) and empirical research has provided some evidence consistent with these models (Silles 2007). Nevertheless, using Chinese census data, I have also found an increase in the number of college graduates across all working-age cohorts over time. In this paper, I provide a simple explanation for the human capital accumulation of adults during the 1980s and 1990s in China. My paper is based on the premise that the dramatic decline in the young cohort size induced by the family planning polices in China increased the return to college education and the relative supply of college educated workers in the young cohort. In the presence of imperfect substitutability between college educated workers of different ages, this

[1] National adult education statistical bulletin on http://www.edu.cn and The world at work: Jobs, pay, and skills for 3.5 billion people by McKinsey&Company.

Figure 2.1: The National Share of Birth Cohort with a College Degree



increase in the relative supply of young college educated workers can lead to an increase in the productivity of older college educated workers. This productivity difference may raise the wage differential between older college educated workers and non-college educated workers, increasing the incentives for adults without a college education to upgrade.

My research can be linked to the literature on the impact of birth cohort size on wage rate and human capital levels. A research by Card and Lemieux (2001) found that the rising return to college educated workers in North American and United Kingdom was caused by the slowdown in the rate of growth of college graduates in birth cohorts born from the early 1950s onward. Other researchers have documented the large influx of young workers from the baby boom has had significant impacts on the schooling attainment of those born before and after the baby boom in the US (Stapleton and Young 1988, Wachter and Wascher 1984, Falaris and Peters 1992). However, the potential concern about these studies is that neither the supply of young educated workers nor the birth cohort size of young workers stems from an exogenous source. The fertility change and schooling outcomes could be individuals' behavioral response to the underlying economic and demographic factors that drive economic development. This could bias conclusions on the relationship between the supply of young workers and college returns or college attainment.

In this study I contribute to the existing research by using China's family planning policies as a source of exogenous variation in birth cohort size. A handful of studies have documented the effects of China's national one-child family planning policies on the total fertility rate in the recent decades (Gu et al. 2007, Attane 2002, Short and Zhai 1998). My research examines a longer period of the family planning policies from the initiation in the 1960s to the later 1970s. Specifically, I use the variation in the timing of the shocks to the young cohort size caused by polices which were adopted at different times in different regions. Under the general guideline of fertility control by the central government, some provinces had access to contraceptive measures as early as 1964, which others did not receive until the 1970s. Many provinces carried out the family planning practices around the mid-70s during the campaign called "Later Longer and Fewer". By 1979 fertility polices had reached all provinces in China with the broad principle of "one child per couple". I compare the differential declines in the young cohort size across regions where they carried out family planning earlier or later. The strong administrative power to enforce polices within a province allows me to predict the decline in the young cohort size sixteen years later, starting from different points of time for different regions.

I use aggregate province-year level data constructed from multiple Chinese censuses to measure the size of the young cohort aged between sixteen and twenty-four. I also use the aggregate data to measure the size of older college and non-college graduates for each birth cohort group in the age range between twenty-five and forty-nine. To apply the empirical strategy of a quasi-experiment to the relationship between the young cohort size and the older adults' education levels, I link the cross-regional changes in the number of college/non-college graduates before and after the decline in the young cohort size to the cross-regional declines in the young cohort size caused by family planning policies. This comparison separates the effect of the change in the young cohort size due to the fertility policies from the effect due to other factors that might affect regions more broadly. For example, a region can have a faster education upgrade over time because of factors associated with job specialization and the factors may be correlated with the changes in fertility as well. The empirical strategy will be valid as long as the variation in the timing of the initial implementation of fertility policies is orthogonal to changes in other confounding factors across regions and over time.

I also examine the impact on the number of adult college/non-college graduates by their age to gauge whether the increase is consistent with the cross-age distribution of the migration cost and the cost for older workers to go back to school. Further I distinguish the channel for the increase in the number of adult college graduates via: 1) in-migration of college graduates from another province to a province that has a decline in the young cohort size versus 2) educational upgrading of non-college graduates who are already in the province.

My research shows a strong impact of the decline in the young cohort size on the educational upgrading for all adult age groups. The decomposition of the increase in the number of adult college graduates suggests that the migration of college graduates seems to be less important for older adults aged above 25. The results are consistent with the prediction that older workers have higher costs of migration. The empirical results also show that the young cohort size does not have any significant impact on the size of adult non-college graduates, which means non-college educated migrants will replace the reduction in the number of local non-college graduates who upgrade their educational levels.

As a theoretical framework for the empirical analysis, I introduce the model of aggregate production with age-specific supplies of college educated workers from Card and Lemieux (2001). In their model, under the assumption that college educated workers of different ages are not perfect substitutes, the college premium for a given age group depends on both the aggregate relative supply of college educated workers in the labor force, and on the age-specific relative supply of college educated workers. Here, I maintain the assumption of imperfect substitutability of college educated workers across age groups. I interpret the change in the number of college graduates relative to non-college graduates across age groups as a response to the change in the age-specific college premium.

I also consider a model of "cohort crowding" from Bound and Turner (2007), in which a larger cohort that competes for limited public resources for education will end up with a smaller share that become college graduates. Bound and Turner show an inelastic pattern of adjustment where colleges might not increase enrollments when the cohort size is increasing, and enrollments might not drop during the time when the cohort size is shrinking. The theory predicts that the average education level will be higher for a smaller birth cohort. I show empirical evidence that the young cohort born after the family planning policies has greater college enrollment compared to their counterpart, the young cohort born prior to the policies. This is consistent with a relaxation on the constraint of the supply of college education. It might also be consistent with the quantity-quality trade-off of children at the household level.[2]

This paper can broadly contribute to the literature on demographic transition and wage inequality. My analysis of how a fertility decline affects human capital accumulation complements research by Bloom and Canning (2004) and Bloom, Canning and Malaney (2000), who argue that during the demographic transition when the mortality rate declines earlier than the fertility rate, it will generate a population age distribution with more working adults and fewer children. A boom of working-age population with a low dependency ratio will provide an opportunity of a period of fast income growth and economic development. In this paper, I provide an example of how the decline of young population caused by the fertility fall can lead to the human capital accumulation of people born earlier. Through quantifying these results, I contribute to the literature on wage inequality (Autor and Katz 1999), and Mincer's analysis (1996) of economic development and growth of human capital.

2.2 Impact of Young Cohort Size on Adult Educational Upgrade

I use the model created by Card and Lemieux (2001) to explain how the young cohort size will increase returns to college education for older adults. The model predicts that the increased education level among the young cohort will increase the productivity of older college educated workers and therefore their returns to college education, and thus lead to adult college upgrading.

In the Card and Lemieux model, workers with the same education level but different ages are imperfect substitutes. This is modeled in (2.2.0.1) using CES functions for aggregate labor of each education level, high school and college, where i indicates cohort, t indicates time, η is a parameter defining the elasticity of substitution between workers of the *same* education level, and α and β are efficiency parameters.

$$H_t = \left[\sum_i (\alpha_i H_{it}^{\eta})\right]^{1/\eta}, C_t = \left[\sum_i (\beta_i C_{it}^{\eta})\right]^{1/\eta} \tag{2.2.0.1}$$

The production function itself, equation (2.2.0.2), is also a CES function with ρ defining the elasticity of substitution between workers of *different* education levels and the θ parameters measuring the technological efficiency of each education type at time t. The authors assume that the ratio of wages for different groups is equal to the ratio of their marginal products and thus derive in equation (2.2.0.3) the ratio of wages ("college

[2]In another chapter of my thesis, I will examine the potential mechanisms by which the educational levels of young birth cohort can be affected by the family planning policies. Here I show the reduced empirical results of the impact of family planning policies on the educational attainment of young birth cohort.

premium") as a function of efficiency parameters, elasticities of substitution, cohort specific sizes of each worker type, and the aggregate quantities of each worker type.

$$y_t = (\theta_{ht} H_t^{\rho} + \theta_{ct} C_t^{\rho})^{1/\rho} \tag{2.2.0.2}$$

$$ln\left(\frac{w_{it}^c}{w_{it}^h}\right) = ln\left(\frac{\theta_{ct}}{\theta_{ht}}\right) + ln\left(\frac{\beta_i}{\alpha_i}\right) + \left[\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right] ln\left(\frac{C_t}{H_t}\right) - \frac{1}{\sigma_A} ln\left(\frac{C_{it}}{H_{it}}\right) + \varepsilon_{it} \tag{2.2.0.3}$$

The key term for this paper is $\left[\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right] ln\left(\frac{C_t}{H_t}\right)$, which shows how the aggregate ratio of college educated workers to high school workers can affect the cohort-specific college premium. The elasticity of substitution of workers of the same education levels but different ages is $\sigma_A = \frac{1}{1-\eta}$ while the elasticity of substitution for workers of different education levels is $\sigma_E = \frac{1}{1-\rho}$. If workers of different education levels are more substitutable than workers of the same education but different ages, $\sigma_E > \sigma_A$, then an increase in the aggregate ratio of college educated workers to high school workers can raise the cohort specific college premium. In the case of China, I will argue that regions which initiated fertility policies earlier experienced earlier declines in cohort size. If these smaller cohorts had access to the same total supply of schooling then a greater percentage of the cohort would become college educated and thus the aggregate ratio ($\frac{C_t}{H_t}$) would rise. Assuming that $\sigma_E > \sigma_A$, this would raise the education premium for older cohorts, providing an incentive for older high school educated workers to upgrade their education.

However, this result will depend on the supply elasticity of college education. The study by Bound and Turner (2007) suggests that fixed resources allocated to higher education institutions tend to decease the share of the college educated in a larger birth cohort. Adult workers studied in this paper were born in a larger birth cohort, so limited educational resources can be a constraint to their college enrollment. Even though they may have anticipated that the college premium would increase in the near future, they were crowded out of college when they were young and had to upgrade their education as adults. As colleges or adult institutions gradually adjust their cost to provide more seats for adults we should see less educational upgrading.

Finally, a key assumption for this argument is that $\sigma_E > \sigma_A$. Card and Lemieux estimate these elasticities with wage data but I do not have access to appropriate wage data. I cannot directly estimate these elasticities and therefore just note this model is consistent with my findings.

2.3 Research Design for Causal Relationship

This study aims to estimate the causal effect of the young cohort size on the educational upgrade of older adults. A simple cross-sectional comparison between the young cohort size and the number of adult college educated workers may be biased. While it may be reasonable to expect fertility changes to affect cohort sizes, there could be other unobserved factors affecting both cohort sizes and adult human capital accumulation. I will exploit an exogenous source of variation in the young cohort size caused by the family planning policies in China.

It is likely that the bias of the cross-sectional estimate is negative, meaning the estimated effect is more negative than the true one. For example, a shift in the parental preference for fewer but better educated children could be correlated with a secular trend of skill biased technological improvement which increases the productivity and the wage returns of skilled labor. As a result, more adults may acquire college education. Another possible case for a downward bias is related to population density or urban congestion. Denser areas could have both greater demand for college educated labor (perhaps as a result of agglomeration economies, foreign trade, or the location of government offices in big cities) and also experience a greater decline in fertility (as would result from higher housing costs or other costs of raising a child). This correlation could also lead to a downward bias.

An upward bias, i.e., the cross-sectional estimate is biased towards zero or being positive, is also possible. One omitted variable for an upward bias could be the local subsidies for higher education. Local subsidies could be correlated with many underlying characteristics in that area, such as the local GDP level, the local financing capacity for education, the existing institutional facilities, and/or political favoritism to educational development. Greater subsides allow a greater supply of adult college educated labor. At the same time, greater education resources may attract more young people.

The labor mobility of young people can also lead to an upward bias. Suppose young people are perfectly mobile. Any shock to a local labor market resulting from the fertility change will be neutralized, because young people can move across locations to equalize their marginal productivity. The cross-sectional relationship between the young cohort size and the number of adult college educated workers is insignificantly different from zero as long as the effect on the adult educational upgrading is the same across locations.

Since a cross-sectional comparison is possibly misleading, a second identification strategy is to link the change in the adult educational upgrading to the change in the young cohort size. The reliability of this strategy depends on whether the unobserved characteristics determining the two variables of interest are constant over time. If the unobserved characteristics do not stay constant over time and the changes in the unobserved characteristics are correlated with the changes in the variables of interest, then this strategy is problematic.

The advantage of my identification strategy is to utilize the timing of the initiation of the family planning policies in China. These policies create a random shock to the supply of young labor across regions. Specifically, I compare the change in the number of adult college graduates over time in regions where there is a large decrease in the young cohort size caused by earlier adoption of the policies to the change over time in regions where there is no decrease or a smaller decrease in the young cohort size caused by later adoption of the policies.

The key assumption is that the geographical variation in the timing of the initiation of the policies is exogenous to the geographical and temporal variations in other confounding factors. In other words, the region specific changes in the cohort size of young people are not capturing changes due to other unobservable factors. Additionally, all the influences from the income level, preferences for more education, population, population density, and labor mobility will be taken into account by controlling for fixed effects and other characteristics that are varying across regions and over time.

Although I can not directly test the key assumption, the variation in the timing of young cohort declines matches exactly with the variation in the initiation times of the policies. Other factors driving the fertility change are unlikely to change so dramatically at a point in time that growth of the young cohort size is reversed abruptly. It is hard to imagine the timing of the family planning policies planned by the central government in the 1960s will be correlated with the future changes in certain unobserved factors that influence the adult educational upgrading.

2.4 Background Knowledge on Family Planning Policies and Adult Educational Upgrading in China

2.4.1 Family Planning Policies

The earliest family planning work was organized by the Ministry of Public Health in 1956, but the impact of the early work seemed fairly temporary. The total fertility rate in 1957 was even higher than that before 1955. The concept of reducing population received a lot of debate among top leaders and family planning work was not carried through the Great Famine period (1958-1961). After 1963, the government began to implement stronger policies.

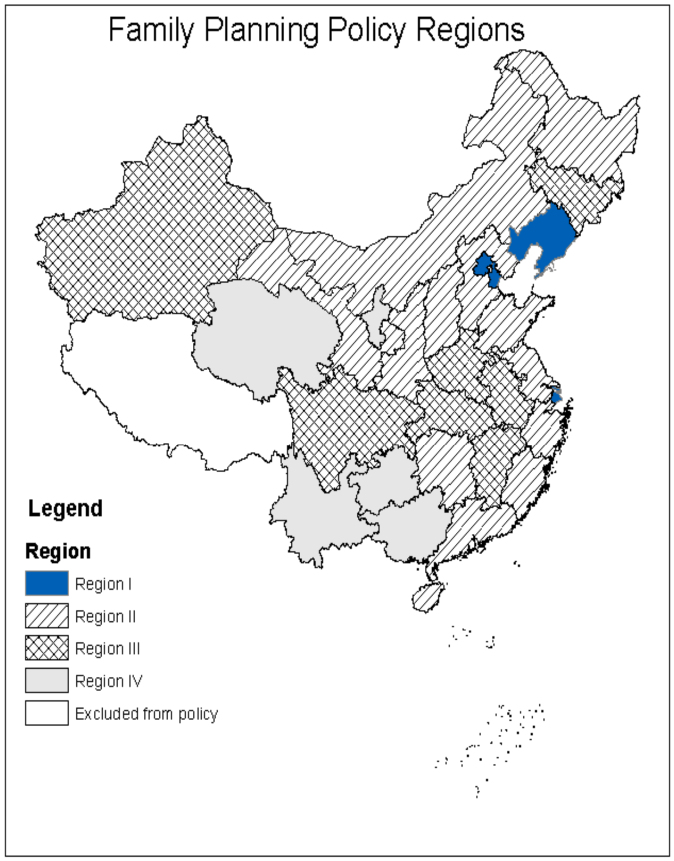
In 1964, the national agency, known as the Family Planning Committee of the State Council, was created to be responsible for making family planning policies and cooperate with local agencies. However, the main function of the committee was to educate people and encourage them to adopt family planning measures; these measures included a recommended later marriage age, technical guidance on birth control methods and surgeries, and rewards such as housing benefits for compliance with family planning work, etc. There was no enforcement on the policy implementation. Minority residence areas were completely exempt from family planning policies. Inclusive of other forces that might play a role, the family planning policies were effective in reducing the population. The total fertility rate declined from 7.5 in 1963 to 5.3 in 1967, a near 30 percent decrease. However, the event of the Cultural Revolution interrupted the policies and they were resumed after 1970.

After 1970, the main steps in family planning work included making birth control pills and other methods and surgeries free in the entire country, ordering local governments to establish family planning agencies, increasing education intensity and media coverage to less dense areas, improving quality of birth control methods, improving health care of women and children at birth, and investing more in research of birth control pills. The family planning programs during the period covered most provinces including some minority areas in case people would have a demand for family planning. However, the minority concentrations still remained the exemption. Meanwhile, the central government decided to increase the intensity of education over family planning adoption. In 1973, the State Council initiated the "Later Longer Fewer" family planning program. It encouraged every young couple to marry at a later age, adopt a longer birth spacing, and have a smaller family size[3].

The family planning work was quite successful in the 1970s. The total fertility rate declined from 5.8 in 1970 to 2.75 in 1979. In 1979 the government realized that it was difficult to meet the targeted population

[3] "Later" means a marriage age of 25 for men and 23 for women. Some more stringent rules specify age 28 for men and age 25 for women. "Longer" means the birth spacing should be more than three years. "Fewer" means that a couple is recommended to have two children or fewer. In some circumstances, rural couples are encouraged to have three or four children or fewer.

Figure 2.2: Map of Provinces Implementing Family Planning Policies



growth rate because the large cohort born in the late 1950s was in the primary reproductive stage. Therefore the government developed more forceful measures. New family planning policies, a.k.a., the one-child policy, rewarded couples who decided to have only one child. In 1982, the policies became that one child was just enough; two children were allowed in certain circumstances; and three children were generally not allowed. Minority areas were ruled by less strict policies. Urban couples could have two children, while rural couple could have three to four children under certain circumstances.

Based on the adoption timing of family planning policies, I divide all provinces into four broad regions. The earliest region includes Beijing, Shanghai, Tianjin and Liaoning. These provinces adopted the policies in 1963-1964. In the next region I group Fujian, Gansu, Guangdong, Hebei, Heilongjiang, Hunan, Inner Mongolia, Jiangsu, Shaanxi, Shanxi, Shandong and Zhejiang. They started family planning work around 1971-1973. Following them, Anhui, Henan, Hubei, Jiangxi, Jilin, Sichuan and Xinjiang came to adopt the policies during 1974-1975. The last group of provinces includes Guangxi, Guizhou, Ningxia, Qinghai and Yunnan, which began to implement the polices from 1979-1983. Figure 2.2 shows the geographic distribution of provinces that are included in these four regions.

The municipalities of Beijing, Tianjin and Shanghai had effective family planning programs in the early 1960s. Beijing and Tianjin provided free birth control methods and surgeries; they had also extended the medical leave and provided goods or monetary compensation for birth control surgeries. Shanghai offered free birth control surgeries and started initiating clinical trials of birth control pills. Two kinds of birth control pills for women were approved in 1967. Some other provinces also provided free birth control surgeries, but the policy effectiveness in those provinces was weaker than in the three municipalities. In a strongly planned economy with a system of hierarchical administrative bureaucracies, the municipalities may have directly received more support and supervision from the central government.

The majority of provinces adopted family planning policies before 1979. I divide them into two groups by the criterion whether they had initiated the policies before the national program "Later Longer Fewer" in 1973. In this way, I create a clear time cut-off between the two groups. In the 1970s, these provinces started to receive technical support on birth control methods from the central government. With access to birth control methods and the influence from the "Later Longer Fewer" program, the family planning work seemed to be effective.

Having implemented family planing policies only since 1979, the last group of provinces belong to the reference group. This group consists of major provinces with minority residents. They were least likely to be affected by family planning policies before 1979 and affected by the least degree after 1979 because of relaxation of the one-child policy in this group. I will treat them as the control group in evaluating the effects of the policies.

2.4.2 Adult Educational System and Educational Upgrading

The Chinese adult education system was initially created in the early 1950s and became popular in the late 80s. The development of the adult educational system can be divided into three stages: the startup in the early 1950s, the expansion from the late 1970s to the early 1990s, and the further development from 1993.

The adult educational system consists of two different tracks: a diploma-oriented system and a non-diploma-oriented system. The non-diploma-oriented system caters to the demand for knowledge or skill improvement related to work or personal interests. Programs such as job training and vocational qualification belong to this system. I will focus on adult educational progress in the diploma-oriented system, in which the goal is an academic diploma. In this study particularly, it is a diploma at the tertiary level of education.

The diploma-oriented system, combined with social assistance, was designed to help adults achieve academic degrees through self-learning. To meet academic requirements, adult students normally take an entrance exam and enroll in a higher educational institution, and earn credits for a diploma, just like a young college student. A typical program usually takes two or three years to complete, depending on whether the learning is full-time or part-time. The credit requirement of the degree depends on the academic difficulty set by the programs or schools. This form of adult study has existed since the early 1960s. Alternatively, adult students can participate in a self-study program, which provides a concentrated period of self-studying in the classrooms of public schools like night schools, correspondence schools, or television schools organized by local colleges or universities[4]. College teachers would be available if there was a need from students for

[4]Adult education at the tertiary level in China can have various forms. Night colleges for adults utilize the time during the night to teach and learn. Correspondence colleges focus on self-study and sometimes provide a concentrated study period. TV colleges use a broadcast system to learn and most often are equivalent to the correspondence colleges. Vocational colleges are special colleges founded by larger enterprises or local trade unions, in which learning is done in spare time and mainly linked to work practice. In

authoritative assistance. At the end of the study, candidates are required to take a set of exams and have to pass with a minimum score to get the diploma. There is no time constraint with respect to the period of study or the period of time it takes to pass all the exams. This "self-study" (zi-kao) regime was introduced as a policy for adults to achieve higher educational levels in 1981.

Although it had existed for a long time, the diploma-oriented adult education started to emerge as a popular tool of achieving higher levels of education for people with some work experience and at least a high school diploma in the late 1980s. In 1999 there were about 0.8 million adults graduating from adult colleges, which was roughly the same number of college graduates from normal colleges in that year. Another 0.4 million adults became college graduates through the self-study examination[5].

2.5 Date Sources and Trends in Young Cohort Size and Adult Educational Upgrading

2.5.1 Data Sources

To track the educational upgrade of adults and link it to the change in the young cohort size, I need comprehensive data on the national geographic and temporal distribution of young and older adults with information over their education levels, migration status, and demographic characteristics. I will also use macro-level data on economic and demographic conditions to control for potential confounding effects across regions.

I use the Chinese censuses from the consecutive survey years 1982, 1992 and 2000 for information on demographics and education. The census data are sampled at the level of household, either domestic household or collective living unit[6]. They are a representative sample of all individuals who have Chinese nationality and live in the country at the time of the survey. The sampling rates are 1%, 1% and 0.095% for the three survey years. The census data provide many advantages for the purpose of my study. They provide detailed personal information on sex, year of birth, month of birth, ethnicity, province of residence, educational levels, educational status, migration history, etc. The censuses also survey characteristics at a household level, including the number of persons living in the household, the number of births and deaths in the previous year, the number of registered persons absent over one year, and the total number of alive births to a woman[7].

The main variables to construct are the size of college graduates and the size of non-college graduates for older adults at the cohort-province-year level[8], and the size of young cohort between age 16 and 24 at the level of province and year. I define college graduates as those with at least a college degree; non-college graduates as those with no more than a high school degree. I will track a particular birth cohort in a province over time to examine the educational increase for this group. I also examine the educational improvement of

recent years, the tuition of correspondence colleges ranges from 700-1500 US dollars per year. The learning experience could be either full-time or spare-time.

[5] National adult education statistical bulletin on http://www.edu.cn.

[6] A domestic household contains a group of individuals who are relatives or non-relatives living in the same household. A collective living unit can be a company's dorm for migrant workers or a university's dorm for college students.

[7] For detailed information, see the section of appendix.

[8] A year indexes a census year; a province indexes an administrative division as coded in the place-of-residence in the censuses. A cohort is a group of people born in the same year.

the young cohort themselves. Assuming they have not finished education before age 24, I look at the number of college students and college graduates together at the age-province-year level.

Second, I construct the size of the young cohort aged 16-24 at the province-year level. For example, to construct the cohort size in Liaoning province in 1982, I count the number of persons born between 1958 and 1966, living in Liaoning province in 1982 from the 1982 census. In this way, I can calculate the size of the young cohort in any province in the year of 1982, 1990 and 2000 [9].

A challenge for this paper is to distinguish the increase in the number of college educated adults driven by the in-migration of college adults from the actual educational upgrading from a high school graduate. To achieve this goal, I use migration information in censuses to identify migrant labor and labor native to a region. In 1990 and 2000 censuses, the migration question asks which province a person lived in five years ago. I will define an individual as a migrant if they lived in a different province five years ago.

Using the geographic variation in the initiation of family planning polices, I group twenty-eight provinces in the sample into four different regions. Region-I includes provinces that initiated the policy measures in 1964. It had a continuous fall in the number of births from 1964. This means starting in 1980, 16 years after the initiation of the policies, the young cohort size started to decline in region-I. Region-II consists of provinces that initiated the policies in 1971, leading to a decrease in the young cohort size from 1987. Region-III initiated the policies in 1974 and had a decrease in the young cohort size from 1990 and region-IV, being the last adopter of policies in 1978, had a decrease in the young cohort size from 1994.

Appendix Table 2.7 lists sample statistics. I divide the length of my study into two periods: 1982-1990 and 1990-2000. I list the age range of the young cohort and the birth year range of the older adults in my study. The age range of the young cohort is 16 to 24 years old. The birth year range for adults is 1933-1957 in the first period and 1941-1965 in the second period. For each period, I list the change in the size of the young cohort and the changes in other economic and demographic variables, the change in the number of college students and graduates within province and age for young people, and the change in the number of college and non-college graduates within province and birth cohort for older adults. I also list the change in the migrant share for young and older people. Lastly, I separate the statistics by regions of family planning policies.

At the national level, the natural logarithm of the size of the young cohort, or the log size of the young cohort on average increased by 0.24 in 1982-1990 and decreased by 0.3 in 1990-2000. The log number of young college students or graduates within an age group on average increased by 0.34 in 1982-1990 and increased by 1.67 in 1990-2000. The log number of adult college graduates within a birth cohort on average increased by 0.53 in 1982-1990 and increased by 0.44 in 1990-2000. The log number of adult non-college graduates within a birth cohort on average decreased but did not change much in both periods. Among the young college students or graduates, the share of them who lived in a different province five years ago decreased from 0.14 in 1990 to 0.04 in 2000. The share of migrants among the older college graduates decreased from 0.04 in 1990 to 0.01 in 2000, and the share of migrants among the older non-college graduates

[9] In the appendix, I describe a method to construct the size of young cohort in each province for any year between 1970 and 2000. Using the constructed size of young cohort at the province-year level, I can estimate the region-specific time trend which illustrates the impacts of family planning policies on the youth population over the entire period from 1970 to 2000. I will show the region-specific time trend in the next subsection.

Table 2.1: Sample Means by Family Planning Program Region in 1982 and 1990

	Family Planning Program Regions				Test of Equal Means (F-Stat)	
	Region-I	Region-II	Region-III	Region-IV	without region-I	All regions*
1982						
Log real GDP Per Capita (1982RMB)	7.0[0.6]	5.9[0.3]	5.8[0.2]	5.6[0.5]	2.21	12.8
Log real Fixed Asset Investment Per Capita (1982RMB)	5.6[0.6]	4.4[0.3]	4.2[0.4]	4.1[1.0]	0.66	6.7
No. Tertiary Teachers per 100,000 people	164[116]	25.5[9.2]	26.4[12.0]	19.1[7.5]	0.95	12.64
Population density (100 persons per sqaure kilometer)	2.0[1.0]	1.0[0.5]	1.1[0.4]	1.0[0.5]	0.08	3.56
1990						
Log real GDP Per Capita (1982RMB)	7.3[0.4]	6.4[0.3]	6.3[0.3]	6.1[0.6]	1.79	8.63
Log real Fixed Asset Investment Per Capita (1982RMB)	6.1[0.5]	5.0[0.4]	4.7[0.5]	4.5[0.8]	1.15	7.13
No. Tertiary Teachers per 100,000 people	184[124]	32.0[11.1]	33.6[16.8]	24.3[9.0]	0.87	13.1
Population density (100 persons per sqaure kilometer)	2.4[1.2]	1.1[0.6]	1.2[0.4]	1.2[0.5]	0.05	3.9

did not change over the period.

Additionally in Table 2.7, I break down the provinces by regions of policies. I find that provinces in region-I and region-II experienced a larger decline in the log young cohort size relative to the national average than provinces in region-III and region-IV in 1982-1990; provinces in region-III experienced a larger decline than provinces in other regions in 1990-2000. I also find that in the first period, both the number of young college students and the number of older college graduates increased relative to the national average in region-I, region-II and region-III, but decreased relatively in region-IV. In the second period, both measures increased relatively in region-II and region-III and decreased relatively in region-I and region-IV. For the number of older non-college graduates, the changes in the four regions did not differ from the national average level except region-I in the first period.

Further, I show the change in the migration share among college and non-college groups for each region. I find that the migrant share of young college students or graduates decreased more in region-I and region-III than in region-II and region-IV in the second period 1990-2000. However, the change in the migrant share of older college graduates and older non-college graduates did not differ across the four regions and did not differ from the national average change.

As I mentioned earlier, the key identification assumption is that the variation in the initiation of family planning policies across regions is not correlated with the variation in other economic conditions. I provide some evidence that the regions are similar to each other in dimensions other than the timing of the policies by testing the differences in these controlled variables at the beginning year of the two periods. Results are shown in Table 2.1. Except for region-I, provinces in region-II, region-III and region-IV do not have significant differences in log real GDP per capita, log real value of fixed asset investment per capita, the number of tertiary teachers per capita and the population density. Nevertheless, the identification assumption I impose is weaker than the test of non-differences in the initial levels of these controlled variables. The identification assumption is that the variation in the timing of the initiation of the policies is uncorrelated with the change in the number of college graduates net of the controlled variables except through the change in the young cohort size.

2.5.2 Trends in Young Cohort Size and Adult Educational Upgrading

I first show patterns in young cohort sizes across regions and illustrate the strategy to identify the adult educational upgrade. Figure 2.3 is the relative annual size of the young cohort aged between 16 and 24 in each year and in each region. Each dot in these sub-figures is calculated as the ratio of the regional young cohort size to the national young cohort size. The three lines in each sub-figure are generated from three census samples (see the appendix for more details). Essentially, any year t on the x-axis corresponds to the range of birth cohorts born between the year $t-24$ and the year $t-16$. The differences between the three lines for a given year t represent the intensity of the migration of the birth cohorts across the years 1982, 1990 and 2000. The lines from the 1982 sample and the 1990 sample match well with each other except in region-I and region-III there exists a small gap between two years. This implies the internal migration across regions is not an important issue in the period 1982-1990. The migration is more intensive for the period 1990-2000 across all regions. Those migrants are about twenty to forty years old in 2000. In additional analysis not included in this paper I have also found that the migration outside the country is not an issue in my study.[10]

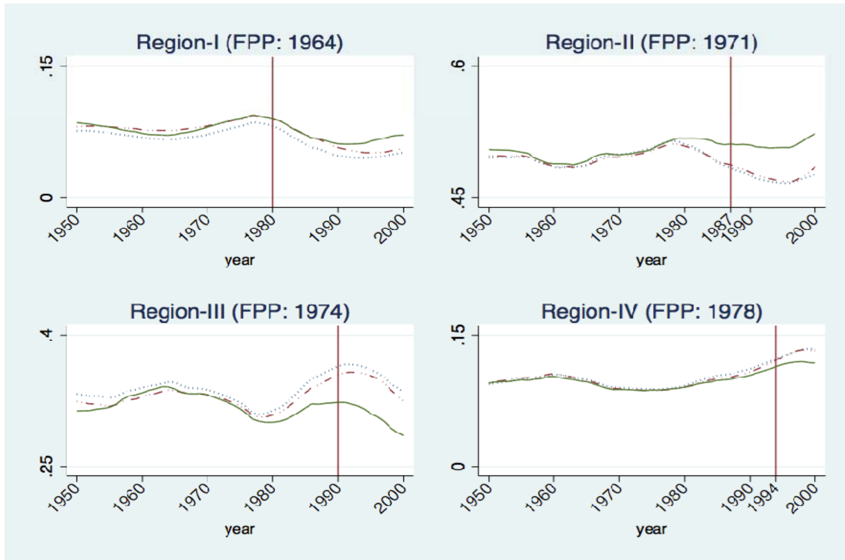
Figure 2.3 shows the pattern of the relative young cohort size that is aggregated at the regional level. I further use the provincial level data on the log size of the young cohort and subtract the national mean from it. I estimate the region-specific trend in the log size of the young cohort and plot the trend in Figure 2 (see the appendix for more details on the regression). Figure 2 shows a similar pattern of growth of the young cohort size as Figure 2.3. There are distinct patterns of growth of the young cohort size in four regions. The actual decline in the relative log number of the young cohort size happened in 1978 in region-I. The young cohort size increased slightly before 1978 and it kept decreasing fast through the 1980s and the earlier 1990s. After 1995, there was a slight increase in the young cohort size in region-I. In region-II, the relative log number of the young cohort size was at a constant level before 1985 and declined afterwards. The pattern of the log number of the young cohort size in region-III is slightly different. It was stable before 1991 and started declining from 1992 onward. The pattern in region-IV is constant before 1981 and increasing gradually from 1982 to 1996 and declining after 1997.

The vertical lines in Figure 2.3 and Figure 2.4 represent the different timing of the decline in the young cohort size predicted by the initiation time of the family planning policies. The year of the policy initiation is 1964 in region-I, 1971 in region-II, 1974 in region-III and 1978 in region-IV. Accordingly, the year of the decline in the size or the log size of the young cohort should be 1980 in region-I, 1987 in region-II, 1990 in region-III and 1994 in region-IV. Note that in Figure 2.4 the timing of the actual decline in the young cohort size conforms with the timing of the initiation of the policies, despite the fact that young cohorts might migrate between these regions. In addition, the pattern of the young cohort size is roughly the same across regions before 1980, which again suggests that there are not differences in other confounders that might determine the change in the young cohort size.

To illustrate the impact of the size of the young cohorts on the average educational level of themselves and the educational upgrading of the older adults, I plot the change in the size of young cohort aged 16-24 across regions on the x-axis, and the change in the college enrollments (i.e, college students and graduates)

[10] I plot the national size of each birth cohort born between the year 1940 and the year 1980 from three census samples, the gap between census samples is small. The figure is available upon request.

Figure 2.3: Young Population Distribution by Region



The plotted lines are the ratio of the regional young cohort size to the national young cohort size for each year (see the appendix for more details). The dot line is computed from the 1982 census sample. The dash line is from the 1990 sample. The solid line is from the 2000 sample. The family planning policies have been implemented since 1964 in region-I, 1971 in region-II, 1974 in region-III and 1978 in region-IV. The red vertical lines indicate the years when the young cohort size (or the rate of growth of it) is treated in the regions.

on the y-axis (see appendix for details on the regression), instrumented by the region-specific change due to the family planning policies. Figure 2.5 shows that on the x-axis, the variation of the change in the size of young cohort is large across regions. We can see a negative relationship between the young cohort size and the college enrollments of the young cohort. The magnitude of the slope is slightly greater in the period 1982-1990 than in the period 1990-2000.

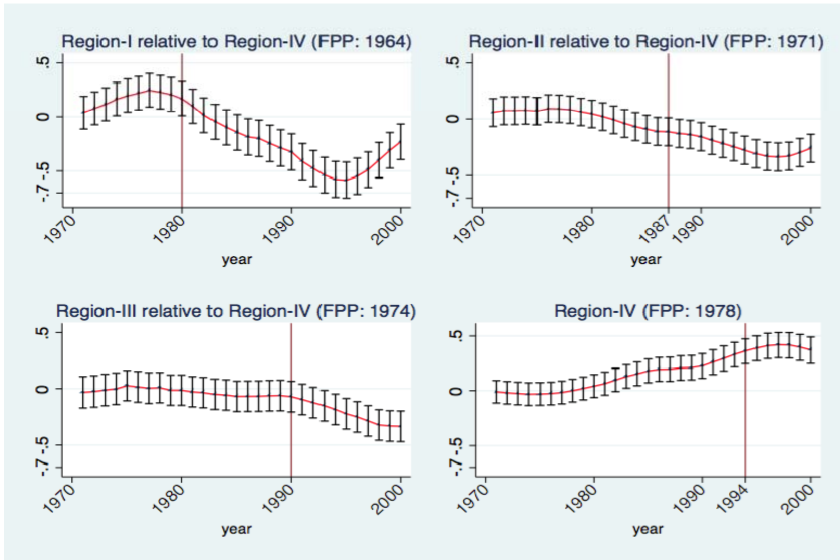
Figure 2.6 plots the change in the number of older college graduates aged 25-49 on the y-axis by age group and by period. The relationship shows that the decline in the size of young cohort increases the number of older college graduates. The magnitude differs by age group and by period. In the following sections I will lay out an econometric model and discuss the results.

2.6 Econometric Specification

In this section, I lay out a simple econometric model to estimate the impact of the young cohort size on adult educational upgrading. To illustrate, I assume the impact is homogenous across regions and over time. First, I will point out the potential problem with a simple cross-sectional comparison. Then I will show the econometric method in this study. A simple cross-sectional regression setup is:

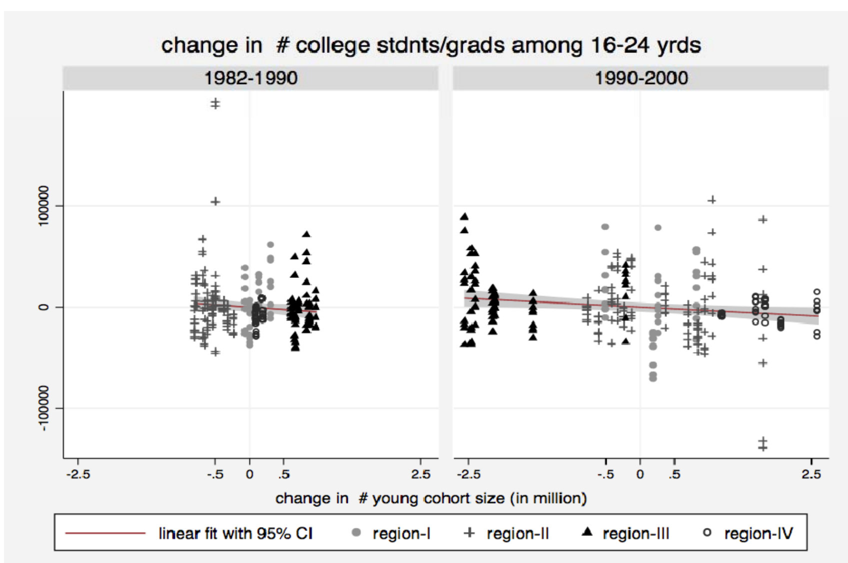
$$TC_{jpt} = \alpha ln(YC_{pt}) + \sum_{p} \beta_p Province_p + \gamma_t Year_t + \lambda X_{pt} + \sum_{j} \mu_j Z_j + \varepsilon_{jpt} \qquad (2.6.0.4)$$

Figure 2.4: Predicted Young Population Distribution by Region



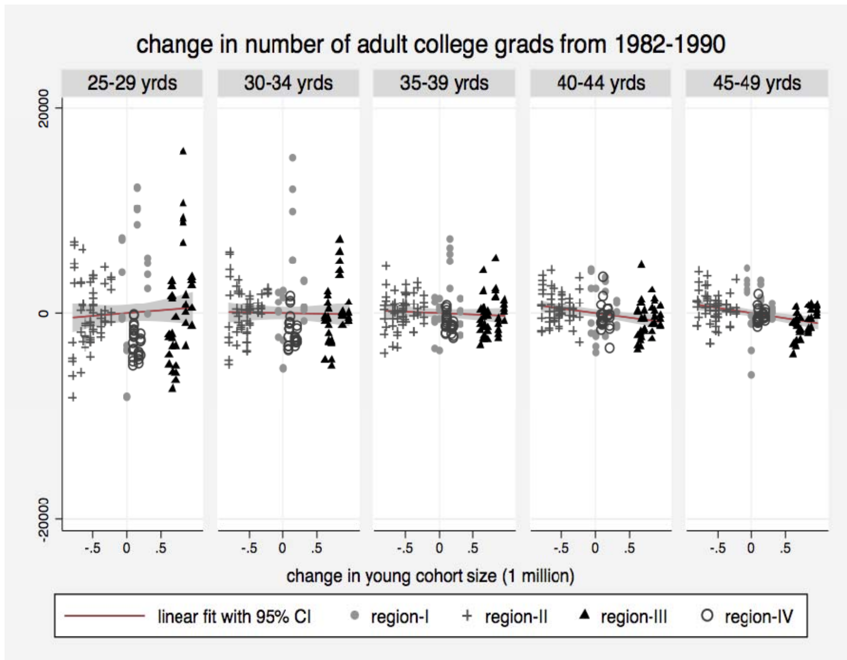
I use the demeaned province-year-level log number of young people aged 16-24 to predict the region-specific year trend of log number of the young cohort size (see the appendix for more details). The red lines with bars are the coefficients on the year dummies (Region-IV (FPP: 1978)), and the coefficients on the year dummies interacted with region dummies (Region-I relative to Region-IV (FPP: 1964), Region-II relative to Region-IV (FPP: 1971), Region-III relative to Region-IV (FPP: 1974)). The bars represent the 90% confidence intervals. The red vertical lines indicate the years when the log number of the young cohort size is treated in the regions.

Figure 2.5: Change in Educational Levels of Young Cohort Corresponding to Change in Young Cohort Size

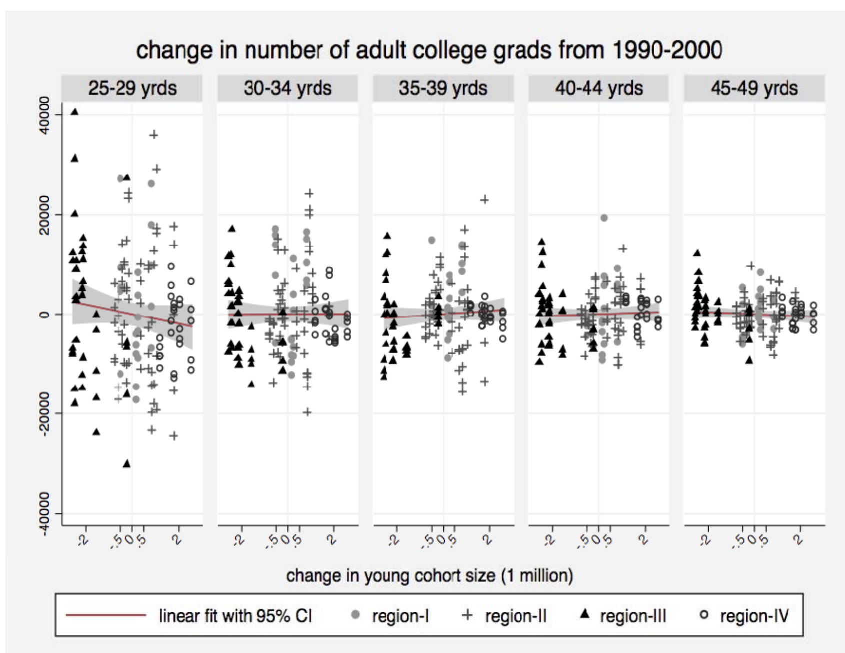


The plot is the relationship between the change in the number of young college students and graduates of age 16-24 within province and the change in the young cohort size. See the appendix for more details.

Figure 2.6: Change in Educational Levels of Older Adults Corresponding to Change in Young Cohort Size by Age Group



(a) 1982-1990 by Age



(b) 1990-2000 by Age

This figure by adult age group is the relationship between the change in the number of adult college graduates within province and birth cohort for 1982-1990 and 1990 to 2000, shown against the change in the young cohort size. See the appendix for more details.

where j indexes an adult birth cohort, p indexes province and t indexes time. The term TC_{jpt} is the total number of college graduates in each birth cohort, each province and each year while $ln(YC_{pt})$ is the natural logarithm of the number of the young cohort aged 16-24 in each province in each year. Here I use $ln(YC_{pt})$ instead of YC_{pt} because the latter is too large in magnitude compared with the change in the adult educational upgrading; $ln(YC_{pt})$ also has a better goodness-of-fit in the regression. The variables $Province_p$, $Year_t$, and Z_j are the province fixed effect, the time fixed effect, and the birth cohort fixed effect; X_{pt} are other characteristics at the province-year level. I use the birth cohort fixed effect to control for any impact on the aggregate number of college graduates that is from the population compositional change. The birth cohort fixed effect also captures any effect on the adult educational level from characteristics that are specific to a birth cohort, such as if a birth cohort was delayed to enter college during the Cultural Revolution. I use the province fixed effect to capture determinants for the adult education level that is constant in each province, such as if big provinces have more adult college educated workers. I use time fixed effect to control for general demand for college educated workers at the national level or the general supply shock to college education.

The OLS regression of the equation (2.6.0.4) will probably be an inconsistent estimate of the effect of the young cohort size on the number of adult college graduates. The unobserved characteristics in ε_{jpt} can be correlated with the young cohort size $ln(YC_{pt})$; such examples of characteristics would include agglomeration economies, local subsidies for college, or migration of young labor.

The identification strategy is to compare the change in the number of college graduates in regions with a large decrease in the young cohort size caused by family planning policies to the change in regions with no decrease or a smaller decrease in the young cohort size. I use a group of instrumental variables $region_1 \times year_t$, $region_2 \times year_t$ and $region_3 \times year_t$ for the endogenous variable $ln(YC_{pt})$. The interpretation is the region-specific time change in the log number of the young cohort size. I choose region-IV as the reference region. For example, the term $region_1 \times year_t$ indicates how much is the change in the log number of the young cohort size in region-I relative to the change in the reference region. The identification assumption is that the instrumental variables are uncorrelated with ε_{jpt}.

As a robustness check, I create a post-policy dummy interacted with region dummies $post_t \times region_r$. The variable $post_t$ is an indicator for whether the year t is after the treatment for a region. In the previous section, I stated that region-I will be treated from 1980, region-II from 1987, region-III from 1990 and region-IV from 1994. The interpretation is that for those regions who are treated, there may be a regional difference in the treatment effect. This may be due to a different length of time after the treatment, different measures adopted by provinces, and/or different degrees of the enforcement in these regions.

To examine the impact of the size of the young cohort on their own educational achievement, I regress the number of college students or graduates of the age group 16-24 on the natural logarithm of the number of the young cohort aged 16-24, where the instrumental variables for the log size of the young cohort are the same as for the adult education regression. The only difference in the regression specification is that I control for the age effect of the young cohort instead of the birth cohort fixed effect.

$$TC_{ipt} = \alpha ln(YC_{pt}) + \sum_p \beta_p Province_p + \gamma_t Year_t + \lambda X_{pt} + \sum_i \mu_i Age_i + \varepsilon_{ipt} \qquad (2.6.0.5)$$

Similarly, I also examine the impact of the young cohort size on the number of adult non-college graduates using the timing of the family planning polices as instrumental variables for the young cohort size. The specification is

$$TNC_{jpt} = \alpha ln(YC_{pt}) + \sum_{p} \beta_p Province_p + \gamma_t Year_t + \lambda X_{pt} + \sum_{j} \mu_j Z_j + \varepsilon_{jpt} \qquad (2.6.0.6)$$

where TNC_{jpt} is the total number of non-college graduates in each birth cohort, in each province and each year. I use this regression to address whether the non-college graduates migrate and to what extent they migrate in response to the changes in the young cohort size.

2.7 Results

2.7.1 Impact of Family Planning Policies on Young Cohort Size

Table 2.2 shows the impact of the family planning policies on the log number of the young cohort size. The results show that the effect of the policies on the log young cohort size, as indicated by the instrumental variables in the first three rows, is the same for both the young cohort and older adults but different in the two periods. In Panel A (1982-1990), controlling for other characteristics, the change in the log number of the young cohort size in region-I was 0.065 lower than the change in region-IV, and the change was 0.113 lower in region-II than in region-IV. Region-III did not have a significant difference in growth of the size of the young cohort than region-IV controlling for other characteristics.

Coefficients on other characteristics indicate that the log number of the young cohort size increased in region-IV by 0.542 which is expected growth in the young cohort size without the family planning policies. Log real GDP per capita and log real value of fixed asset per capita did not have a significant effect on the young cohort size in this period. The number of tertiary teachers per capita was negatively associated with the young cohort size. Population and population density were also negatively associated with the young cohort size in the period.

In the second period 1990-2000, the impact of the policies was substantial and negative for all regions except region-IV. The change in the log number of the young cohort size was 0.379 lower in region-I than in region-IV, 0.34 lower in region-II than in region-IV, and 0.469 lower in region-III than in region-IV. The year fixed effect indicates that the log number of the young cohort size declined by 0.063 insignificantly in region-IV. Log real GDP per capita was positively associated with the young cohort size; both log real value of fixed asset per capita and the number of tertiary teachers per capita were negatively associated with the young cohort size. Population and population density were positively associated with the young cohort size in the second period.

In the robustness check, I group provinces into the treated group and the control group. In the first period 1982-1990, the treated groups are region-I and region-II. The treatment effect was a 0.125 decrease in the log number of the young cohort size in region-II and a 0.077 decrease in region-I. Note that both provinces in region-I and region-II were exposed to the “Later Longer Fewer” program in this period; the difference is that region-II had access to the birth control methods in the period. In the second period 1990-2000, the treated groups are region-III and region-IV. All four regions were affected by the one-child policy. Additionally, region-III was affected by the "Later Longer Fewer" program and had access to the birth control methods. On top of what region-III had experienced in the period, region-IV was treated by the relaxed one-child policy. The results show that relative to region-I and region-II, region-III had an additional 0.125 decrease in the log number of the young cohort size, which is the same as the treatment effect for region-II in the first period. Conversely, region-IV had a 0.467 increase relative to region-III in the log number of the young cohort size because of the relaxed one-child policy.

In the first period, the general time trend in the log number of the young cohort size was 0.551. In the second period, the log number of the young cohort size declined by 0.418 because of the one-child policy. Other characteristics had the same effects as in the previous setup specifying a region-specific trend in the young cohort size.

2.7.2 Impact of Young Cohort Size on College Attainment of Young Cohort

The relationship between the log number of the young cohort size and their own educational attainment is important in its own right. Additionally, the results for the young cohort combined with that for the older adults can inform the mechanism for the adult educational upgrading.

In Table 2.3, the OLS regression results suggest the relationship between the log number of the young cohort aged 16-24 and their college enrollments is insignificantly different from zero in both periods. The estimates from the 2SLS regression indicate a negative relationship between the log number of the cohort size and their college enrollments. A one percent decrease in the young cohort size increased the number of college enrollments by 679 persons in the first period and 394 persons in the second period. The robustness check shows that the results do not change substantially. When I examine the effect on the college enrollments separately by gender, I find that the effect of the log number of the young cohort size on the college enrollments is larger for male youth than for female youth. In the second period, the effect for female youth is essentially indifferent from zero.

To sum up, empirical results show that a smaller young cohort size implies more young people who enroll in college. This is consistent with relaxation of the constraint on college education supply (cohort crowding). It can also be consistent with the quantity-quality trade-off of children within household.

2.7.3 Impact of Young Cohort Size on Adult Educational Increase

Table 2.4 presents the OLS results for the relationship between the log number of the young cohort size and the number of adult college graduates. Panel A (the period 1982-1990) results show that a one percent increase in the young cohort size decreased the number of older college graduates by 48-100 persons for the age group 25-39; the relationship is insignificant for the age group 40-49. The time trend for adult college graduates was positive over time with exception of the group over 40 for whom there was a negative change. Log real GDP per capita had no significant role in increasing the number of college graduates for those below 40, but the effect was significantly positive for those above 40. Log real value of fixed asset per capita had a negative effect on the number of college graduates. The number of tertiary teachers per capita had a positive

Table 2.2: First stage regression of impact of family planning policies on provincial supply of young labor

Dependent var: Provincial-year-level log number of young people aged 16-24

	Panel A: 1982-1990 sample		**Panel B: 1990-2000 sample**
VARIABLES		VARIABLES	
region-I X 1990	-0.065*	region-I X 2000	-0.379***
	(0.036)		(0.088)
region-II X 1990	-0.113***	region-II X 2000	-0.340***
	(0.016)		(0.060)
region-III X 1990	0.020	region-III X 2000	-0.469***
	(0.016)		(0.059)
year FE=1990	0.542***	year FE=2000	-0.063
	(0.034)		(0.105)
lg_real_GDP_percapita	0.036	lg_real_GDP_percapita	0.600***
	(0.056)		(0.128)
lg_real_fasset_percapita	-0.013	lg_real_fasset_percapita	-0.354***
	(0.043)		(0.103)
tertiary_teachers_percapita	-0.019***	tertiary_teachers_percapita	-0.012***
	(0.001)		(0.001)
population	-0.017***	population	0.017**
	(0.003)		(0.007)
density	-0.140***	density	0.105*
	(0.053)		(0.058)
Age/Cohort FE	Y	Age/Cohort FE	Y
Provnce FE	Y	Provnce FE	Y
R-squared	0.997	R-squared	0.983
Robustness Check			
VARIABLES			
Post-policy dummy	-0.125***	Post-policy dummy	-0.125***
	(0.011)		(0.042)
region-I X Post-policy dummy	0.048	region-IV X Post-policy dummy	0.467***
	(0.035)		(0.059)
year FE=1990	0.551***	year FE=2000	-0.418***
	(0.032)		(0.062)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

I regress the provincial-year-level log number of young people aged 16-24 on instrumental variables indicating the region-specific change caused by family planning policies. I also control for province fixed effect, time fixed effect, age fixed effect for the young cohort or birth cohort fixed effect for adults, and economic and demographic variables. In the robustness check, the instrumental variables are the post-policy period dummy interacted with regions. Panel A and panel B list the results for the young cohort and older adults for 1982-1990 and 1990-2000 respectively. The first stage results are the same for the young cohort and older adults.

Table 2.3: Impact of provincial supply of young labor on the number (in thousands) of college students/grads among the young cohort

Dependent var: Provincial-year-age-level number (in thousand) of college students or grads

	Panel A: 1982-1990					Panel B: 1990-2000				
	OLS	IV	IV Robust	IV Males	IV Females	OLS	IV	IV Robust	IV Males	IV Females
VARIABLES										
log # of young cohort	16.04	-67.92**	-71.04***	-44.67**	-23.25***	-16.04	-39.38**	-38.45**	-30.69***	-8.69
	(9.82)	(26.99)	(26.73)	(18.85)	(8.68)	(9.78)	(15.57)	(15.67)	(10.46)	(6.25)
year FE=1990	-33.00**	8.89	10.45	2.53	6.36	2.19	-7.07	-6.70	-4.78	-2.29
(or year FE=2000 in Panel B)	(13.26)	(10.24)	(10.18)	(6.66)	(3.92)	(8.71)	(10.09)	(10.19)	(6.57)	(4.42)
lg_real_GDP_percapita	25.73	31.29*	31.50*	22.83*	8.47	78.74***	80.19***	80.13***	43.29***	36.89***
	(17.78)	(18.84)	(18.87)	(13.26)	(5.82)	(16.58)	(16.16)	(16.14)	(10.69)	(6.45)
lg_real_fasset_percapita	-5.77	-10.61	-10.79	-4.97	-5.63**	-29.87***	-29.16***	-29.19***	-16.48**	-12.68***
	(6.30)	(8.01)	(8.09)	(5.33)	(2.78)	(10.87)	(10.80)	(10.80)	(7.03)	(4.71)
tertiary_teachers_percapita	1.01***	-0.53	-0.59	-0.41	-0.12	0.01	-0.23	-0.22	-0.15	-0.08
	(0.27)	(0.50)	(0.50)	(0.34)	(0.17)	(0.19)	(0.23)	(0.23)	(0.15)	(0.10)
population	4.25***	2.93***	2.88***	1.96***	0.97***	-1.19	-1.01	-1.02	-1.49	0.48
	(0.74)	(0.63)	(0.64)	(0.44)	(0.21)	(1.62)	(1.52)	(1.52)	(1.05)	(0.53)
density	48.02**	31.28**	30.66**	25.12**	6.16	-16.33**	-13.77*	-13.87*	-9.56*	-4.21
	(18.66)	(15.42)	(15.40)	(10.77)	(4.94)	(8.16)	(7.73)	(7.73)	(5.18)	(2.94)
Age FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Provnce FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	504	504	504	504	504	504	504	504	504	504
R-squared	0.49	0.46	0.46	0.45	0.48	0.69	0.69	0.69	0.63	0.71

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the number of college students or graduates, in each province in each census year of each one-year age group. The main independent variable is the provincial supply of young labor, defined as the log number of young people aged 16-24 in each province in each census year; it is instrumented by variables indicating the region-specific change in the log number of the young cohort size due to family planning policies. In the IV robustness check, the instrumental variables are the post-policy period dummy interacted with regions. I also control for province fixed effect, time fixed effect, age fixed effect, and economic and demographic variables. Panel A and panel B are the results for 1982-1990 and 1990-2000 respectively.

effect. Population and population density had a positive effect. Panel B (the period 1990-2000) results show no relationship between the log number of the young cohort size and the number of older college graduates. Log real GDP per capita had a greater impact on the number of college graduates in the second period than in the first period. The number of tertiary teachers per capita had an insignificant effect except for the age group 25-30. Other controlled variables show similar effects as in the previous period.

Table 2.5 presents the 2SLS estimates of the effects of the log number of the young cohort size on the number of adult college graduates by age group. In the period 1982-1990, a one percent decrease in the young cohort size on average increased the number of adult college graduates by 67-104 persons within a birth cohort. The effects of the log number of the young cohort size are heterogeneous across age groups, with the smallest effect among the age group 35-39. Hence, the relationship between the log number of the young cohort size and the educational increase appears to be a U-shape with respect to age. In the second period 1990-2000, the effects of the log number of the young cohort size become insignificant.[11] In case the young cohort size affects the change rather than the stock of the number of adult college graduates, I also perform a test on the change of the upgrading. Because of the data limitation I only analyze the change from 1990 to 2000. Specifically, the upgrading in 1990 is calculated as the number of adult college graduates in 1982 subtracted from the number in 1990. The result is listed in Table 2.11 and shows that an increase in the log number of the young cohort size will increase the upgrading in the college attainment for adults 30 to 44 years of age. Combined with the previous result that a smaller young cohort size implies more college educated young people, the result here implies young college educated workers and older college educated workers may be substitutes.

Effects of other characteristics on the adult educational increase are similar to the effects in the OLS estimation for both periods. In the robustness check, using the post-policy variation in the change in the young cohort size, I find that the effects of the log number of the young cohort size are of the same magnitude as before. Comparing with the 2SLS estimates, the OLS estimates of the effects of the log number of the young cohort size on the adult educational increase tend to be biased upward, suggesting factors in the unobservables are likely to be local subsidies for higher education, and/or the nature of young people being more mobile across locations.

I also examine whether the number of non-college graduates changes with respect to the young cohort size. The 2SLS estimates of the effects of the log number of the young cohort size on the number of non-college graduates for all age groups show insignificant effects.[12]

In short, 2SLS regression results suggest that the log number of the young cohort size has a negative causal relationship with the number of adult college graduates. The relationship is strong for all age groups in the period 1982-1990. The effect is insignificant from zero in 1990-2000.

[11] The reason for the insignificant results for the second period may be due to the fact that the migration of the young cohort becomes easier so that the education levels of the young cohort are less affected, which affects older adults to a lesser degree. This is confirmed by the results of the college enrollments in Table 2.3. It shows the impact of the log number of the young cohort size is smaller in 1990-2000 than in 1982-1990. Additionally, this may be due to the fact that older adults have predicted the increasing returns for adult college education, therefore they have gone through the educational upgrading in the first period.

[12] See Table 2.9.

Table 2.4: OLS regression of impact of provincial supply of young labor on the number (in thousands) of adult college grads

Dependent var: Provincial-year-cohort-level number (in thousand) of college grads										
	Panel A: 1982-1990 sample					Panel B: 1990-2000 sample				
	Adutls' Age in 1982					Adults' Age in 1990				
	25-29	30-34	35-39	40-44	45-49	25-29	30-34	35-39	40-44	45-49
VARIABLES										
log # of young cohort	-10.03***	-7.57***	-4.78***	-1.39	-2.17	-3.60	2.09	2.45	1.46	-1.55
	(3.31)	(2.66)	(1.68)	(1.34)	(1.41)	(6.37)	(2.61)	(2.49)	(1.83)	(1.32)
year FE=1990	2.76	3.65**	1.57	-3.57***	-2.60**	9.36	5.94*	0.09	-1.82	-1.44
(or year FE=2000 in Panel B)	(2.30)	(1.72)	(1.10)	(1.19)	(1.15)	(6.51)	(3.55)	(2.87)	(2.41)	(1.59)
lg_real_GDP_percapita	4.07	-0.74	-1.98	5.68***	6.23***	21.47*	22.37***	15.88***	15.46***	4.78*
	(2.70)	(1.99)	(1.52)	(1.51)	(1.45)	(11.58)	(5.72)	(4.65)	(3.42)	(2.58)
lg_real_fasset_percapita	-2.40	-2.34*	-2.11**	-4.24***	-2.52***	-9.41	-8.85**	-3.84	-4.79	-1.03
	(1.88)	(1.33)	(1.01)	(0.78)	(0.86)	(7.96)	(3.91)	(3.64)	(2.94)	(2.00)
tertiary_teachers_percapita	0.40***	0.28***	0.25***	0.31***	0.19***	0.41***	0.09	0.00	0.00	0.03
	(0.09)	(0.06)	(0.04)	(0.05)	(0.05)	(0.12)	(0.07)	(0.05)	(0.07)	(0.03)
population	0.74***	0.53***	0.42***	0.49***	0.24***	1.65***	0.62***	0.70***	0.21	0.31***
	(0.11)	(0.07)	(0.06)	(0.06)	(0.06)	(0.39)	(0.17)	(0.19)	(0.14)	(0.08)
density	1.55	3.64**	4.48***	3.38***	0.99	9.00**	3.98*	1.52	-0.98	1.03
	(1.93)	(1.43)	(0.95)	(0.98)	(0.93)	(4.44)	(2.20)	(2.21)	(1.37)	(0.94)
Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Provnce FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	280	280	280	280	280	280	280	280	280	280
R-squared	0.86	0.88	0.92	0.94	0.91	0.84	0.85	0.85	0.82	0.85

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the number of college graduates in each province in each census year of each one-year birth cohort. The main independent variable is the provincial supply of young labor, defined as the log number of young people aged 16-24 in each province in each census year. I also control for province fixed effect, time fixed effect, birth cohort fixed effect, and economic and demographic variables. Panel A and panel B list the results for each 5-year birth cohort bracket for 1982-1990 and 1990-2000 respectively.

Table 2.5: Two-stage least squares estimates of impact of provincial supply of young labor on the number (in thousands) of adult college grads

Dependent var: Provincial-year-cohort-level number (in thousand) of college grads										
	Panel A: 1982-1990 sample					Panel B: 1990-2000 sample				
	Adults' Age in 1982					Adutls' Age in 1990				
	25-29	30-34	35-39	40-44	45-49	25-29	30-34	35-39	40-44	45-49
VARIABLES										
log # of young cohort	-8.33*	-9.52***	-6.66***	-8.82***	-10.40***	-13.33	-2.64	1.37	1.73	-2.55
	(4.74)	(3.38)	(2.52)	(2.27)	(2.27)	(9.71)	(4.13)	(4.24)	(3.08)	(2.16)
year FE=1990	5.03*	6.24***	2.78**	-0.70	1.13	10.79	7.53**	2.66	-0.41	-1.41
(or year FE=2000 in Panel B)	(2.71)	(2.02)	(1.39)	(1.49)	(1.40)	(6.59)	(3.48)	(3.01)	(2.51)	(1.75)
lg_real_GDP_percapita	-2.51	-3.95*	-2.41	7.90***	7.56***	22.73**	23.09***	16.32***	15.60***	4.90**
	(2.53)	(2.06)	(1.64)	(1.59)	(1.62)	(10.52)	(5.24)	(4.15)	(3.17)	(2.35)
lg_real_fasset_percapita	-3.88**	-3.27**	-2.36**	-4.25***	-2.80***	-18.31**	-14.73***	-9.04***	-7.08**	-1.75
	(1.78)	(1.36)	(0.99)	(0.81)	(0.90)	(7.74)	(3.98)	(3.35)	(2.85)	(1.98)
tertiary_teachers_percapita	0.50***	0.27***	0.23***	0.15**	0.03	0.19	-0.05	-0.08	-0.03	0.01
	(0.10)	(0.07)	(0.05)	(0.06)	(0.06)	(0.16)	(0.08)	(0.06)	(0.07)	(0.04)
population	1.12***	0.69***	0.42***	0.28***	0.07	2.93***	1.45***	1.39***	0.51***	0.42***
	(0.13)	(0.09)	(0.07)	(0.08)	(0.08)	(0.63)	(0.23)	(0.30)	(0.19)	(0.12)
density	-6.25**	-0.94	3.40**	4.08***	0.33	5.12	1.26	-1.18	-2.23*	0.74
	(2.51)	(1.90)	(1.45)	(1.31)	(1.32)	(4.35)	(2.40)	(1.91)	(1.32)	(1.05)
Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Provnce FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	280	280	280	280	280	280	280	280	280	280
R-squared	0.87	0.88	0.92	0.94	0.90	0.84	0.85	0.86	0.82	0.85
Robustness Check										
VARIABLES										
log # of young cohort	-10.56**	-10.64***	-7.23***	-8.73***	-10.14***	-12.84	-1.61	2.48	1.96	-2.36
	(4.64)	(3.39)	(2.51)	(2.26)	(2.28)	(9.68)	(4.13)	(4.27)	(3.09)	(2.14)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the number of college graduates in each province in each census year of each one-year birth cohort. The main independent variable is the log number of young people aged 16-24 in each province in each census year; it is instrumented by variables indicating the region-specific change due to family planning policies. I also control for province fixed effect, time fixed effect, birth cohort fixed effect, and economic and demographic variables. In the robustness check, the instrumental variables are the post-policy period dummy interacted with regions. Panel A and panel B list the results for each 5-year birth cohort bracket for 1982-1990 and 1990-2000 respectively.

2.7.4 Adult Educational Increase: Migration or Upgrading

Having established the negative causal link between the young cohort size and the number of older college graduates, I want to separate the effect of educational upgrading from that of in-migration of college educated labor. The increase in the number of adult college graduates may result from the geographical redistribution of college educated labor. This concern is unlikely to persist, since the decline in the young cohort size becomes prevalent in the entire country. The theory of complementarity between college educated workers of different ages predicts the increased productivity of older college educated workers, which will eventually drive adults to acquire college education.

If the decline in the young cohort size affects a local non-college graduate and a potential college educated migrant equally, the chance that an additional college educated worker is local, rather than a migrant, should be equal to the original composition of college educated workers in that given province. Therefore, the effect of the young cohort size on the migrant share should be positive if it has a stronger effect on the local non-college graduates; the effect should be negative if the college educated migrants are more strongly affected.

To examine the effect of the log number of the young cohort size on the migrant share, I use the econometric model:

$$MigC_{jpt} = \eta ln(YC_{pt}) + \sum_{p} \beta_p Province_p + \gamma_t Year_t + \lambda X_{pt} + \sum_{j} \mu_j Z_j + \varepsilon_{jpt} \tag{2.7.4.7}$$

where $MigC_{jpt}$ is the share of adult college graduates that are migrants in each birth cohort, in each province and each year; $ln(YC_{pt})$ is the log number of the young cohort size in each province and each year. Again $Province_p$, $Year_t$, and Z_j are province fixed effects, time fixed effects, and birth cohort fixed effects and X_{pt} are other characteristics at the province-year level. The endogenous variable $ln(YC_{pt})$ can be potentially correlated with ε_{jpt} and is instrumented by variables $region_1 \times year_t$, $region_2 \times year_t$ and $region_3 \times year_t$ indicating the region-specific change due to family planning policies. I also use the same specification for the share of non-college graduates that are migrants.

$$MigNC_{jpt} = \eta ln(YC_{pt}) + \sum_{p} \beta_p Province_p + \gamma_t Year_t + \lambda X_{pt} + \sum_{j} \mu_j Z_j + \varepsilon_{jpt} \tag{2.7.4.8}$$

where $MigNC_{jpt}$ is the share of non-college graduates that are migrants.

I estimate these specifications using migration information in the 1990 and 2000 census samples[13] and show the results in Table 2.6. The effect on the migrant share is insignificant, implying that the increase in adult college graduates from 1990 to 2000 is most likely to be through the upgrading of local non-college graduates for all age groups[14]. The results for the migrant share of adult non-college graduates show that the log number of the young cohort size has no significant effect on the migrant share[15]. To sum up, the results

[13] Migration information in not available in the 1982 census sample, so I cannot test the migration part for the period 1982-1990. As shown in Figure 1, the migration issue is not important in the first period. Results obtained here will serve as the upper bound for the effect of the log number of the young cohort size on the migration share of college adults in the first period.

[14] One interesting finding is that if I exclude the number of tertiary teachers per capita in the regression, the results show differently, which is, for the age group 25-29, the increase in the number of college graduates is more likely to be driven by the in-migration than by the upgrading. It implies that the channel for the increase in the migrant college educated workers aged 25-29 may be through the expansion of the college resources.

[15] Results are available upon request.

Table 2.6: Two-stage least squares estimates of impact of provincial supply of young labor on share of adult college grads who are migrants

	Dependent var: Provincial-year-cohort-level share of				
	Panel A: 1990-2000 sample				
	Adults' Age in 1990				
	25-29	30-34	35-39	40-44	45-49
VARIABLES					
log # of young cohort	-0.01	-0.02	-0.00	-0.03	-0.04
	(0.03)	(0.02)	(0.02)	(0.04)	(0.02)
year FE=2000	-0.07***	-0.03*	-0.00	0.02	-0.02
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
lg_real_GDP_percapita	0.04	0.02	-0.02	-0.06	-0.01
	(0.04)	(0.02)	(0.02)	(0.06)	(0.03)
lg_real_fasset_percapita	-0.02	-0.00	0.01	0.02	0.01
	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
tertiary_teachers_percapita	0.00***	-0.00	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
population	0.00	-0.00	-0.00***	-0.00**	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
density	-0.02	-0.03**	0.00	0.01	-0.00
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Cohort FE	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y
Observations	280	278	278	272	272
R-squared	0.68	0.42	0.42	0.25	0.51
Robustness Check					
log # of young cohort	-0.02	-0.02	-0.00	-0.03	-0.04
	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the share of adult college graduates that are migrants, in each province in each census year of each one-year birth cohort. The main independent variable is the provincial supply of young labor, defined as the log number of young people aged 16-24 in each province in each census year; it is instrumented by variables indicating the region-specific change due to family planning policies. I also control for province fixed effect, time fixed effect, birth cohort fixed effect, and economic and demographic variables. In the robustness check, the instrumental variables are the post-policy period dummy interacted with regions. Panel A lists the results for each 5-year birth cohort bracket for 1990-2000.

show that the increase in the number of adult college graduates resulting from the decline in the young cohort size is mostly driven by the upgrading from the local non-college graduates for all adult age groups.

2.7.5 Adult Educational Upgrading by Gender

In the previous subsection, I show the effect of the log number of the young cohort size on the college enrollments by gender for young people. I find that the effect of the log number of the young cohort size on the college enrollments is larger for male youth than for female youth. In the second period, the effect for female youth is essentially insignificantly different from zero.

Here, I show the effect of the log number of the young cohort size on the adult educational upgrading by gender. In Table 2.10, I use the region-specific change in the log number of the young cohort size due to family planning policies as an exogenous source of the variation in the young cohort size; I find the effect of the log number of the young cohort size on the number of college graduates is greater for men than for women in 1982-1990. For men, it shows a similar U-shaped pattern of the relationship between the two variables of interest with respect to age. Men aged 40-49 had a greater increase in the number of college graduates than men aged 35-39 when the young cohort size decreased. For women, this U-shape feature is somewhat weaker; women aged 45-49 had a slightly greater impact than women aged 40-44.

In the second period 1990-2000, the effect of the log number of the young cohort size seemed to be insignificant for all age groups for both men and women except women aged 35-39 had a significant increase in the number of college graduates. In the robustness check, the effects of the log number of the young cohort size on the number of college graduates remain the same for both periods and for both gender. The U-shaped pattern of the relationship between the young cohort size and the adult educational upgrading still exists.

2.7.6 Summary and Discussion

In the section of results, I have shown the family planning policies that started from 1964 and were implemented through the 1970s significantly reduced the size of young cohort aged between 16 and 24. Specifically, the program "Later Longer Fewer" combined with access to birth control methods reduced the size of the young cohort aged 16-24 by 13 percent. The reduction in the young cohort size resulting from the policies increased the college enrollments of the young cohort in both periods 1982-1990 and 1990-2000. The reduction in the young cohort size also increased the number of college graduates among adults aged 25-49 in the period 1982-1990. I provide some evidence that the increase in the adult college graduates is mostly likely to be driven by the educational upgrading from non-college graduates in the local labor markets.

The results on the college enrollment of the young cohorts and the educational upgrading of older adults together suggest that college educated workers of different ages may be complements in the production function. The reason is that if we assume college educated workers of different ages are perfect substitutes, the increase in the college graduates should reflect the slope of the supply curve of college education for each age group. Given that the cost of going back to college may be higher for older people and the lifelong returns are lower for older people, the slope of the supply curve of college education should be steeper for older adults than for younger adults. Nevertheless, I obtain a U-shaped pattern in the relationship between the young cohort size and the number of adult college graduates. This suggests college educated workers of age 40 and above may be complements to young college educated workers aged 16-24.

The demand for college education from adults is based on the premise that the family planning polices reduce the number of young people and increase the number of college educated workers among young people. The increase in the supply of young college educated workers can lead to an increase in the productivity of older college educated workers and therefore creates incentives for older people to acquire a college degree. Although a smaller family size resulting from the fertility policies can relax the budget constraint and lead to a greater supply of adult college educated workers within a family, the incentives for older workers to upgrade will not be lessened by this supply shift. By controlling for log real GDP per capita at the province-year level, and the average number of alive children per woman at the province-birth-cohort level, I find roughly the same effect on the educational upgrading for women as before.

2.8 Conclusion

In this paper, I examine the effect of the reduction in the young cohort size caused by family planning policies in China on the educational upgrading of older adults born before the policies. The log number of the young cohort size decreases substantially once a province is treated by the policies. I find the decline in the young cohort size leads to educational improvement for the young cohort themselves and the educational upgrading for adults. A one percent decrease in the size of the young cohort increases the number of college students or graduates by about 700 persons within an age group among themselves during 1982-1990 and by about 400 persons during 1990-2000; a one percent decrease in the size of the young cohort increases the number of college graduates by 70-100 persons within a birth cohort among those aged 25-49 during 1982-1990. The effect is greater for adults aged above 40 than adults aged 35-39. It suggests the complementarity of college educated workers of different ages in the production function.

I further find that the increase in the number of older college graduates resulting from the reduction in the young cohort size is mostly driven by the educational upgrading from non-college graduates who have already been in the local labor markets. The results are consistent with the assumption that older college educated workers have higher costs of migration or a distaste for migration. In addition, I find the decrease in the young cohort size does not affect the number of non-college graduates for older adults. This means the in-migration of non-college educated migrants will offset the reduction in the local non-college graduates who have upgraded their educational levels.

One special finding is that the reduction in the young cohort size resulting from the fertility policies does not have a significant impact on the number of adult college graduates in the period 1990-2000[16]. It might be due to the weaker migration restrictions in the 1990s which leads to more equalized number of young college educated workers across regions. Therefore, the productivity of older workers would not be affected differently across regions. To gain a better understanding of this, future studies could look into the wage returns for the age groups.

[16]Could the adult educational upgrading be driven by the schooling re-investment of cohorts affected by the Cultural Revolution? It is unlikely. Previous research (Han, Suen and Zhang 2011) has found that the birth cohorts that were interrupted in the educational progress were born between 1947 and 1961. These birth cohorts are 21 to 35 years of age in 1982. But I have found strong effects for those above age 35, which is not consistent with this confounding hypothesis. Moreover, for my results to be explained by the Cultural Revolution, the geographical pattern of the interruption generated by the Cultural Revolution should coincide with the geographical variation in the fertility policies. This is unlikely to be true because the Cultural Revolution affected the birth cohorts the same across regions.

2.9 Appendix

2.9.1 Census data in 1982, 1990 and 2000

The samples I use are from China's national population censuses in 1982, 1990 and 2000. The data are sampled at the household level, either domestic household or collective household living unit, with a sampling rate of 1%, 1% and 0.095%. The data are representative samples of all individuals who have Chinese nationality and live in China.

To examine the effect of the young cohort size on the educational upgrading, I need individual information below to construct the educational composition within each birth-cohort-province-year cell: the number of persons with a college or advanced education (called college graduates) and the number of persons with a high school or lower education (called non-college graduates). The information includes:

1. age: age by July 1st of 1982 (only for 1982 census)
2. year of birth and/or month of birth: available in 1990 and 2000 censuses
3. residential place: the current residential province or county

The provinces in censuses include the municipalities, provinces, autonomous regions and special administrative regions in China. They are Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong (including Hainan), Guangxi, Sichuan (including Chongqing), Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang, Hongkong and Macau. Tibet, Hongkong and Macau had not implemented the family planning polices in the period of my study, so they are excluded in my samples.

4. educational level: a discrete variable categorized into six groups: illiterate or semi-literate, primary school, junior high school (7th - 9th grade), senior high school (10th - 12th grade), college or semi-college, or graduate education.
5. educational status: a discrete variable categorized into four groups: graduate, undergraduate, drop-out, or else.
6. adult education: adult education degree (only available in 2000 census)

By tracking the individuals within the birth-cohort-province-year cell over time, I can identify the increase in the number of college graduates over time. One problem about the increase is whether it is because of measurement error in education. Information on (6) adult education helps to illustrate how much of the increase could be potentially caused by measurement error. This variable provides a second way to count the amount of college graduates who have acquired the college degrees in their adulthood. This amount matches with 67% of the increase calculated from tracking individuals within birth cohort.

I construct the migration percentage for each educational and demographic group, i.e., the fraction of the college graduates or the non-college graduates that are immigrants in each birth-cohort-province-year cell. To identify the migration status, I use the migration information from census data:

7. previous residential place: residential province five years ago (only available in 1990 and 2000 censuses)

The migration share within an education-age group helps to decompose the increase in the group size into a part that is because of migration and a part that is because of the actual upgrading from lower education. If the migration part is more important in response to a decreased young cohort size, we will expect the migrant share within the group to increase.

However, identifying the migration based on a five-year-span comparison is likely to overestimate the part of the increase that is because of the upgrading. For example, a college-educated person had moved to a place six years ago. He is counted in the part of upgrading rather than migration. It can underestimate it as well. If a person was living in another province five years ago temporarily, but was residing in the current province for most of his life, then he should have been counted in the part of upgrading rather than migration had he acquired a college education. Both directions of bias are likely to be mitigated in the cross-region comparison of the changes in the migrant share over time, unless the timing of the migration induced by the young labor shortage, changes dramatically at the cutoff (i.e. five years ago) over time and across region. This kind of bias emerging in the cross-regional comparison of changes over time seems to be less likely.

Another subtle issue is that I cannot know the exact educational level upon one's arrival. They could be college graduates upon arrival or non-college upon arrival but upgrade to college graduates afterwards. Such two scenarios are treated the same as being the contribution of migration to the increase in the number of the college graduates.

2.9.2 Other Yearly Provincial Macro-economy Statistics

GDP per capita, fixed asset investment per capita, the number of tertiary teachers per 100,000 people, population are collected from China Yearly Provincial Statistical Yearbooks.

2.9.3 Note on Figure 2.3

Essentially, counting the number of people aged 16-24 in a location (i.e., a province, a region, or the whole country) in a year t is to count the number of people born in $t-24$ to $t-16$ who live in that location in that year. To get the number of young people aged 16-24 living in a region in the year t, I add up the number of persons born between the year $t-24$ and $t-16$ who were sampled to be living in the region in the 1982 census sample. This produces the dot line in Figure 1. If we assume people do not move across regions, this number is an accurate estimate of the number of young people for any year t. However, if people move across the regions, the more further away a year is from 1982, the less accurate this way of estimating the number of young people *for the year* will be. For each region, I produce the pattern of the youth size computed from the census sample of 1982, 1990 and 2000 respectively. Normally for a year t, where $t <= 1982$, the estimate of the youth size would be more accurate if it is computed from the sample of 1982 than from the sample of 1990 or 2000. For a year t, where $1982 < t <= 1990$, the estimate of the youth size would be more accurate from 1982 or 1990 than from 2000, because given the size of them (i.e., the size of those born between $t-24$ and $t-16$) is precisely computed for 1982, 1990 and 2000, it is less likely for the size to drop to the lower level predicted by the 2000 sample in a year, say 1985, and rise back to the level predicted by the 1990 sample in 1990, and drop again to the lower level predicted by the 2000 sample in 2000 (see the subfigure named as Region-III(FPP: 1974) in Figure 1). Likewise, for a year t, where $1990 < t <= 2000$, the estimate of the youth size would be more accurate from 1990 or 2000 than from 1982.

For each region, Figure 1 presents the pattern of the youth size computed from the census sample in 1982, 1990 and 2000 respectively. Instead of presenting the youth population by year, it presents the ratio of the youth population in each region to the national youth population. The latter, the national youth population in a year t, can be calculated by adding up the number of young people born in $t-24$ to $t-16$ in the whole country (precisely, the whole set of provinces included in my study) from the census sample of 1982, 1990 and 2000 respectively. The ratio illustrates the region-specific evolution of the youth population more clearly than the absolute number does.

2.9.4 Note on Figure 2.4

As described in the previous subsection, the method to calculate the size of young cohort aged 16-24 in a province in a year t is to count the number of persons born between the year $t-24$ and $t-16$ who live in the province in the year t from the census sample of the year t. In this way, I can calculate the size of young cohort in a province in the year of 1982, 1990 and 2000. Additionally, the census sample of 1990 and 2000 asks which province a person lived in five year ago, so I can calculate the size of young cohort in a province in the year of 1985 and 1995 as well.

Next, using the size of young cohort in the year of 1982, 1985, 1990, 1995 and 2000, I calculate the young cohort size in a given year t between these years as the weighted sum of the sizes in two closest years. The weight is inversely proportional to the time distance to two years. For example, the young cohort size in 1984 is the weighted sum of the size in 1982 and the size in 1985. The weight is two thirds to the year of 1982 and one third to the year of 1985. For any year t prior to 1982, I count the number of persons born between the year $t-24$ and $t-16$ who live in the province in 1982 from the 1982 census. Here the assumption is that people did not move across provinces prior to 1982, which is probably realistic for this period. Then I convert the size of young cohort aged 16-24 into the log number of the size of young cohort for each province in each year. In the same way, I also construct the log number of the size of young cohort at the national level from 1970-2000.

Using the constructed natural logarithm of the size of young cohort $ln(YC_{pt})$ at the level of province and year, I create a demeaned series of the log size of young cohort by subtracting the national level from it. I regress the demeaned log size of young cohort $ln(dmYC_{pt})$ on a set of year fixed effects, a set of region dummies interacted with year fixed effects, a set of province fixed effects, and economic and demographic variables such as log real GDP per capita and log real fixed asset investment per capita, the number of tertiary teachers per capita, population and population density.

$$\begin{aligned} ln(dmYC_{pt}) = & \sum_{t=1971}^{2000} \alpha_{1t} Region_1 \times Year_t \\ & + \sum_{t=1971}^{2000} \alpha_{2t} Region_2 \times Year_t \\ & + \sum_{t=1971}^{2000} \alpha_{3t} Region_3 \times Year_t \\ & + \sum_{t=1971}^{2000} \gamma_t Year_t + \sum_{p=2}^{28} \beta_p Province_p + \lambda X_{pt} + \varepsilon_{pt} \end{aligned} \tag{2.9.4.9}$$

where,

$ln(dmYC_{pt})$	$=ln(YC_{pt}) - ln(YC_t)$
$ln(YC_{pt})$	=the log number of young cohort aged 16-24 at the province and year level
$ln(YC_t)$	=the log number of young cohort aged 16-24 at the national and year level
$Region_1$	=the indicator for region-I
$Region_2$	=the indicator for region-II
$Region_3$	=the indicator for region-III
$Year_t$	=the indicator for each year t
$Province_p$	=the indicator for each province
X_{pt}	=a set of economics and demographic variables
ε_{pt}	=error term for the demeaned log number of young cohort at the province and year level

I plot the coefficients on the year dummies γ_t in the subfigure Region-IV (FPP: 1978), and the coefficients on region dummies interacted with year fixed effects α_{1t} in the subfigure Region-I relative to Region-IV (FPP: 1964), α_{2t} in the subfigure Region-II relative to Region-IV (FPP: 1971), α_{3t} in the subfigure Region-III relative to Region-IV (FPP: 1974) in Figure 2.4.

2.9.5 Note on Figure 2.5 and Figure 2.6

Figure 2.5 illustrates the causal relationship between the size of young population and the educational improvement of young people. Figure 2.6 illustrates the causal relationship between the size of young population and the educational upgrading of older adults. To produce the figures, I construct the time difference of all the variables within a province p and an age group j for young people (or within a province p and a birth cohort j for older adults), then I run the following regression:

$$dYC_p = \alpha_1 \times Region_1 + \alpha_2 \times Region_2 + \alpha_3 \times Region_3 + \beta \times dX_p + \varepsilon_p$$

Predict residual from the above equation: $\widehat{\varepsilon_p}$, then run the following regressions:

$$dYC_p = \gamma \times \widehat{\varepsilon_p} + \lambda \times dX_p + \eta_p$$
$$dTC_{jp} = \delta \times \widehat{\varepsilon_p} + \tau \times dX_p + \mu_{jp}$$

where,

dYC_p =change in the size of young cohort aged 16-24 over time within a province

$Region_1$ =the indicator for region-I

$Region_2$ =the indicator for region-II

$Region_3$ =the indicator for region-III

dX_p =change in the economic and demographic variables over time within a province

dTC_{jp} =change in the total number of college students and graduates over time within a province, if j indexes an age group of the young cohort

=change in the total number of college graduates over time within a province, if j indexes a birth cohort of the older adults

Plot $\widehat{\mu_{jp}}$ against $\widehat{\eta_p}$ by time period in Figure 2.5 for the young cohort. Plot $\widehat{\mu_{jp}}$ against $\widehat{\eta_p}$ by adult birth cohort group and by time period in Figure 2.6 for older adults..

2.9.6 Additional Statistics

Table 2.7: Summary Statistics: National and by Region

Sample Statistics: National and Region I

	1982-90 sample			1990-2000 sample		
	1982	1990	Growth rate	1990	2000	Growth rate
National						
number of provinces	28			28		
# of longitudinal young age groups in each province	9			9		
# of longitudinal older birth cohorts in each province	25			25		
young people: age range	16-24			16-24		
young people: year of birth	1958-1966	1966-1974		1966-1974	1976-1984	
older adults: age range	25-49	33-57		25-49	35-59	
older adults: year of birth range	1933-1957			1941-1965		
popualtion of young labor, age 16-24, 1 million (n=28)	6.3 [3.8]	8.12 [5.25]		8.12 [5.34]	5.83 [3.8]	
log population of young labor, age 16-24 (n=28)	15.41 [0.8]	15.65 [0.81]	0.24	15.65 [0.81]	15.35 [0.74]	-0.3
young population with coll enrollment/degree, 1 thousand	4.99 [4.24]	20.55 [32.78]		20.55 [32.78]	47.45 [37.87]	
log young population with coll enrollment/degree	7.98 [1.41]	8.32 [2.71]	0.34	8.32 [2.71]	9.99 [2.23]	1.67
migrant share of young coll students/degree holders (avg. of all age groups (n=239))*				0.14 [0.24]	0.04 [0.05]	
older population with coll degree, 1 thousand	4.92 [3.13]	8.96 [5.93]		12.24 [9.66]	22.04 [19.39]	
log older population with coll degree	8.26 [0.83]	8.79 [0.96]	0.53	9.07 [0.95]	9.51 [1.48]	0.44
migrant share of older coll degree holders (avg. of all birth cohorts (n=690))*				0.04 [0.05]	0.01 [0.04]	
older population with non-coll degree, 1 thousand	444.1 [335.92	435.5 [321.33]		549.4 [398.91	531.1 [395.06]	
log older population with non-coll degree	12.69 [0.88]	12.69 [0.85]	0	12.92 [0.86]	12.87 [0.88]	-0.05
migrant share of older non-coll degree holders (avg. of all birth cohorts (n=700))				0.01 [0.02]	0.01 [0.02]	
log real GDP per capita 1982RMB (n=28)	5.99 [0.56]	6.45 [0.51]		6.45 [0.51]	7.18 [0.6]	
log real fixed asset per capita 1982RMB (n=28)	4.5 [0.69]	4.99 [0.67]		4.99 [0.67]	6.14 [0.66]	
tertiary teachers per 100,000 persons (n=28)	44.27 [62.38]	52.79 [68.29]		52.79 [68.29]	52.54 [58.85]	
population, 1 million persons (n=28)	35.51 [23.22]	39.82 [25.81]		39.82 [25.81]	43.44 [28.39]	
population density, 100 persons per km2 (n=28)	1.16 [0.65]	1.34 [0.76]		1.34 [0.76]	1.25 [0.91]	
Regional (family planning program regions)						
Region-I, number of provinces (*Beijing, Shanghai, Tianjin and Liaoning*)	4			4		
popualtion of young labor, age 16-24	3.25 [2.66]	3.24 [2.58]		3.24 [2.58]	2.91 [1.87]	
log population of young labor, age 16-24	14.79 [0.68]	14.81 [0.65]	0.02	14.81 [0.65]	14.75 [0.58]	-0.06
young population with coll enrollment/degree	7.53 [4.88]	29.15 [22.91]		29.15 [22.91]	57.46 [36.95]	
log young population with coll enrollment/degree	8.53 [1.14]	9.02 [3.07]	0.49	9.02 [3.07]	10.39 [2.01]	1.37
migrant share of young coll students/degree holders				0.51 [0.34]	0.09 [0.08]	
older population with coll degree	7.59 [3.81]	16.37 [6.78]		21.06 [10.88]	30.04 [18.7]	
log older population with coll degree	8.8 [0.54]	9.6 [0.48]	0.8	9.84 [0.49]	10.12 [0.63]	0.28
migrant share of older coll degree holders				0.06 [0.07]	0.02 [0.04]	
older population with non-coll degree	222 [180.02	238.4 [174.78]		310 [229.18	292.8 [233.65]	
log older population with non-coll degree	12.06 [0.67]	12.18 [0.61]	0.12	12.43 [0.64]	12.33 [0.7]	-0.1
migrant share of older non-coll degree holders				0.04 [0.03]	0.04 [0.04]	
log real GDP per capita 1982RMB	7.01 [0.62]	7.28 [0.45]		7.28 [0.45]	8 [0.51]	
log real fixed asset per capita 1982RMB	5.6 [0.61]	6.05 [0.55]		6.05 [0.55]	7.04 [0.68]	
tertiary teachers per 100,000 persons	163.6 [115.7]	184.3 [123.9]		184.3 [123.9]	161.2 [107.71]	
population, 1 million persons	16.16 [13.28]	17.75 [14.39]		17.75 [14.39]	18.72 [15.18]	
population density, 100 persons per km2	2.04 [1]	2.4 [1.19]		2.4 [1.19]	2.58 [1.42]	

Sample statistics are computed from China census samples in 1982, 1990 and 2000, and from provincial yearbook data. All the provincial statistics are at the level of province and year; all the birth cohort statistics are at the level of province and year and one-year birth cohort. Sample standard deviations are in brackets. 1982-1990 and 1990-2000 are sub-samples following adult birth cohorts over time. They are birth cohorts aged 25-49 in 1982 and in 1990 respectively. I divide the provinces into four regions based on the timing of the initiation of family planning policies in China. Certain young age groups do not have any young college students or graduates; certain older birth cohorts do not have any college educated adults.

Sample Statistics: Regions II, III, IV

	1982-90 sample			1990-2000 sample		
	1982	1990	Growth rate	1990	2000	Growth rate
Region-II, number of provinces	12			12		
(*Fujian, Gansu, Guangdong, Hebei, Heilongjiang, Hunan, Inner Mongolia, Jiangsu, Shaanxi, Shanxi, Shandong and Zhejiang*)						
popualtion of young labor, age 16-24	7.44 [3.2]	9.1 [3.76]		9.1 [3.76]	7.14 [4.22]	
log population of young labor, age 16-24	15.73 [0.45]	15.95 [0.41]	0.22	15.95 [0.41]	15.64 [0.53]	-0.31
young population with coll enrollment/degree	5.25 [3.88]	24.7 [43.03]		24.7 [43.03]	52.62 [38.06]	
log young population with coll enrollment/degree	8.18 [1.22]	8.54 [2.55]	0.36	8.54 [2.55]	10.38 [1.5]	1.84
migrant share of young coll students/degree holders				0.08 [0.18]	0.04 [0.05]	
older population with coll degree	5.05 [2.39]	8.86 [4.06]		5.05 [2.39]	8.86 [4.06]	
log older population with coll degree	8.42 [0.48]	8.97 [0.53]	0.55	8.42 [0.48]	8.97 [0.53]	0.55
migrant share of older coll degree holders				0.04 [0.04]	0.02 [0.05]	
older population with non-coll degree	515.4 [294.01	504.9 [284.26]		515.4 [294.01	504.9 [284.26]	
log older population with non-coll degree	13 [0.55]	12.98 [0.54]	-0.02	13 [0.55]	12.98 [0.54]	-0.07
migrant share of older non-coll degree holders				0.01 [0.005]	0.01 [0.01]	
log real GDP per capita 1982RMB	5.92 [0.26]	6.42 [0.31]		6.42 [0.31]	7.23 [0.51]	
log real fixed asset per capita 1982RMB	4.43 [0.34]	4.95 [0.39]		4.95 [0.39]	6.14 [0.53]	
tertiary teachers per 100,000 persons	25.45 [9.15]	32 [11.15]		32 [11.15]	35.75 [9.5]	
population, 1 million persons	40.04 [18.96]	44.93 [21.34]		44.93 [21.34]	49 [23.12]	
population density, 100 persons per km2	0.97 [0.55]	1.13 [0.64]		1.13 [0.64]	1.18 [0.76]	
Region-III, number of provinces	7			7		
(*Anhui, Henan, Hubei, Jiangxi, Jilin, Sichuan and Xinjiang*)						
popualtion of young labor, age 16-24	8.14 [4.57]	11.55 [7.1]		11.55 [7.1]	6.69 [3.5]	
log population of young labor, age 16-24	15.75 [0.66]	16.07 [0.7]	0.32	16.07 [0.7]	15.6 [0.52]	-0.47
young population with coll enrollment/degree	5.42 [4.47]	20.53 [23.73]		20.53 [23.73]	55.22 [38.84]	
log young population with coll enrollment/degree	8.22 [1.01]	8.69 [2.42]	0.47	8.69 [2.42]	10.55 [1.04]	1.86
migrant share of young coll students/degree holders				0.09 [0.13]	0.02 [0.02]	
older population with coll degree	5.33 [3.05]	9.33 [4.87]		13.05 [8.52]	24.25 [19.61]	
log older population with coll degree	8.42 [0.59]	9 [0.55]	0.58	9.31 [0.57]	9.81 [0.8]	0.5
migrant share of older coll degree holders				0.02 [0.02]	0 [0.02]	
older population with non-coll degree	597.1 [415.99	574.9 [397.03]		713.8 [484.01	681.3 [458.48]	
log older population with non-coll degree	13.05 [0.73]	13.02 [0.73]	-0.03	13.24 [0.72]	13.2 [0.72]	-0.04
migrant share of older non-coll degree holders				0.01 [0.01]	0.01 [0.02]	
log real GDP per capita 1982RMB	5.84 [0.23]	6.34 [0.26]		6.34 [0.26]	7.07 [0.25]	
log real fixed asset per capita 1982RMB	4.24 [0.43]	4.75 [0.47]		4.75 [0.47]	5.91 [0.38]	
tertiary teachers per 100,000 persons	26.35 [12.01]	33.62 [16.8]		33.62 [16.8]	36.74 [17.48]	
population, 1 million persons	48.97 [30.37]	54.78 [33.21]		54.78 [33.21]	59.75 [35.4]	
population density, 100 persons per km2	1.05 [0.42]	1.17 [0.45]		1.17 [0.45]	0.97 [0.37]	
Region-IV, baseline region, number of provinces	5			5		
(*Guangxi, Guizhou, Ningxia, Qinghai and Yunnan*)						
popualtion of young labor, age 16-24	3.44 [2.62]	4.9 [3.59]		4.9 [3.59]	3.85 [2.82]	
log population of young labor, age 16-24	14.64 [1.14]	15.02 [1.11]	0.38	15.02 [1.11]	14.8 [1.08]	-0.22
young population with coll enrollment/degree	1.77 [1.58]	3.72 [4.65]		3.72 [4.65]	16.15 [14.5]	
log young population with coll enrollment/degree	6.75 [1.84]	6.72 [2.63]	-0.03	6.72 [2.63]	7.95 [3.6]	1.23
migrant share of young coll students/degree holders				0.05 [0.07]	0.03 [0.05]	
older population with coll degree	1.91 [1.3]	2.75 [2.29]		3.89 [3.8]	7.81 [8.07]	
log older population with coll degree	7.2 [1.08]	7.41 [1.16]	0.21	7.77 [1.12]	7.93 [2.52]	0.16
migrant share of older coll degree holders				0.02 [0.04]	0.01 [0.02]	
older population with non-coll degree	236.3 [174.2]	231.2 [170.78]		287.4 [227.39	276.4 [216.39]	
log older population with non-coll degree	11.93 [1.07]	11.91 [1.07]	-0.02	12.12 [1.07]	12.07 [1.09]	-0.05
migrant share of older non-coll degree holders				0.01 [0.01]	0.01 [0.01]	
log real GDP per capita 1982RMB	5.56 [0.53]	6.05 [0.57]		6.05 [0.57]	6.57 [0.54]	
log real fixed asset per capita 1982RMB	4.13 [0.97]	4.54 [0.84]		4.54 [0.84]	5.73 [0.76]	
tertiary teachers per 100,000 persons	19.07 [7.54]	24.34 [8.98]		24.34 [8.98]	27.97 [10.67]	
population, 1 million persons	21.26 [16.07]	24.31 [18.35]		24.31 [18.35]	27.02 [20.29]	
population density, 100 persons per km2	1.05 [0.46]	1.22 [0.51]		1.22 [0.51]	0.72 [0.23]	

Sample statistics are computed from China census samples in 1982, 1990 and 2000, and from provincial yearbook data. All the provincial statistics are at the level of province and year; all the birth cohort statistics are at the level of province and year and one-year birth cohort. Sample standard deviations are in brackets. 1982-1990 and 1990-2000 are sub-samples following adult birth cohorts over time. They are birth cohorts aged 25-49 in 1982 and in 1990 respectively. I divide the provinces into four regions based on the timing of the initiation of family planning policies in China. Certain young age groups do not have any young college students or graduates; certain older birth cohorts do not have any college educated adults.

Table 2.8: Impact of provincial supply of young labor on share of college students/grads who are migrants for the young cohort

Dependent var: Provincial-year-age-level share of migrants among young college			
	Panel A: 1990-2000		
	OLS	IV	IV Robust
VARIABLES			
log # of young cohort	0.22***	0.33***	0.31***
	(0.04)	(0.06)	(0.06)
year FE=2000	-0.04	0.00	-0.00
	(0.04)	(0.05)	(0.04)
lg_real_GDP_percapita	0.10	0.09	0.09
	(0.08)	(0.08)	(0.08)
lg_real_fasset_percapita	-0.07	-0.07	-0.07
	(0.05)	(0.05)	(0.05)
tertiary_teachers_percapita	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
population	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
density	-0.00	-0.02	-0.01
	(0.03)	(0.03)	(0.03)
Age effect	Y	Y	Y
Province FE	Y	Y	Y
Observations	478	478	478
R-squared	0.71	0.71	0.71

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the share of college students or graduates that are migrants, in each province in each census year of each one-year age group. The log number of young people aged 16-24 is instrumented by variables indicating the region-specific change due to family planning policies. In the robustness check, it is instrumented by the post-policy period dummy interacted with regions. Results are available only for 1990-2000.

Table 2.9: Two-stage least squares regression of impact of provincial supply of young labor on the number (in thousands) of adult non-college grads

	Dependent var: Provincial-year-cohort-level number (in thousand) of non-college grads									
	Panel A: 1982-1990 sample					Panel B: 1990-2000 sample				
	Adults' Age in 1982					Adults' Age in 1990				
	25-29	30-34	35-39	40-44	45-49	25-29	30-34	35-39	40-44	45-49
VARIABLES										
log # of young cohort	-20.29	-9.30	-51.65	-44.97	-24.82	212.63	117.93	41.21	35.76	8.58
	(87.66)	(106.42)	(89.77)	(55.07)	(42.16)	(200.35)	(99.41)	(55.88)	(41.29)	(29.02)
year FE=1990	19.39	14.58	22.14	17.25	8.11	60.06	28.24	-9.14	5.89	-11.31
(or year FE=2000 in Panel B)	(36.06)	(49.55)	(48.58)	(29.78)	(19.82)	(125.45)	(59.15)	(34.90)	(32.61)	(20.81)
lg_real_GDP_percapita	-9.04	-0.10	18.77	21.08	14.84	22.53	2.66	-6.67	-14.08	-0.77
	(65.90)	(71.50)	(43.25)	(34.00)	(32.99)	(177.85)	(79.18)	(46.51)	(47.83)	(29.43)
lg_real_fasset_percapita	6.24	2.18	-11.78	-11.27	-4.89	-34.83	-19.10	0.32	-4.97	-2.53
	(30.05)	(37.39)	(34.70)	(15.98)	(14.44)	(133.61)	(62.25)	(34.52)	(36.21)	(24.33)
tertiary_teachers_percapita	-1.47	-0.77	-1.22	-1.09	-0.84	2.33	1.20	0.22	0.42	0.32
	(1.57)	(2.02)	(1.68)	(1.18)	(0.81)	(3.19)	(1.44)	(0.84)	(0.66)	(0.39)
population	-7.31**	-5.69	-5.34	-5.52**	-7.07***	-1.16	2.00	2.64	3.02	0.35
	(3.44)	(4.16)	(3.59)	(2.29)	(1.69)	(11.17)	(5.55)	(3.33)	(2.87)	(1.69)
density	124.95**	89.64	75.67**	32.19***	71.21**	7.59	-8.49	-11.59	-5.19	-0.86
	(59.87)	(60.95)	(36.21)	(31.07)	(28.18)	(65.52)	(28.63)	(18.04)	(20.14)	(10.74)
Cohort FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Provnce FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	280	280	280	280	280	280	280	280	280	280
R-squared	0.99	0.98	0.98	0.99	0.99	0.89	0.95	0.99	0.98	0.98

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the number of non-college graduates (including uncompleted college education, high school graduates or lower education levels) in each province in each census year of each one-year birth cohort. The log number of young people aged 16-24 is instrumented by variables indicating the region-specific change due to family planning policies.

Table 2.10: Impact of provincial supply of young labor on the number (in thousands) of adult college grads by gender

	Dependent var: Provincial-year-cohort-level number (in thousand) of college grads									
	Panel A: 1982-1990 sample					Panel B: 1990-2000 sample				
	Adults' Age in 1982					Adults' Age in 1990				
	25-29	30-34	35-39	40-44	45-49	25-29	30-34	35-39	40-44	45-49
All										
log # of young cohort	-8.33*	-9.52***	-6.66***	-8.82***	10.40***	-13.33	-2.64	1.37	1.73	-2.55
	(4.74)	(3.38)	(2.52)	(2.27)	(2.27)	(9.71)	(4.13)	(4.24)	(3.08)	(2.16)
Males										
log # of young cohort	-3.99	-5.55**	-4.20**	-6.54***	-7.99***	-5.53	-2.98	5.31*	1.53	-0.99
	(3.33)	(2.62)	(1.92)	(1.84)	(1.90)	(7.06)	(2.97)	(3.23)	(2.52)	(1.74)
Females										
log # of young cohort	-4.34**	-3.97***	-2.47**	-2.29***	-2.41***	-7.80*	0.34	-3.94**	0.20	-1.56
	(1.85)	(1.25)	(1.01)	(0.87)	(0.75)	(4.32)	(1.88)	(1.64)	(1.35)	(1.23)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regressions use gender-specific samples. The dependent variable is the number of college graduates in each province in each census year of each one-year birth cohort for a specific gender. The log number of young people aged 16-24 is instrumented by variables indicating the region-specific change due to family planning policies.

Table 2.11: Impact of provincial supply of young labor on the number (in thousands) of the change of adult college grads

Dependent var: Provincial-year-cohort-level number (in thousand) of the change in college grads					
	Panel B: 1990-2000 sample				
	Adults' Age in 1990				
	25-29	30-34	35-39	40-44	45-49
VARIABLES					
log # of young cohort	10.04	11.95***	13.34***	8.02**	-0.57
	(8.10)	(4.20)	(4.51)	(3.16)	(2.51)
year FE=2000	3.10	2.71	-1.60	-2.50	-3.03
	(5.37)	(3.28)	(2.97)	(2.43)	(1.98)
lg_real_GDP_percapita	-6.41	7.88*	4.77	6.52**	0.03
	(8.18)	(4.07)	(3.93)	(3.20)	(2.67)
lg_real_fasset_percapita	-0.83	-3.54	0.24	-1.50	1.36
	(6.03)	(3.18)	(2.90)	(2.69)	(2.27)
tertiary_teachers_percapita	0.81***	0.24***	0.16**	0.20***	0.15***
	(0.14)	(0.07)	(0.07)	(0.06)	(0.05)
population	1.28***	0.99***	1.11***	0.31*	0.30**
	(0.47)	(0.21)	(0.30)	(0.19)	(0.13)
density	2.83	0.86	-2.18	-1.92	0.60
	(3.46)	(1.84)	(1.82)	(1.35)	(1.15)
Cohort FE	Y	Y	Y	Y	Y
Provnce FE	Y	Y	Y	Y	Y
Observations	280	280	280	280	280
R-squared	0.76	0.71	0.66	0.55	0.42

Bibliography

Angeles, Gustavo, David K Guilkey, and Thomas A Mroz, "The effects of education and family planning programs on fertility in Indonesia," *Economic Development and Cultural Change*, 2005, *54* (1), 165–201.

Attane, I., "China's family planning policy: an overview of its past and future," *Studies in Family Planning*, 2002, *33* (1), 103–113.

Autor, D. and L. Katz, "Changes in the wage structure and earnings inequality," *Handbook of labor economics*, 1999, *3*, 1463–1555.

Bailey, Martha J, "More power to the pill: the impact of contraceptive freedom on women's life cycle labor supply," *The Quarterly Journal of Economics*, 2006, *121* (1), 289–320.

Becker, G. and G. Lewis, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, 1973, *81* (2), S279–88.

____ and N. Tomes, "Child Endowments, and the Quantity and Quality of Children," *NBER Working Paper 123*, 1976.

Becker, Gary Stanley, *A Treatise on the Family*, Harvard university press, 1981.

Bloom, D.E. and D. Canning, "Global demographic change: dimensions and economic significance," *NBER Working Paper 10817*, 2004.

____ , ____ , and P.N. Malaney, "Population dynamics and economic growth in Asia," *Population and Development Review*, 2000, *26*, 257–290.

Bound, J. and S. Turner, "Cohort crowding: How resources affect collegiate attainment," *Journal of Public Economics*, 2007, *91* (5-6), 877–899.

Card, D. and T. Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

Carneiro, P. and J. Heckman, "Human capital policy," *NBER working paper series*, 2003.

chao Chen, Pi and Adrienne Kols, "Population and birth planning in the Peoples Republic of China.," *Population Reports. Series J: Family Planning Programs*, 1982, (25), 577–619.

Coale, Ansley J, *Rapid population change in China, 1952-1982*, Vol. 27, National Academies Press, 1984.

Falaris, E.M. and H.E. Peters, "Schooling choices and demographic cycles," *Journal of Human Resources*, 1992, pp. 551–574.

Greenhalgh, Susan, "Controlling births and bodies in village China," *American Ethnologist*, 1994, *21* (1), 3–30.

Gu, Baochang, Feng Wang, Zhigang Guo, and Erli Zhang, "China's local and national fertility policies at the end of the twentieth century," *Population and Development Review*, 2007, *33* (1), 129–148.

Han, Jun, Wing Suen, and Junsen Zhang, "Picking up the Losses: The Impact of the Cultural Revolution on Human Capital Re-investment in Urban China," *Unpublished working paper*, 2011.

Jiang, *Zhongguo jihua shengyu gongzuo shouce (China Family Planning Workbook)*, China Population Press, 1996.

Joshi, Shareen and T Paul Schultz, "Family planning as an investment in development: evaluation of a program's consequences in Matlab, Bangladesh," Technical Report, IZA Discussion Papers 2007.

Mao, Zedong, Roderick MacFarquhar, Timothy Cheek, Eugene Wu, Merle Goldman, and Benjamin Isadore Schwartz, *The secret speeches of Chairman Mao: from the hundred flowers to the great leap forward*, Vol. 6, Harvard University Press, 1989.

Maurer-Fazio, M., "Earnings and education in China's transition to a market economy Survey evidence from 1989 and 1992," *China Economic Review*, 1999, *10* (1), 17–40.

Mincer, J., "Economic development, growth of human capital, and the dynamics of the wage structure," *Journal of Economic Growth*, 1996, *1* (1), 29–48.

Mitrut, Andreea and François-Charles Wolff, "The impact of legalized abortion on child health outcomes and abandonment. Evidence from Romania," *Journal of health economics*, 2011, *30* (6), 1219–1231.

Peng, *Zhongguo jihua shengyu quanshu (China Family Planning Book)*, China Population Press, 1997.

Pop-Eleches, Cristian, "The impact of an abortion ban on socioeconomic outcomes of children: Evidence from Romania," *Journal of Political Economy*, 2006, *114* (4), 744–773.

Roy, N. and A.D. Foster, "The Dynamics of Education and Fertility: Evidence from a Family Planning Experiment," *University of Pennsylvania, Working Papers 073*, 1996.

Short, S.E. and Fengying. Zhai, "Looking locally at China's one-child policy," *Studies in Family Planning*, 1998, pp. 373–387.

Silles, M.A., "Adult education and earnings: evidence from Britain," *Bulletin of Economic Research*, 2007, *59* (4), 313–326.

Stapleton, D.C. and D.J. Young, "Educational attainment and cohort size," *Journal of Labor Economics*, 1988, pp. 330–361.

Valente, Christine, "Access to abortion, investments in neonatal health, and sex-selection: evidence from Nepal," *Journal of Development Economics*, 2014, *107*, 225–243.

Wachter, M.L. and W.L. Wascher, "Leveling the peaks and troughs in the demographic cycle: An application to school enrollment rates," *The Review of Economics and Statistics*, 1984, pp. 208–215.

Yang, Dennis Tao and Dandan Chen, "Transformations in China's population policies and demographic structure," *Pacific Economic Review*, 2004, *9* (3), 269–290.

Zhang, J., Y. Zhao, A. Park, and X. Song, "Economic returns to schooling in urban China, 1988 to 2001," *Journal of Comparative Economics*, 2005, *33* (4), 730–752.

Zhang, Weiguo, "Economic reforms and fertility behaviour: A Study of a North China village," *Richmond: CURZON*, 2002, *304*, 81–7.