Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

__ECONOMICS__

Ciclo __31__

Settore Concorsuale: _13/A1_____

Settore Scientifico Disciplinare:__SECS-P/01_____

Three Essays in Economics of Education and Economic History

Presentata da: _Yu Bai

Coordinatore Dottorato **Supervisore**

__Prof. Marco Casari **_Prof. Daniela Iorio**

Esame finale anno 2019

Three Essays in Economics of Education and Economic History

Yu Bai

Department of Economics, Università degli Studi di Bologna

Abstract

This dissertation consists of three essays in Economics of Education and Economic History.

The first essay "Disentangling the Effects of Education on Health: A Sibling-pair Analysis" examines the causal effect of education on long-run physical health, using survey data on matched siblings. By adopting a sibling-differences strategy, we are able to obtain estimates that are not biased by unobserved genetic factors and family background which affect both education and health. To address the potential endogenous shocks that affect siblings differently within the family, we further employ an instrumental variable approach by exploiting a profound disturbance in the education system during the Cultural Revolution in China. In the first stage, we investigate the impact of cohort exposure to the disruption on individual completed years of education. The within-sibling estimates in the second stage suggest that an additional school year is found to be positively related to health status later in life (measured by better self-reported health, lower probability of feeling uncomfortable, getting chronic diseases, and being underweight or overweight). We also unravel the potential mechanisms by examining the potential roles of income and cognition in effects of education on health. Controlling for factors that vary within sibling pairs does not alter the conclusion of this paper.

The second chapter "Good Bye Chiang Kai-shek? The Long-Lasting Effects of Education under the Authoritarian Regime in Taiwan" analyzes whether experiencing an authoritarian regime at an early age have long-lasting effects on people's political outcomes, such as political attitudes, voting behavior, and national identity? This paper aims to explore this causality by employing an example from Taiwan. After the former leader Chiang Ching-kuo lifted martial law in 1987, Taiwan ended the

authoritarian regime and began democratization. During that period, the Ministry of Education removed the major part of ideological content in the textbooks and required primary and secondary school to change the extra-curriculum activities and teaching practices, in order to correspond to the needs of the society. Exploiting this historical event, we utilize cut-off birth dates for school enrollment that lead to variation in the length of exposure to the authoritarian education system from 1979 to 1987 within the same birth cohort. Based on around 2,800 observations from the Taiwan Social Change Survey (TSCS), we find that one additional year of exposure to authoritarian education during youth reduces subsequent satisfaction with democracy and political participation, increases the likelihood to support and vote for the KMT party, and drops the probability of self-declared Taiwanese identity rather than Chinese in later life. Our results persist after ruling out alternative interpretations and are robust to the sensitivity tests and different groups definitions.

The third chapter "Quantity-Quality Trade-off in Northeast China during the Qing Dynasty" paints a detailed picture of how, in the absence of an industrialization process, the regional popularization of education shaped the trade-off between human capital and fertility decisions in Northeast China during the Qing Dynasty. Using data from the China Multi-Generational Panel Dataset-Liaoning (CMGPD-LN), we investigate 19,490 adult males born between 1760 and 1880 and originating from nearly 700 different villages. After clearing out the birth-order effects and controlling for a rich set of individual, parental, household-level, and village-level characteristics, we apply both logit and linear probability model as well as an instrumental variable approach to test our hypothesis and address endogeneity issues. Our findings suggest that sibship size, as instrumented by twins at last birth, has a substantial negative effect on the probability of receiving an education, indicating the emergence of a child quantity-quality trade-off for large parts of China's population in the northeastern region since the mid-18th century. We draw attention to new evidence showing that the child quantity-quality trade-off is not necessarily a consequence of industrialization, but rather a result of rational behaviors perpetuated by households in response to high educational returns and accessibility. Such findings provide a new explanation for birth controls, contributing to a better understanding of fertility transition in ancient China.

Keywords. Education; Siblings; Health outcomes; Returns to schooling; The Cultural Revolution; An Authoritarian Regime; Political attitude; The long-term effects; Contemporary Economic History; Quantity-Quality Trade-off; Birth Controls; Taiwan; China; Ancient China

Contents

List of Tables

List of Figures

Chapter 1

Disentangling the Effects of Education on Health: A Sibling-pair Analysis

*

1.1 Introduction

An ageing population presents a phenomenal challenge in China, just as it does in the rest of the world (Lee and Mason, 2011; Huang and Zhou, 2013). Dramatic demographic changes have brought renewed focus to the physical and mental health of the elderlies. Being the most visible of the various dimensions of health, physical health including strength, flexibility, endurance, nutrient intake, and other aspects, is understood to be the accumulation of influence from life experiences. A very general point of view links physical status among the elderly with their educational achievement at early ages. Since education has been discussed as playing a critical role in accumulating knowledge, forming lifestyle and habits (Lleras-Muney, 2005; Clark and Royer, 2013), scholars assume that education generates persistent impacts, in addition to its contemporary ones.

*This study has been published in the B.E. Journal of Economic Analysis & Policy, 2018: 18(4).

In this paper, the causal effect of education on long-term health is disentangled in a novel way. We base our estimates on a sample of siblings from the China Family Panel Studies (CFPS). Generally, siblings share similar genetic factors and, to a large extent, common family background. Associating within-sibling gaps in education with within-sibling differences in health status, we are therefore able to difference out impacts from unobserved genetic endowments and common family characteristics that may bias the education coefficient.

Previous literature utilizes a variety of natural experiments to estimate the causal effect of schooling on health, such as educational reforms (Lleras-Muney, 2005; Oreopoulos, 2006). These studies have, without any doubt, enhanced our understanding about the causal relationship between education and health. However, natural experiments that these studies rely on, affect those whose return to schooling is different from the average returns in the entire population (Cutler and Lleras-Muney, 2008). For example, compulsory schooling laws are proposed to raise the years at school of those at the lower tail of the educational distribution of the population, while having a little influence on those who plan to continue studies anyway. Hence, the estimates imply simply a local average treatment effect (LATE), but not the effect of increasing the schooling years for the whole population. In response to these limitations, scholars have come up with a twin-differencing strategy (Lundborg, 2013). Twin design depends on the occurrence of differences in education within twins. If such differences are equally distributed across twins, the estimates obtained could be closer to an average treatment effect (ATE) instead. Hence the twin-differences strategy has great potential to provide important findings in term of the health returns from schooling in the population.

Unfortunately, the only public and nationwide dataset enabled to match twins across China is the 1% mini population census. However, the 1% mini population census does not involve individual health status.[2] Due to data limitations, we match siblings instead of twins using data from the CFPS survey. However, similar to twin data, the criticism remains that there may be within-pair differences in unobserved elements that affect education. One major candidate for such within-pair differences,

[2]The dataset used in Behrman et al. (2015) is not an open-sourced data and it contains only individuals from Chengdu, Chongqing, Harbin, Hefei and Wuhan, but not a nationwide sample.

as Bound and Solon (1999) address, is individual weight at birth, given that the anecdotal evidence indicates that low birth weight may be correlated with personal ability and health in childhood. Besides this factor, Ashenfelter and Rouse (1998) suggest that parental treatment to children may be another contributor to the within-pair differences.

To address the potential within-sibling endogeneity, we control for individual heterogeneity and employ an instrumental variable approach. We exploit a profound shock that happened in China in the 1960s and 1970s, during the Cultural Revolution. After 1966, the standard public education system was massively destroyed and abandoned due to Chairman Mao's distrust toward intellectuals. This disrupted education system caused by social and political upheavals damaged individuals' human capital accumulation, measured by quantity and quality of education, and sequentially the outcomes including labor market performance in later life (Meng and Gregory, 2002; Giles et al., 2015). This paper utilizes two-stage least squares (2SLS) exploring cohort exposure to the interrupted education system during 1966-1972 as an instrumental variable to investigate the causal effects of education on health. In the first stage, we find that larger exposure to this disruption leads to fewer years of schooling achieved. This contributes to the literature by providing an exogenous source of differences in schooling within siblings.

In the second stage, we employ five variables that link to individuals physical health conditions in the middle ages, which are: self-reported health, physical discomfort, chronic diseases, underweight, and overweight. After applying 2SLS, the sibling-fixed effects estimates show that an extra year of education results in better health conditions, measured by better self-reported health, lower probability of feeling uncomfortable, getting chronic diseases, and being underweight. Controlling for specific factors which differ within siblings, such as weight at birth and early life health, does not alter the conclusion.

In the last part, this paper unravels potential mechanisms by which these outcomes occur by examining the potential roles of income and cognition in the effects of education on health. The estimates suggest that all these two mediators explain well the impact of education achievement on individual health.[3]

[3]Especially for self-rated health and probability of chronic conditions, the mediate power of the

The rest of the paper is organized as follows: Section 1.2 briefly introduces the literature and discusses some recent findings. Section 1.3 introduces the empirical model. After that, we discuss the data used in the analysis from the China Family Panel Studies (CFPS) in section 1.4. Section 1.5 reports the empirical results and section 1.6 includes robustness checks. Mechanism analysis is presented in section 1.7. Section 1.8 concludes the paper.

1.2 Literature Reviews

The first strand of relevant literature maps the education-health linkage.

The effect of education on health may be both intuitive and deliberative. First, schooling and health may be correlated through unobservable genes and family background. An example is from Fuchs (1982), who proposes time preferences as a candidate of "unobserved" factors. It argues that more future-oriented people spend more on both education and health, both as long-run investments, than less patient people. Second, health may affect educational attainment. Illness in childhood may interfere with the studying process and education-related choices and may surely be related to late-life health (Behrman and Rosenzweig, 2004; Black et al., 2007). The omission of these factors in a regression biases the coefficient of interests.

To deal with these issues, early literature considers the instrumental variable approach, by instrumenting children's education with variables such as parental educational achievement (Sander, 1995; Leigh and Dhir, 1997). Berger and Leigh (1989) use per capita income and local per capita expenditures on education in the birthplace, and parental educational achievement as instruments. However, the exogeneity of this set of instruments may be questioned.

More recent papers utilize "natural experiments" to estimate the causal effect of schooling on health. Among these "natural experiments," educational reforms are the most frequently exploited. However, these studies suffer from weak instruments; therefore, most of them find no effect of education. For instance, Lleras-Muney (2005) does not find any significant effect on health when instrumenting for educa-

above two variables is quite strong.

tion. Besides educational reforms, Adams (2002) applies the strategy of Angrist and Kreuger (1991) by adopting birth quarter as the instrument. Their F-values suggest the existence of weak instruments, and therefore, no significant effect of education on health is obtained. Moreover, as addressed in the introduction, these studies identify "local" effects instead of average effects across the entire population.

Other scholars have provided a twin-differences design, which avoids the drawbacks above and brings novel perspectives to the literature. Twin strategy was first employed by Ashenfelter and Krueger (1994) and developed by Ashenfelter and Rouse (1998). Subsequent studies have been done by Miller et al. (2005) using Australian twin data, as well as Ashenfelter and Rouse (1998) applying data collected from the Twinsburg Festivals; similarly Behrman and Rosenzweig (1999) utilize data from the Minnesota Twin Registry, and Bonjour et al. (2003) utilize UK twins data. These papers disentangle the link between education and wage. Lundborg (2013), as the first paper that adopts twin strategy to examine the health returns to schooling, finds that higher educational attainment is positively associated with self-evaluated health and negatively with number of chronic conditions. Behrman et al. (2015) develop this strategy by using the Chinese Adults Twins Survey, and find that schooling-health associations are significant for health-related behaviors, but not for health outcomes. However, these twin studies make strong assumptions that the differences in schooling within pairs are all exogenous.

In this paper, we combine the "natural experiment" method and sibling-differences strategy together to overcome the drawbacks of each, aiming at a higher precision of identification and better representativeness of the results. It has the potential to avoid the problems described above and may, therefore, bring some new findings.

Second, the paper is also related to the literature about the Cultural Revolution in China.

Several studies have examined the effects of the Cultural Revolution from 1966 to 1976. Scholars link the Cultural Revolution to outcomes in educational achievement and subsequent wages. Both Deng and Treiman (1997) and Meng and Gregory (2002) estimate the effect of the Cultural Revolution on educational achievement. A comparison of the Cultural Revolution (CR) cohort and non-CR cohort allows them to find this adverse impact on human capital accumulations. Meng and Gregory

(2007) take a step forward and show that given the educational level achieved, the impact of these missed years of schooling on subsequent earnings is trivial.

This paper, particularly the first stage of the estimation, also supplements the literature by showing that the Cultural Revolution indeed destroyed the human capital accumulation and sequentially, weakened the long-lasting health conditions of the treated cohorts, and this remains true even within sibling pairs.

1.3 Identification Strategy

1.3.1 Sibling Fixed Effects Model

In this section, we employ the sibling fixed effects model to address household-level unobserved factors, including genetic health conditions, and parental characteristics and social class. The identifying assumption is that after controlling for sibling fixed effects, we link the educational gap of each sibling pair to their later difference in health status, obtaining within-sibling estimations as mentioned in the introduction. Family background, social status and occupations of parents, which linked closely to household-level educational decisions and individual health, could be considered invariant for sibling pairs.

Given that the outcomes are binary or ordinal, we apply a linear probability model (LPM) with sibling fixed effects.[4] The outcome of the equation describing the health status can be expressed as:

$$H_i = \gamma_0 + \gamma_1 Edu_i + X_i\delta + u_f + \delta_w + \epsilon_i \tag{1}$$

where the subscripts i and f index the individual i of sibling pair f, respectively. The sibling fixed effects, denoted u_f, address the heterogeneity common to the siblings, including genetic health conditions, parental characteristics and social class. Edu_i stands for individual i's actual years of education; γ_1 interprets that for a difference of an additional year of schooling, how the sibling's relative health outcome

[4]We run linear regressions to ease the interpretation of the coefficients after considering potential heteroscedasticity problem, but results are robust to running Probit or ordered-Probit specifications.

changes. X_i is a vector of control variables, including age, age squared, gender, urban dummy (current and at 12-years-old), marital status, ethnicity, a dummy that indicates whether individual i lives with his/her siblings, δ_w put a set of dummies, including survey year w when i is interviewed, which approximate potential systematic difference existing across the survey waves. ϵ_i is the idiosyncratic term.

After applying for sibling fixed effects, several criticisms need to be noted here: this design does not control for any endogenous factors that affect siblings differently within the household, which would be concerned if there are differences between the siblings that affect parents' decisions on each child's years of education. Hence we firstly control for observable personal characteristics, for instance, the individual's birth weight and health status in childhood, or whether one smokes or drinks from an early age for further checks. Then in the next subsection, we employ the instrumental variable approach to solve the unobservable ones.

Another criticism is that measurement error bias still exists. To counter measurement error, however, we incorporate data from three survey waves. The baseline survey asked respondents "The enrollment time and completion time for each education stage, and how long they spent at each stage." In the following two surveys, interviewers were repeatedly asked the same questions about the respondents' education information. If the responses were inconsistent, the interviewer would remind respondents to reconsider the answer carefully. From the dataset and survey report, we know that the inconsistency rate is lower than 3 percent, which means that measurement error is not a serious problem in this paper. Repeatedly asking the same questions in different surveys might not eliminate measurement error, but to a certain extent, it reduces the occurrence of such a situation.

1.3.2 Instrumental Variable Approach

To solve the potential endogeneity originating from the endogenous sibling-differences mentioned above, we exploit the most profound event in 1960s-1970s China, the Cultural Revolution. Particularly, we focus on education system disruption, an important part of the Cultural Revolution, which affected exogenously and differently on siblings born in different years. An advantage compared to twin strategy shows

up here that sibling pair design allows observations of individuals born in different years, who were thus exposed to different extents to the exogenous event, as the next section illustrates.

Background: School Interruption during the Cultural Revolution In May 1966, Chairman Mao launched the Cultural Revolution in order to stave off capitalist-roaders and reassert his control over the Communist party. With his distrusted towards intellectuals, Mao held the view that the intellectuals were not real believers in socialism and communism, and should be punished and revolutionarily reborn through constant criticism and self-criticism. The idea at that time shook the status of educated people dramatically, while the standard public education system that cultivated intellectuals was massively abandoned and disrupted in China for the entire six years from 1966 to 1972 (MacFarquhar, 1997; MacFarquhar and Schoenhals, 2006).

Since 1966, national education was severely affected by two aspects of the Cultural Revolution. The first is characterized by the closure of schools and therefore a shorter schooling period. The universities were shut down for a decade, elementary and high schools were closed for two years and subsequently reopened with a low-quality expansion. Even when schools were reopened during 1968 and 1972, the public schooling became shorter, shrinking from 6-3-3 primary-junior-senior high school years to 5-2-2. The second aspect involves curriculum changes after the reopening of the schools. Neither primary nor secondary schools followed the standard school curriculum from before the Cultural Revolution, instead emphasizing leader worship, political ideology, and peasant and worker techniques.[5] The students also were required to participate intensively in the revolution and were organized as Red Guards.

To sum up, between 1966 and 1972, the disrupted public education system re-

[5]From historical documentation, we can see how the curriculum changed: before 1966, pupils learned basic Math and Chinese, while between 1966 and 1972, for students after Grade 3, Chinese lessons were replaced by the study of the poems of Chairman Mao and articles about class conflicts, and pupils were required to take a new course named "Quotations from Chairman Mao." Middle school students were required to take the new course "Quotations from Chairman Mao" as well, while contents of other courses reduced sharply.

duced the schooling period of the cohorts affected. On the one hand, individuals needed fewer years at school to obtain a diploma, and on the other hand, both closure of schools and universities, and ideological infiltration in the curriculum discouraged the treated cohorts to continue their studies. The quality of education experienced a sharp reduction conditioned on years of schooling since students learned less knowledge at school compared to students before or after that period; instead, most of their time involved production practices in factories or farms and political activities. The abnormal schooling system during the Cultural Revolution has been demonstrated by Cai and Du (2003) and Meng and Gregory (2007), which suggest that as a result of the disruption, both education quantity and quality was largely affected during the revolutionary years. We focus particularly on the reduced schooling years and the corresponding outcomes in the following parts.

Instrument: Cohort Exposure to Disrupted Education We exploit the exposure to the disrupted education of enrollment cohort to which the individual belongs, as an instrumental variable.[6] To capture the cohort variation in exposure, we check the distribution of the schooling cohorts during disruption years (1966-1972). In CFPS, rich information on individual enrollment time and duration for certain levels of education is provided. Following Wu (2017), we calculate the proportion of people at school during 1966-1972 by enrollment cohorts, denoted $CohortExposure_j$, where the subscript j represents the enrollment cohort that individual i belongs to. The larger the proportion is, the larger ratio of people whose education has been disrupted in the movement, to the larger extent a specific cohort was exposed to schooling interruption. Figure 1.1 shows the distribution, whereby the schooling cohorts, who were born in the 1950 and 1960s, are quite likely to be at school during the period 1966-1972, and inevitably, are the ones whose education has been interrupted. Hence, we take individuals who were born between September 1949 and September 1967 (and therefore more intensively involved in the disruption) as IE cohorts. Within IE cohorts, the different exposure ratios shown by the black line allow the comparison

[6]To define enrollment cohorts, we follow the literature by assuming individuals who were born after September would enroll one year later compared with individuals born in the same year but before September. With 1 September the cut-off date for school entry, we thus define the enrollment cohorts.

of siblings that were both affected, but to a different extent. Meanwhile, the non-IE cohorts, including the pre-IE cohorts (those born before September 1947) and the post-IE cohorts (those born after September 1966), were almost unaffected by this educational channel. We exploit a cutoff of cohort exposure equal to 10 percent in defining IE-cohorts and not-IE cohorts. Table 1.1 transmits the same messages, and documents the detailed proportions of being at school by enrollment cohort and indicates clearly whether an enrollment cohort belongs to the IE-cohorts or not. An issue to be addressed here is that all of these 32 cohorts listed in Table 1.1 were exposed to the Cultural Revolution, and the direct impact from the political movement and social chaos other than education disruption can be considered the same for all these cohorts conditional on birth cohort fixed effects.

Two-Stage Least Squares We then conduct 2SLS estimation. The first-stage model is illustrated as follows:

$$Edu_i = \phi CohortExposure_j + \psi X_i + u_f + \delta_{bw} + \epsilon_i \tag{2}$$

Where the subscript j represents the enrollment cohort where individual i belongs. In this stage, we utilize this $CohortExposure_j$ to instrument years of education. Besides all the covariates in equation (1), birth cohort fixed effects are further included to control for any birth cohort[7] trend and heterogeneity from the social unrests during the Cultural Revolution apart from the educational disturbance.

The estimate of ϕ is expected to be negative.[8] This expectation on first-stage results is supported by Meng and Gregory (2002, 2007) and Giles et al. (2008), which investigate the negative impact of this interrupted education system during the revolution on individual educational attainment, measured by years in school.

[7]We employ a three-year window to the years of birth from 1943 to 1976.

[8]Two underlying mechanisms explain the first-stage association: higher proportion of exposure of a given enrollment cohort, lower school years achieved of its belonging individuals. One is that the individual i, as the part of cohort j's proportion attending school when the disruption occurred, was those whose education has been directly suspended and shortened. The other mechanism is the spillover effects. For the individual i who had left school before 1966, the estimated impact would be only caused by the externalities or education of others enrolled in the same year. In section 1.6.2, we report that the second mechanism is not the main driver of the first-stage results.

Figure 1.2 shows that the expected findings may also be supported by our dataset by reporting the actual years of education by three educational levels, primary school, junior high school and senior high school, respectively. For primary school achievement, the affected cohorts have fewer years of schooling than their younger and older cohorts. The pattern is similar for junior high and senior high school students that the IE cohorts have relatively fewer years of schooling on average. These differences are attributed to the closure of schools and the shorter public educational system faced by the IE cohorts.

Finally, we then conduct 2SLS to estimate the effects of education on health:

$$H_i = \lambda_0 + \lambda_1 \widehat{Edu_i} + \psi X_i + u_f + \delta_{bw} + \epsilon_i \tag{3}$$

$\widehat{Edu_i}$ is the predicted education value of equation (2) and all the other variables are the same as those in equation (2). The sibling fixed effects appear in both stage, denoted u_f, therefore λ_1 is again interpreted as a sibling-differences estimate.

1.4 Data and Variables

1.4.1 Dataset

We utilize the China Family Panel Studies (CFPS), a nationally representative, longitudinal survey of Chinese families and individuals launched by the Institute of Social Science Survey (ISSS) of Peking University, China.

The CFPS consists of four questionnaires (Community, Family, Adolescent, and Adult), with economic information, as well as information on the wellbeing of the Chinese population. It contains abundant information on demographic and socioeconomic characteristics, such as gender, date of birth, registered residency (hukou in Chinese[9]), type of residency (rural or urban), ethnicity, marital status, educational attainment, employment status, etc., and covers research topics such as education outcomes, economic activities, family dynamics and relationships, and health. In the 2010 baseline survey, the CFPS successfully surveyed nearly 15,000 households

[9]Hukou is a system of household registration in China.

and approximately 42,590 individuals within those households. In 2012 and 2014, the CFPS carried out two full-scale follow-up surveys on tracked samples of 2010 reservations and further included abundant new samples. The 2012 and 2014 questionnaires shared similar structures to those in 2010. We employ responses from 2010, and from new participants in 2012, and 2014 as well, creating a pooled cross-sectional dataset, in order to obtain within-birth cohort variations in treatment and age.

The unique parental ID in the cross-sectional dataset allows us to match the siblings. Out of the 35,889 individuals interviewed, around 550 individuals of 180 households were chosen for our sibling sample.

1.4.2 Physical Health Outcomes

The CFPS requires interviewees to answer multiple questions about their physical and mental health. Among those, five questions are directly linked to the respondent's physical health conditions, which we use as outcome variables.

Self-reported health is an ordinal variable created by the responses to the question: "How would you describe your health in the last six months?" Individuals are required to evaluate their recent health status from 1 to 5: 1 for excellent, 2 very good, 3 good, 4 fair and, 5 poor. We employ *Self-reported health* as our first variable, which is the standard practice in the health economics literature. Although it is a subjective outcome variable, normally individuals are more aware of their own health than anyone else is. In the sibship sample, 13.5 percent of respondents reported poor health, while nearly 20 percent report fair health.

We collected responses to the question "During the past two weeks, have you felt physically uncomfortable?" which is denoted as *physical discomfort*, with value zero for "no" and one for "yes." It is still a subjective variable but is vital: discomfort may be considered a subtle sign to warn people of health problems. Around 29 percent of the sibling sample have felt uncomfortable in the survey.

The third measure is the *chronic disease*, which captures doctor-diagnosed disease or injury. *chronic disease* equals one if the respondent has been diagnosed by the doctor with a chronic disease, and zero otherwise. This could be regarded as an

important objective health variable: whether the doctor confirmed the respondent has a chronic disease. 14 percent of the sample was diagnosed with a *chronic disease*.

The last two are denoted as *underweight* and *obesity*, indicating whether an individual is underweight (BMI$\leq$19) or overweight (BMI$\geq$28), which have been widely used in the literature in describing the subjects' nutrition (Huang, 2017). BMI is a value derived from the weight and height of an individual, and since the dataset offers us this information, we calculate the index by the formula. The reported underweight rate is 9 percent, while the obesity rate is 23 percent.

Descriptive Statistics Column (1) of Table 1.2 summarizes descriptive statistics for the health outcome variables above, years of schooling, cohort exposure to the educational disruption and the covariates of the whole sibling sample. It shows that the total number of siblings is 550, and overall, the sample had 7.46 years of schooling. For the covariates, 85 percent of the respondents are male. This large percent is driven by the fact that for the CFPS, sibling pairs appear mainly in households with all or some of the siblings living together, with a share of 90 percent so all siblings could be surveyed. And in rural China, it is more common that male siblings lived together, while female siblings left when they got married. Hence the matched siblings pair are mainly male siblings.[10] The average age in the sample is 44. 8 percent are ethnic minorities, 35 percent are urban residents, 10 percent lived in urban areas during the Cultural Revolution. Around 9 percent were retired, while nearly 78 percent had gotten married.

The descriptive statistics by cohort are reported in columns (2) and (3). Around 35 percent of respondents belong to the IE cohorts.[11] Comparing to the older and younger cohorts, IE cohorts were 33 percent more likely to be at school during the Cultural Revolution, and that is how we distinguish IE groups from the non-IE ones. The raw p-values testing for the statistical significance of the differences in characteristics of an interviewee in IE and non-IE groups, are presented in column (4). The health status of the siblings whose education was interrupted is significantly

[10]However, this limitation does not threaten our analysis, since this pattern shows up in both IE-cohorts and non-IE cohorts.

[11]Their projected birth dates are defined in Table 1.1.

worse, i.e. a lower state of self-reported health, a higher probability of feeling physical discomfort, getting a chronic disease, or being underweight. Moreover, siblings that belong to IE cohorts are relatively less educated, measured by around one fewer year of schooling. Units in the IE and non-IE groups are balanced on almost all other covariates, which indicates that the two groups are nearly identical on observable characteristics.

1.5 Empirical Results

1.5.1 Baseline Results

In Table 1.3, we adopt an LPM model as a model specification with the sibship sample and examine the effect of schooling on health profile.[12] Odd columns report raw correlations between schooling and health condition. In even columns, we include sibling fixed effects that allow within-sibling estimations. The raw correlations suggest that higher level of education is negatively and significantly associated with bad health status, measured by poorer self-rated health, higher odds of discomfort, getting chronic diseases, and being underweight as well, except for obesity. The results conditional on the sibling fixed effects show consistent findings. Education decreases the self-reported indicator by 2.5 percentage points in column (2). An additional year in school corresponds to a 1.4 percentage points decrease in the likelihood of discomfort as shown in column (4) and a 0.7 percentage points reduction in being diagnosed with a chronic disease as column (6). Columns (8) and (10) capture the significant effect of school-age education on abnormal BMI, with a conditional correlation of -0.013 and -0.009.

In Table 1.4, we conduct the sibling fixed effects model as a model specification with the sibship sample, controlling for age, age squared, gender, ethnicity,

[12] Choosing siblings out of the full sample does potentially generate selection issues. We therefore experiment with traditional Heckman selection models, using household size in the selection equation. With a larger household size, there is a higher probability of an individual living with parents and siblings and, therefore, they are more likely to be interviewed and recorded together. However, we found no evidence that selection affects our estimates from the insignificant inverse Mills ratios shown in Appendix Table A.1.

urban dummy (both current and at 12 years old), marital status, retirement dummy, and whether the individual lives with his/her siblings and wave fixed effects in odd columns. We further include province fixed effects in even columns. The conditional results estimated by the sibling fixed effects model show the same pattern as Table 1.3 and stay robust no matter if our set of covariates is included or not: one extra year of education is associated with a reduced probability of poor health conditions. The signs of the covariates are what we expected: females and married couples are healthier than males and singles, respectively. Individuals who are retired have better self-reported health and less likelihood to be underweight, conditional on age and age squared. If individuals obtained the urban *hukou*, the likelihood of being underweight drops while the probability of chronic disease increases. The limitation of data prevents us from exploring the reasons. One probable reason may be China's serious urban pollution, which leads to chronic respiratory diseases.

Appendix Table A.2 further divide the sibship sample by type of residency and gender to study the heterogeneous within-sibling effects of the education on health. The place of residence has long been considered as a key health determinant, which is particularly the case in China, where disparities among urban and rural populations are tremendous due to legal designations implemented by *hukou* system and the differentiated economic, health, and other social policies applied to these two departments. Consistent with the policy implications, the results indicate that the effect of education is greater for sibling pairs with rural *hukou*. For sibling pairs with urban *hukou* in their schooling ages, the effect is statistically insignificant from zero for most of the outcome variables. The difference may be attributed to alternative channels to access knowledge in developed urban areas. This accessibility can reduce the gap in school between the urban siblings.

Heterogeneity between the Siblings The crucial assumption made in the literature is that schooling differences within siblings are exogenous, conditional on the sibling fixed effects and included covariates. There is perhaps no way of proving that differences in schooling within siblings are exogenous. It could shed some light on the issue if abundant information on early life differences between siblings is available. Since the CFPS has such information on early life factors, we would like to address

these problems. In this part, we control for individual weight at birth, which is a continuous variable, and poor health status in early age, which is assigned one if the respondents claimed that the reason they could not graduate/pursue studying is a health problem, and zero otherwise.[13] Heterogeneity in young-age habits is also considered, such as information about drinking and smoking from an early age.

After controlling for heterogeneity across siblings, the robustness of the effect remains, as shown in Table 1.5. It is not surprising that early life health status is positively correlated with health in later life, while an individual's weight at birth does not play a role, conditional on household health-related genetic factors. However, these cannot control for unobservable heterogeneity within the family, which may affect the parents' decisions on a child's years of schooling, and subsequent health conditions in old age before implementing 2SLS.

1.5.2 Instrumented Results

The first-stage results are available in Table 1.6. In all columns, we control for birth cohort fixed effects and sibling fixed effects, and set of covariates. The estimates remain negative and statistically significant, indicating an adverse impact of exposure to educational disruption on the actual years of schooling of the individuals from the schooling cohorts. The F-statistics in the first stage and Hansen tests for the instrumental variable are reported at the bottom of Table 1.6. The F-values provide evidence of a significant first stage no matter if the covariates are included or not, suggesting the instrument is by no means a weak instrument. Moreover, the instrument also passes the Hansen tests (over-identification tests). The excludability of the instrument cannot be tested directly, but indirect evidence can be offered in support of such an assumption. Hence in Table A.3 of Appendix, we regress the variables that correlate with health outcomes on the instrument. In these regressions, our instrument is statistically insignificant in any regression, the validity of which is further confirmed.

In Table 1.7, we report the second-stage results by regressing health outcomes

[13] In addition to these, there are other ways to do sensitivity analysis, such as excluding sibling pairs that faced very different early life health conditions from the analyses or controlling for parental time investments and attitudes towards each of their children.

on instrumented years of schooling. For self-evaluated health conditions, the sibling-differences 2SLS estimates in columns (1) reveal that an additional year of education reduces the indicator by 0.06 (4 percent of the mean). Columns (2) suggest the effect of education on the odds of reporting feeling uncomfortable is 5 percentage points (17 percent of the mean). From columns (3), we know the sibling-differencing effect on chronic disease is -0.035. Column (4) show that an additional one year of schooling results in a drop of 2 percentage points in the underweight rate (around 17 percent of the mean), suggesting that schooling can strongly improve nutritional status. The sibling-differences effects on obesity are around -0.020, which is statistically insignificant though. These within-sibling results suggest that the affected cohorts have been placed in disadvantaged positions in term of health as a result of various interruptions in human capital accumulation in their youths.

Some Confounding Factors A source of heterogeneity comes from a confounder along with the instrument, the send-down policy. During the Cultural Revolution, Chairman Mao also initiated the forced rustication movement by sending millions of urban middle school students (10-19 years) to the countryside, which is called the "up to the mountains and down to the countryside" (or "send-down") movement. The scheme applied to all eligible urban individuals who would graduate from high school, exactly those who were exposed to the closure of schools. The eligible urban youths were therefore exiled to the countryside with extremely heavy farm work instead of being educated during their school-age years, even after the reopening of schools. Being estranged from their family for several years, they experienced a dramatic decline in the standard of living, and consequentially they were more likely to develop health problems later in life, as discussed by Xie et al. (2007), Li and Walder (2010), Liang and Li (2014) and Gong et al. (2014). Hence, this variable should be explicitly controlled for. Based on the sent-down information from the CFPS dataset, we would like to put send-down in the model to conduct the following exercise.

In this part, we again control for individuals' early-age characteristics to solve heterogeneity within-siblings: birth weight and early life health. The results conditional on these individual heterogeneities are shown in columns of Table 1.8, which summa-

rize the regression results on health outcomes conditional on early life characteristics and send down experience. This lends additional credibility to our estimation and to the validity of applying an instrumental approach to examine the health returns to schooling. The estimated parameters of our variable of interests remain negative and significant.

1.6 Robustness Checks

1.6.1 Confounding Historical Factors

Now we focus on a confounding factor which might threaten our identification strategy, the Great Famine, which was a period in China between 1959 and 1961 characterized by serious famine.[14] Both harsh weather and the incorrect policies of the Chinese Communist Party contributed to this widespread famine. The event may also have had persistent adverse consequences on children who were born during the famine and therefore exposed to malnutrition and other harmful environments during their fetal period (Hertzman and Wiens, 1996; Hertzman and Power, 2003; Chen and Zhou, 2007). Since part of our sample was born around the Great Famine, it is necessary to control for the potential consequences of the famine on health in later life (Almond et al., 2010). In Table 1.9 we consider age 0-3 for the shock to the initial nutrition endowment due to famine by creating a dummy for famine and assigning a value of one to an individual who was born between 1956 and 1961. Results of models controlling for the Great Famine are reported in Table 1.9. The effects of schooling on later-life health remain robust regardless of considering or not the effect of the famine. Table 1.9 also reports the significant effect of the Great Famine on self-reported health and the likelihood of physical discomfort, conditional on sibling fixed effects.

[14] According to government statistics and historical documentation, there were about 15 million excess deaths during this period.

1.6.2 Placebo Test

A placebo test is conducted to check in the first stage, whether the disrupted national education system affects the quantity of education of some individuals when it should not. Specifically, we select those who were born as schooling cohorts as defined in Table 1.1, but stopped attending school before 1966 and consider these people as "fake" schooling cohorts, because their education was not in fact interrupted by the Cultural Revolution. In this part, we compare siblings belonging to "fake" IE cohort with siblings born before September 1949 (pre-IE cohorts) or after September 1967 (post-IE cohorts). Since "fake" schooling cohorts were not targeted by the education disruption, we expect no statistically significant effect of the instrument on education in the first-stage, hence no effect on health outcomes. In other words, if there are any significant effects on those outcomes, then one would suspect that the instrument $CohortExposure$ may have captured something else out of the educational channel. The reduced-form results are shown in Table 1.10. No significant effects of the $CohortExposure$ could be found on the outcomes of the "fake" IE cohort.

1.7 Mechanisms

We follow the methodology in Cutler and Lleras-Muney (2010) and Huang (2017) to quantify the possible channels by including proxies of potential mediators (i.e., income and cognition) into the equation:

$$H_i = \gamma_0 + \gamma_1 \widehat{Edu_i} + \psi X_i + u_f + \delta_{bw} + \epsilon_i \tag{4}$$

$$H_i = \gamma_0' + \gamma_1' \widehat{Edu_i} + \psi' X_i + Med_i + u_f + \delta_{bw} + \epsilon_i \tag{5}$$

The outcome variable H_i denotes the five health outcomes mentioned above. We use instrumented education directly because it helps us to get unbiased results. All the other settings are the same as those in equation (3). Med_i proxies the potential mediators (i.e., income and cognition). Following the literature, we exploit the change in the magnitude of the coefficient on instrumented education as the part explained

by Med_i. The explained proportion equals $1 - |\gamma_1'/\gamma_1|$, the larger this proportion is, the larger the mediate power.

Income In Panel A of Table 1.11, an additional year of schooling increases monthly wage by 4.3 percentage points, close to the findings in the previous literature (Angrist and Kreuger, 1991). It further confirms the validity of our instrument and 2SLS. Unsurprisingly, higher education predicts higher income, suggesting people with a better educational level can live a higher quality life, such as buying a house in a safer region, working in a better environment, experiencing less financial pressure, etc., and can purchase better health products, preventive medical care, and health insurance.

Column (1) of Panel B reports the overall effects of the instrumented education. Column (2) shows the conditional effects when logged income is controlled for and column (3) shows the results for the proportions explained by the logged income. The part that can be explained by income reaches its maximum, a 62.59 percent for self-reported health, and 46.85 percent for chronic disease in the sibship sample. The mediate power of income is also large for being underweight, 26.91 percent. In general, income is the main channel through which education affects health.

Cognition The CFPS data offers respondents' performance in verbal tests and number series tests, which allows evaluation of individuals' cognitive skills. As suggested in Panel A, the IE cohorts have lower scores in both verbal tests and number series test on average given educational level. This reveals that educational quantity and quality of schooling cohorts were profoundly affected during the Cultural Revolution. Previous literature also suggests a reduction in educational quality by birth cohort, conditional on one's educational attainment.

Since it has been well demonstrated that education plays an important role in generating social capital (Helliwell and Putnam, 2007; Oreopoulos and Salvanes, 2011), it is very likely that the abnormal education system during the Cultural Revolution damaged health by lowering not only education quantity but also quality. To test this hypothesis, a straightforward way is to control for both quantity and quality of schooling at the individual level, with the results presented in columns

(4) and (5) of Panel B. Column (4) suggests the conditional effects when cognitive skills, measured by word recall test and math test, are controlled for, and column (5) shows the corresponding explained proportion. The proportion that can be mediated by cognition ranges from 6 percent to 26.14 percent, implying that cognition is also an important channel, particularly for chronic disease and underweight status. The findings are consistent with the literature that stresses the importance of cognitive skills (Hanushek and Woessmann, 2008; Aaronson and Mazumder, 2011; Carlsson et al., 2012).

Formation of Beliefs and Habits Education can improve individuals' daily living skills, develop their endurance, and therefore well-educated people may have grown to cherish life and develop healthy habits from an early age. Unfortunately, in CFPS, questions for health-related beliefs and habits do not always appear. For instance, individual hours of physical exercise appear only in the 2010 and 2014 waves. Due to data limitations, we leave the questions to future studies that will give us a better understanding of the mediate power of beliefs and habits.

1.8 Conclusion

This study estimates the effects of the schooling at the young age on long-term health status from a novel perspective, by linking different levels of education within siblings to different health statuses in their later life. Using sibship samples from the CFPS, our results show that after controlling for unobservable genetic factors and family background, less educated siblings suffer adverse effects regarding physical health. To solve the heterogeneity issues within sibling-pair studies, we control for the individuals' health at birth and at a very early age.

To address the issue of endogeneity, that is if the schooling decision of an individual remains endogenous conditional on household characteristics and preferences, we further employ a 2SLS approach. In the first stage, we exploit the disrupted national education system during the Cultural Revolution, which exogenously affected the schooling cohorts (IE cohorts). Between 1966 and 1972, the schooling system became shorter, shrinking from 6-3-3 primary-junior-senior high school years

to 5-2-2, and the lack of a normal curriculum discouraged schooling cohorts from continuing their studies, hence the cohorts exposed more to interrupted education have fewer years of schooling. The proportion of each enrollment cohort attending school during the Cultural Revolution captures the cohort's exposure to interrupted education and negatively explains its belonging individual's length of education in the first stage. The predicted years of schooling are put in the regression instead of the endogenous education level in the second stage, and the sibling-differencing estimation indicates that an additional year of schooling drops self-reported health by 0.06, decreases the probability of reporting feeling uncomfortable by 5 percentage points, and the likelihood of being diagnosed with chronic disease by 2.5 percentage points. Meanwhile, it reduces by about 4 percentage points in the abnormal weights rate.

The rest of this paper explores the potential mechanisms underlying these results. Following the framework in Cutler and Lleras-Muney (2010), we examine the potential roles of income and cognition in effects of education on health. The estimates suggest that these two factors strongly explain the impact on health. These results confirm the causal association between schooling and health and suggest that the 50s-60s cohorts in China have been placed in disadvantaged positions in term of health as a result of the interruptions in their early human capital accumulation.

Bibliography

[1] Aaronson, D., B. Mazumder. 2011. The Impact of Rosenwald Schools on Black Achievement. Journal of Political Economy, 119 (5), 821-888.

[2] Adams, S. 2002. Educational Attainment and Health: Evidence from a Sample of Older Adults. Education Economics, 10 (20), 97-109.

[3] Almond, D., Edlund, L., Li, H. and Zhang, J. 2010. Long-Term Health Effects of the 1959-1961 China Famine: Mainland China and Hong Kong, in T. Ito and A.K. Rose (Eds.) The Economic Consequences of Demographic Change in East Asia.

[4] Angrist, J., A. Krueger. 1991. Does Compulsory School Attendance Affect Schooling and Earnings? Quarterly Journal of Economics, 106 (4), 979-1014.

[5] Ashenfelter, O., Krueger, A. 1994. Estimates of the Economic Return to Schooling from a New Sample of Twins. American Economic Review, 84 (5), 1157-1173.

[6] Ashenfelter, O., Rouse, C. 1998. Income, Schooling and Ability: Evidence from a New Sample of Identical Twins. Quarterly Journal of Economics, 113 (1), 253-284.

[7] Behrman, J., Rosenzweig, M. 1999. "Ability" Biases in Schooling Returns and Twins: A Test and New Estimates. Economics of Education Review, 18 (2), 159-167.

[8] Behrman, J. R. and M. R. Rosenzweig. 2004. Returns to Birthweight. Review of Economics and Statistics, 86, 586-601.

[9] Behrman, J., Rosenzweig, M., Taubman, P. 1994. Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment. Journal of Political Economy, 102 (6), 1131-1174.

[10] Behrman, J., Y. Xiong and J. Zhang. 2015. Cross-Sectional Schooling-Health Associations Misrepresented Causal Schooling Effects on Adult Health and Health-

Related Behaviors: Evidence from the Chinese Adults Twins Survey. Social Science and Medicine, 127, 190-197.

[11] Berger, M. C., and J. P. Leigh. 1989. Education, Self-Selection and Health. Journal of Human Resources, 24 (3), 433-455.

[12] Black, S. E., P. J. Devereux, and K. G. Salvanes. 2007. From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes. Quarterly Journal of Economics, 122 (1), 409-439.

[13] Bonjour, D., L. Cherkas, J. Haskel, D. Hawkes, T. Spector. 2003. Returns to Education: Evidence from UK Twins. American Economic Review, 93 (5), 1799-1812.

[14] Bound, J., G. Solon. 1999. Double Trouble: On the Value of Twin-Based Estimation of the Return to Schooling. Economics of Education Review, 18 (2), 169-182.

[15] Cai F., D. Yang. 2003. The Cultural Revolution on Physical Capital and Human Capital (In Chinese). China Economic Quarterly. 2 (4), 795-806.

[16] Carlsson, M., G. B Dahl, B. Ockert, D. Rooth. 2012. The Effect of Schooling on Cognitive Skills? Review of Economics and Statistics, 97 (3), 533-549.

[17] Chen, Y., and Zhou L. 2007. The Long-Term Health and Economic Consequences of the 1959-1961 Famine in China. Journal of Health Economics, 26 (4), 659-681.

[18] Clark, D., Royer, H. 2013. The Effect of Education on Adult Mortality and Health: Evidence from Britain. American Economic Review, 103 (6), 2087-2120.

[19] Cutler, D., A. Lleras-Muney. 2008. Education and Health: Evaluating Theories and Evidence. In: The Effects of Social and Economic Policy on Health, Editors J. House, R. Schoeni, G. Kaplan, and H. Pollack, Russell Sage Press

[20] Cutler, D., Miller, G., Norton, D.M. 2007. Evidence on Early Life Income and Late Life Health from America's Dust Bowl Era, Proceedings of the National Academy of Science, 104 (33), 13244-13249.

[21] Deng, Z., Treiman, D.J. 1997. The Impact of the Cultural Revolution on Trends in Educational Attainment in the People's Republic of China. American Journal of Sociology, 103 (2), 391-428.

[22] Fuchs, V. R. 1982. "Time Preference and Health: An Exploratory Study", in: Fuchs, V.R., (ed.), Economic Aspects of Health, University of Chicago Press, Chicago.

[23] Giles, J., Park, A., Zhang, J.W. 2004. The Great Proletarian Cultural Revolution, Disruption to Education and Returns to Schooling in Urban China. Working Paper.

[24] Gong, J., Lu, Y., Xie, H. 2014. Rusticated Youth: The Send-Down Movement and Beliefs. Working Paper.

[25] Grossman, Mi. 2006. Education and Non-Market Outcomes. In Handbook of the Economics of Education. Vol. 1, Amsterdam: North-Holland, Elsevier Science, 577-633.

[26] Hanushek, E. A and L. Woessmann. 2008. The Role of Cognitive Skills in Economic Development. Journal of Economic Literature, 46 (3), 607-668.

[27] Helliwell, J.F., Putnam, R.D. 2007. Education and Social Capital. Eastern Economic Journal, Eastern Economic Association, 33 (1), 1-19.

[28] Hertzman, C., and M. Wiens. 1996. Child Development and Long-Term Outcomes: A Population Health Perspective and Summary of Successful Interventions. Social Science & Medicine, 43 (7), 1083-1095.

[29] Hertzman, C., and C. Power. 2003. Health and Human Development: Understandings from Life-Course Research. Developmental Neuropsychology, 57 (10), 778-783.

[30] Huang, W., and Y., Zhou. 2013. Effects of Education on Cognition at Older Ages: Evidence from China's Great Famine. Social Science & Medicine, 98, 54-62.

[31] Huang, W. 2017. Understanding the Effects of Education on Health: Evidence from China. Working Paper.

[32] Lee, S.H., A., Mason and D., Park. 2011. Why does Population Aging Matter So Much for Asia? Population Aging, Economic Security and Economic Growth in Asia, Working Papers, DP-2011-04, Economic Research Institute for ASEAN and East Asia.

[33] Leigh, P. J., and R. Dhir. 1997. Schooling and Frailty Among Seniors. Economics of Education Review, 16 (1), 45-57.

[34] Li, B., Walder, A. 2001. Career Advancement as Party Patronage: Sponsored Mobility into the Chinese Administrative Elite, 1949-1996. American Journal of Sociology, 106 (5), 1371-1408.

[35] Liang, P., Li, J. 2014. Regretless Youth? The Long-Term Effect of Send-Down Experience on Social Trust (in Chinese). World Economic Papers, 2, 90-109.

[36] Lleras-Muney, A. 2006. The Relationships between Education and Adult Mortality in the United States. Review of Economic Studies, 73 (1), 189-221.

[37] Lundborg, P. 2013. The Health Returns to Schooling–What can We Learn from Twins? Journal of Population Economics, 26 (2), 673-701.

[38] MacFarquhar, R. 1997. The Politics of China: The Eras of Mao and Deng. Cambridge University Press.

[39] MacFarquhar, R. Schoenhals M. 2006. Mao's Last Revolution. Cambridge, Mass. Harvard University Press. pp. 693.

[40] Meng, X., Gregory, R.G. 2002. The Impact of Interrupted Education on Subsequent Educational Attainment: A Cost of the Chinese Cultural Revolution. Economic Development and Cultural Change, 50 (4), 935-959.

[41] Meng, X., Gregory, R.G. 2007. Exploring the Impact of Interrupted Education on Earnings: The Educational Cost of the Chinese Cultural Revolution. Working Paper.

[42] Miller, P., C. Mulvey, N. Martin. 2005. Birth Weight and Schooling and Earnings: Estimates from a Sample of Twins. Economics Letters, 86 (3), 387-392.

[43] Oreopoulos, P. 2006. Estimating Average and Local Average Treatment Effects of Education when Compulsory School Laws Really Matter. American Economic Review, 96 (1), 152-175.

[44] Oreopoulos P., Salvanes, K. 2011. Priceless: The Non-Pecuniary Benefits of Schooling, Journal of Economic Perspectives, 25 (1), 159-184.

[45] Sander, W. 1995. Schooling and Quitting Smoking. Review of Economics and Statistics, 77 (1), 191-199.

[46] Wu, L.W. 2017. The Invisible Wound: The Long-Term Impact of China's Cultural Revolution on Trust. Working Paper.

[47] Xie, Y., Jiang, Y., Greenman, E. 2007. Did Send-Down Experience Benefit Youth? A Reevaluation of the Social Consequences of Forced Urban-Rural Migration during China's Cultural Revolution. Social Science Research, 37 (2), 686-700.

[48] Zhang, J., Liu, P.W. and Yung, L. 2007. The Cultural Revolution and Returns to Schooling in China: Estimates Based on Twins. Journal of Development Economics, 84 (2), 631-639.

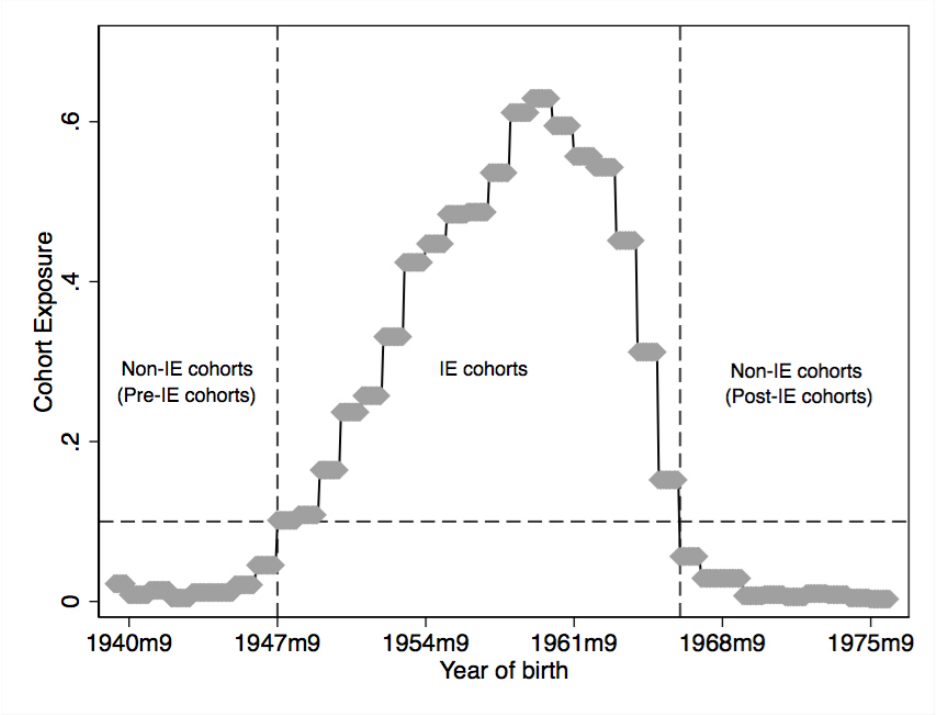

Figure 1.1: Proportion of Being at School during the Educational Disruption

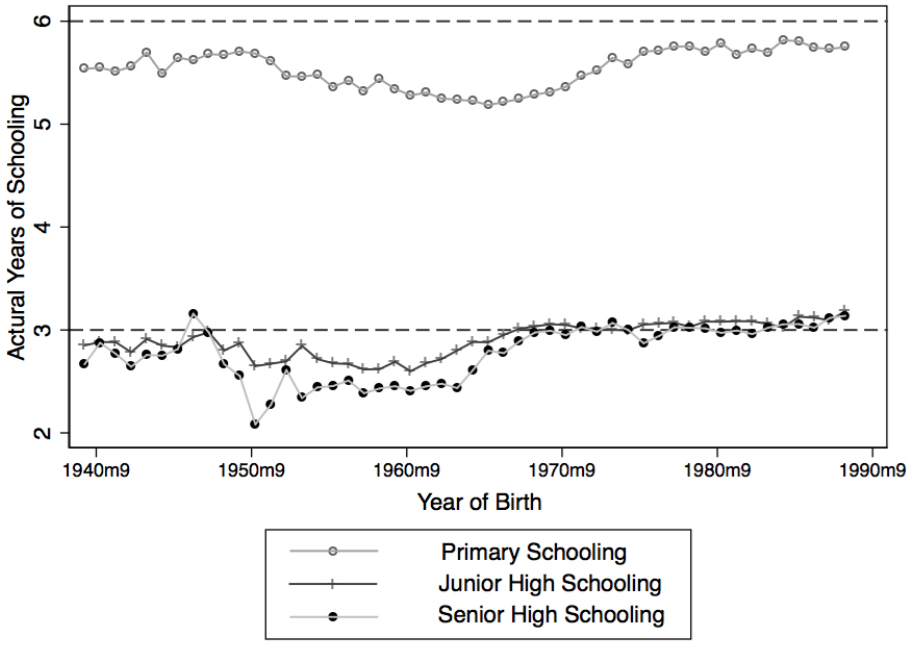

Figure 1.2: Changes in Actual Years of Schooling

Table 1.1: Cohort Exposure to Interrupted Education

Dates of Birth (Enrollment cohort)	Proportion of being at school (Cohort Exposure to IE)	Corresponding cohorts
Sep,1975- Aug,1976	0.002	post-IE cohort
Sep,1974- Aug,1975	0.002	post-IE cohort
Sep,1973- Aug,1974	0.004	post-IE cohort
Sep,1972- Aug,1973	0.007	post-IE cohort
Sep,1971- Aug,1972	0.009	post-IE cohort
Sep,1970- Aug,1971	0.006	post-IE cohort
Sep,1969- Aug,1970	0.007	post-IE cohort
Sep,1968- Aug,1969	0.006	post-IE cohort
Sep,1967- Aug,1968	0.028	post-IE cohort
Sep,1966- Aug,1967	0.029	post-IE cohort
Sep,1965- Aug,1966	0.311	IE cohort
Sep,1964- Aug,1965	0.450	IE cohort
Sep,1963- Aug,1964	0.543	IE cohort
Sep,1962- Aug,1963	0.556	IE cohort
Sep,1961- Aug,1962	0.594	IE cohort
Sep,1960- Aug,1961	0.629	IE cohort
Sep,1959- Aug,1960	0.610	IE cohort
Sep,1958- Aug,1959	0.536	IE cohort
Sep,1957- Aug,1958	0.486	IE cohort
Sep,1956- Aug,1957	0.483	IE cohort
Sep,1955- Aug,1956	0.446	IE cohort
Sep,1954- Aug,1955	0.423	IE cohort
Sep,1953- Aug,1954	0.330	IE cohort
Sep,1952- Aug,1953	0.257	IE cohort
Sep,1951- Aug,1952	0.236	IE cohort
Sep,1950- Aug,1951	0.164	IE cohort
Sep,1949- Aug,1950	0.107	IE cohort
Sep,1947- Aug,1949	0.100	IE cohort
Sep,1946- Aug,1947	0.044	Pre-IE cohort
Sep,1945- Aug,1946	0.021	Pre-IE cohort
Sep,1944- Aug,1945	0.010	Pre-IE cohort
Sep,1943- Aug,1944	0.011	Pre-IE cohort

[1] We employ a cutoff of 10 percent to define schooling cohort, and the assignment into treatment and control groups is similar to Meng and Gregory (2002).

Table 1.2: Descriptive Statistics : Siblings Sample

	(1)	(2)	(3)	(4)
		Mean (S.D.)		p-value
	All	Non-IE cohort	IE cohort	
Health Variables				
Self-reported health	2.947	2.829	3.187	<0.000
(1=excellent,2=very good,3=good,4=fair,5=poor)	(1.247)	(1.224)	(1.261)	
Physical Discomfort	0.288	0.256	0.346	0.046
(1=Yes,0=No)	(0.458)	(0.445)	(0.477)	
Chronic disease	0.142	0.089	0.239	<0.000
(1=Yes,0=No)	(0.355)	(0.297)	(0.428)	
Underweight	0.089	0.068	0.132	0.013
(BMI≤19,1=Yes,0=No)	(0.285)	(0.252)	(0.339)	
Obesity	0.225	0.239	0.198	0.275
(BMI≥28,1=Yes,0=No)	(0.418)	(0.427)	(0.399)	
Explanatory Variable & Instrumental Variable				
Schooling	7.457	7.802	6.735	0.005
	(4.452)	(4.368)	(4.550)	
CohortExposure	12.43	1.218	35.09	<0.000
	(19.74)	(1.234)	(20.17)	
Control Variables				
Age	44.80	41.88	50.71	<0.000
	(6.581)	(5.113)	(5.091)	
Male	0.849	0.850	0.846	0.893
	(0.358)	(0.357)	(0.362)	
Ethnic Minority	0.080	0.082	0.077	0.852
	(0.272)	(0.274)	(0.267)	
Urban	0.347	0.351	0.341	0.819
	(0.477)	(0.478)	(0.475)	
Urban (young-age)	0.098	0.101	0.093	0.792
	(0.298)	(0.301)	(0.292)	
Married	0.782	0.777	0.791	0.708
	(0.413)	(0.417)	(0.408)	
Retired	0.089	0.079	0.110	0.229
	(0.285)	(0.270)	(0.314)	
Co-habiting	0.895	0.889	0.907	0.519
	(0.307)	(0.315)	(0.292)	
N	550	368	182	

[1] Non-IE cohort Includes pre-CR cohort (1943.9-1947.9) and post-CR cohort (1966.9-1976.9).
[2] Column (4) reports raw (unconditional) p-value for a t-test of differences in means across cohort. Significant raw difference in *age* between the two groups disappears after controlling for birth cohort fixed effects.

Table 1.3: LPM and LPM with Sibling Fixed Effects

	Self-rep. health		Phy. discomfort		Chronic disease		Underweight		Obesity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Schooling	-0.041***	-0.025***	-0.012**	-0.014***	-0.005*	-0.007*	-0.008**	-0.013***	-0.004	-0.009***
	(0.011)	(0.009)	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)	(0.003)
Sibling fixed effects	N	Y	N	Y	N	Y	N	Y	N	Y
N	544	544	438	438	449	449	544	544	544	544

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

Table 1.4: LPM: Sibling Fixed Effects with Additional Controls

	Self-rep. health		Phy. discomfort		Chronic disease		Underweight		Obesity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Schooling	-0.026**	-0.025*	-0.019**	-0.011*	-0.017***	-0.015***	-0.010***	-0.008**	-0.013**	-0.011**
	(0.004)	(0.005)	(0.008)	(0.007)	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)
Age	0.018	0.039	0.010	-0.005	0.057	0.067	-0.013	-0.022	0.009	0.035
	(0.054)	(0.062)	(0.046)	(0.057)	(0.041)	(0.046)	(0.029)	(0.032)	(0.021)	(0.022)
Age square	-0.000	-0.000	0.000	0.000	-0.001	-0.001	0.000	0.000	-0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ethnic minority	-0.132	-0.189	-0.037	0.161***	-0.024	0.077	0.124***	0.049*	-0.071*	-0.048*
	(0.071)	(0.075)	(0.025)	(0.038)	(0.039)	(0.075)	(0.021)	(0.025)	(0.037)	(0.028)
Male	-0.138	-0.128	-0.144**	-0.133*	-0.102*	-0.119*	-0.091*	-0.111**	-0.022	-0.028
	(0.130)	(0.152)	(0.064)	(0.072)	(0.056)	(0.068)	(0.049)	(0.053)	(0.073)	(0.070)
Married	0.002	0.015	-0.023	-0.013	-0.067**	-0.079**	-0.001	0.002	0.004	-0.013
	(0.072)	(0.076)	(0.034)	(0.036)	(0.026)	(0.032)	(0.017)	(0.018)	(0.022)	(0.025)
Urban	-0.034	-0.066	0.071	0.076	0.247***	0.210***	-0.053***	-0.053**	-0.002	-0.029
	(0.049)	(0.072)	(0.061)	(0.076)	(0.047)	(0.055)	(0.020)	(0.024)	(0.051)	(0.057)
Urban (young-age)	0.374	0.288	0.267*	0.241*	0.235	0.226	-0.041	-0.025	-0.031	-0.054
	(0.263)	(0.213)	(0.154)	(0.135)	(0.177)	(0.157)	(0.053)	(0.058)	(0.049)	(0.062)
Retired	-0.444***	-0.429***	-0.067	-0.035	0.006	-0.011	-0.069***	-0.060***	0.452***	0.416***
	(0.055)	(0.064)	(0.092)	(0.108)	(0.095)	(0.087)	(0.020)	(0.012)	(0.054)	(0.064)
Co-living	0.065	-0.035	-0.034	-0.044	0.006	-0.000	0.006	-0.020	-0.087**	-0.071
	(0.174)	(0.195)	(0.060)	(0.060)	(0.050)	(0.063)	(0.032)	(0.034)	(0.040)	(0.044)
Sibling FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	N	Y	N	Y	N	Y	N	Y	N	Y
N	544	544	438	438	449	449	544	544	544	544

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy, whether siblings live together.

Table 1.5: Individual Heterogeneity

	Self-rep. health		Phy. discomfort		Chronic disease		Underweight		Obesity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Schooling	-0.024**	-0.025**	-0.019**	-0.020**	-0.017***	-0.017***	-0.010***	-0.010***	-0.012**	-0.013**
	(0.012)	(0.012)	(0.008)	(0.008)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)
Weight at birth	-0.042	-0.045	0.003	0.003	-0.005	-0.006	0.003	0.003	0.001	-0.002
	(0.032)	(0.031)	(0.010)	(0.011)	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)	(0.009)
Early-age health	0.116	0.103	0.017	0.017	0.078*	0.069	0.083***	0.086***	0.009	-0.004
	(0.255)	(0.242)	(0.075)	(0.075)	(0.043)	(0.042)	(0.027)	(0.029)	(0.045)	(0.041)
Alcohol and Cigarette	N	Y	N	Y	N	Y	N	Y	N	Y
Sibling Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	544	544	438	438	449	449	544	544	544	544

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy, whether siblings live together.

Table 1.6: 2SLS: First-stage Results

	Schooling			
	(1)	(2)	(3)	(4)
CohortExposure	-0.105***	-0.111***	-0.111***	-0.096**
	(0.039)	(0.039)	(0.041)	(0.046)
Sibling FE	Y	Y	Y	Y
Birth-cohort FE	Y	Y	Y	Y
Wave FE	N	Y	Y	Y
Controls	N	N	Y	Y
Province FE	N	N	N	Y
N	531	531	531	531
First Stage F-statistics	13.20	13.31	18.47	14.90
P-value for the F-tests	0.000	0.000	0.000	0.000
Over-identification P-values	0.197	0.267	0.313	0.143

[1] Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy, whether siblings live together.

Table 1.7: 2SLS with Sibling Fixed Effects

	Self-rep. health	Phy. discomfort	Chronic disease	Underweight	Obesity
	(1)	(2)	(3)	(4)	(5)
Schooling	-0.058**	-0.049**	-0.026*	-0.021**	-0.020
	(0.026)	(0.021)	(0.015)	(0.009)	(0.019)
Sibling FE	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y
N	524	421	421	524	524

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

Table 1.8: 2SLS: Sibling Fixed Effects with Additional Controls

	Self-rep. health	Phy. discomfort	Chronic disease	Underweight	Obesity
	(1)	(2)	(3)	(4)	(5)
Schooling	-0.058*	-0.051**	-0.026*	-0.019**	-0.019
	(0.034)	(0.021)	(0.015)	(0.010)	(0.019)
Early-life health	0.104	-0.048	0.097***	0.088*	-0.030
	(0.229)	(0.092)	(0.036)	(0.047)	(0.071)
Birth weight	0.029	-0.006	0.030*	0.012	-0.011
	(0.037)	(0.022)	(0.016)	(0.009)	(0.016)
Send-down	0.127	0.233	0.063	-0.088***	-0.250**
	(0.116)	(0.188)	(0.120)	(0.028)	(0.119)
Sibling FE	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y
N	524	421	421	524	524

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Covariates include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy, whether siblings live together, survey wave fixed effects and birth cohort fixed effects.

[3] A source of heterogeneity comes from a confounder along with the instrument, the send-down policy. During the Cultural Revolution, Chairman Mao also initiated the forced rustication movement by sending millions of urban middle school students (10-19 years) to the countryside. The scheme applied to all eligible urban individuals who would graduate from high school, exactly those who were exposed to the closure of schools. Being estranged from their family for several years, they experienced a dramatic decline in the standard of living, and consequentially, may have different health outcomes nowadays, as discussed by Xie et al. (2007).

Table 1.9: Robustness Check: Great Famine

	(1) Self-rep. health	(2) Phy. discomfort	(3) Chronic disease	(4) Underweight	(5) Obesity
Schooling	-0.070**	-0.072***	-0.039**	-0.025**	-0.029
	(0.033)	(0.023)	(0.019)	(0.013)	(0.023)
Great famine	0.335***	0.197*	0.127	0.029	0.145
	(0.118)	(0.112)	(0.110)	(0.052)	(0.114)
Sibling FE	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y
N	534	422	430	534	534

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy, whether siblings live together.

Table 1.10: Robustness Check: Placebo Test

	Reduced form: "Fake" IE cohorts vs. non-IE cohorts					
	(1)	(2)	(3)	(4)	(5)	(6)
	Schooling	Self-rep. health	Phy. discomfort	Chronic disease	Underweight	Obesity
CohortExposure	0.168	0.047	0.040	0.031	0.008	0.040
	(0.222)	(0.087)	(0.040)	(0.029)	(0.021)	(0.027)
Sibling FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Birth cohort & Wave FE	Y	Y	Y	Y	Y	Y
N	357	363	288	297	363	363

[1] Placebo sample selected here is consist of "fake" IE cohorts siblings and non-IE cohorts siblings. Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy, whether siblings live together.

Table 1.11: Mechanism: Intermediate Variables

	(1)	(2)	(3)	(4)	(5)
	Effect of	Income		Cognition	
	Schooling	Cond. Eff.	Explained(%)	Cond. Eff.	Explained(%)
Panel A					
Income	0.043**	-	-	-	-
	(0.018)				
Math test	0.215***	-	-	-	-
	(0.033)				
Word test	0.149***	-	-	-	-
	(0.047)				
Panel B					
Self-report health	-0.058	-0.022	62.59	-0.054	6.156
Physical discomfort	-0.051	-0.045	12.43	-0.038	26.14
Chronic disease	-0.026	-0.014	46.85	-0.023	10.97
Underweight	-0.019	-0.014	26.91	-0.018	6.06

[1] Panel A shows the impact of education, instrumented by cohort exposure during the educational disruption, on the potential intermediate variables: income (measured by logged monthly wage (yuan)), CFPS word test (z-score), CFPS number series test (z-score).
[2] Panel B shows the role of the above intermediate variables in effects of education on health outcomes.

Table A.1: Heckman selection model

	Self-rep. health	Phy. discomfort	Chronic disease	Underweight	Obesity
	(1)	(2)	(3)	(4)	(5)
Schooling	-0.025**	-0.012**	-0.008*	-0.009***	-0.000
	(0.013)	(0.006)	(0.004)	(0.003)	(0.004)
Panel B: First Stage					
Sibship Sample					
Household size	0.102***	0.101***	0.101***	0.102***	0.100***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
mills					
lambda	0.159	0.080	0.017	-0.038	-0.047
	(0.119)	(0.051)	(0.039)	(0.029)	(0.040)
Controls	Y	Y	Y	Y	Y
Wave effect	Y	Y	Y	Y	Y
N	62679	62573	62584	62679	62679

[1] Coefficients are estimated from Heckman selection model (ML). Standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.
[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), current urban dummy, marital status, retirement dummy, ethnic minority dummy.

Table A.2: Rural, Urban subsample

	Self-rep. health		Phy. discomfort		Chronic disease		Underweight		Obesity	
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
IE cohort	-0.048**	-0.015	-0.010	-0.019**	-0.031***	-0.011***	0.000	-0.009***	-0.013*	-0.017***
	(0.024)	(0.014)	(0.019)	(0.007)	(0.008)	(0.003)	(0.011)	(0.003)	(0.007)	(0.006)
Sibling FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Wave FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	191	353	172	266	173	276	191	353	191	353

[1] Coefficients are estimated from linear model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Control variables include age, age square, gender, urban dummy (at 12-year-old), marital status, retirement dummy, ethnic minority dummy, whether siblings live together.

Table A.3: Exclusion Restrictions

	(1) current hukou	(2) childhood hukou	(3) migration	(4) married	(5) retired	(6) co-habiting
CohortExposure	-0.002	0.001	0.005	0.006	0.002	0.003
	(0.004)	(0.003)	(0.004)	(0.011)	(0.004)	(0.003)
Sibling FE	Y	Y	Y	Y	Y	Y
Birth cohort FE	Y	Y	Y	Y	Y	Y
N	550	550	550	550	550	550

[1] Coefficients are estimated from OLS model. Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.

[2] Outcome variables include current hukou (urban dummy), hukou in childhood (urban dummy at 12-year-old), migration (whether the individual has experienced rural-urban migration), marriage dummy, retirement dummy, whether siblings live together.

Chapter 2

Good Bye Chiang Kai-shek? The Long-Lasting Effects of Education under the Authoritarian Regime in Taiwan

2.1 Introduction

Existing literature suggests that people's attitudes and ideology can be formed through a variety of channels: family (Bisin and Verdier, 2001), media and propaganda (Della Vigna and Kaplan, 2011), peer to peer (Sacerdote, 2001), etc. Education is among the most evident ones since individuals spend almost their childhood in schools and their core values are high likely formed during that period. In history, there are full of examples of dictators were using schooling in particular to control the outlook of children and young adults (Voigtlnder and Voth, 2015). As we have known, between 1974 and 1990, many countries ended the rule of authoritarian or autocratic regimes and shifted towards transitional or democratic regimes. However, the end of an authoritarian regime does not necessarily mean that its impacts disappear; it might still affect individuals' ideology throughout their lives. This raises a fascinating question: does an authoritarian political regime, which a person expe-

rienced at an early age, have long-lasting effects on that person's political outcomes (political attitudes, voting behavior, and national identity) through an educational channel?

In this paper, we would like to respond to this question by exploiting evidence from an authoritarian regime and its transition in Taiwan. Historically, the leader of the Kuomintang party (KMT), Chiang Kai-shek, took control of Taiwan in 1945 following the World War II. Then the KMT arbitrary seizure of private property and their economic mismanagement have motivated considerable protests, resistance, and campaigns against them, spontaneously organized by Taiwanese people (Rubinstein, 1999). The protests were violently suppressed by the KMT government, which killed thousands of civilians. Hence the politicians of the KMT realized that it was important to strengthen its control over people, especially youth: the KMT promulgated not only martial law to restrict human rights,[1] but also inserted pro-regime content to the curriculum and increased the number of patriotic activities for primary and secondary schools. The explicit goal of the KMT-led authoritarian educational system was to impose government ideology on students: requiring unconditional support for the KMT; promoting patriotic education, worshiping leadership as well as anti-communism; going counteroffensive against the mainland, etc. These goals found their way into every single school subject (Hsu, 2007; Tsai, 2007).

According to history, the authoritarian regime had controlled Taiwan for almost forty years until 14 July 1987, when former Taiwanese leader Chiang Ching-kuo announced the lifting of martial law, marking an official end of Taiwan's authoritarian regime.[2] Since that, people have regained freedom of speech and have been able to participate in political activities. Facing increasingly overwhelming criticism from the masses about the political indoctrination in primary and secondary schools, the Ministry of Education was aware of the necessity for changes in the content of the textbook along with the dramatic political regime transition. At the same time, more Taiwan-born people disliked the indoctrination in education for a long time and got the power in the government and the Ministry of Education, and they had a chance

[1]The KMT published strict regulations to prevent assembly, association, procession, petition and other forms of organised protest.

[2]It is an exogenous shock for people: no one could anticipate this event before he announced the lifting of martial law.

to change the "brainwashing education." From 1 September 1987, the Ministry of Education has required the schools sharply decreased the number of patriotic activities, deleted the indoctrination content mainly in teaching materials, and published the revised textbooks in 1989 (the revised textbooks, hereafter), and most of the political inculcation had been cut out.

Based on the history mentioned above, we examine the effects of the authoritarian education on political outcomes by exploiting the sharp decrease of ideological content in schools after 1987 in Taiwan. Moreover, we combine a regulation published by the government which requires children who were six years old on or before August 31 of a given year, to enroll in elementary school by September 1 of that year (Spohr, 2003; Tsai, 2009; Tsai, 2010). Those who were born after September 1 would only be enrolled in September of the next year. Applying this regulation, our identification utilizes cut-off birth dates for elementary school enrollment that cause variation in the length of exposure to the authoritarian education system within the same birth cohort. Following Fuchs-Schündeln and Masella (2016), we employ this strategy that relies on comparison for people born on the same year, which helps us not to face the difficulty of dealing with unobserved differences across cohorts. We further include regional fixed effects to control for non-time-varying confounding variables at the regional level.[3]

The findings provide new evidence that an additional year of exposure to authoritarian education has significant effects on the outcomes of interest: it decreases preference for democracy and election turnout rate while increasing the probability of supporting the KMT, and reduces the likelihood of self-identifying as Taiwanese, rather than Chinese. There is no significant difference between control and treatment groups regarding attitude toward unification with China, because there was almost no variation of relevant content in schools before and after the lifting of martial law. When the whole sample is divided into two parts by gender, we can see that the coefficients for men remain significant, but for women, the long-term effects disappear. This could potentially be attributed to the fact that women had been excluded from political participation since the beginning of feudal times in China, and this

[3]Taiwan has 22 administrative divisions, including special municipalities, cities and counties. For easier understanding, we refer to administrative divisions as regions.

cultural heritage persists over time. Hence receiving one more year of authoritarian education does not have a significant effect on women who are the same age.

In the last part of the empirical analysis, the robustness of the results is tested. First, we deal with the argument that there might be other factors from society affecting the treatment group apart from the indoctrination in education. We use a placebo test by using pre-cohorts (those born between 1960 and 1968), who completed all of their education (nine-year compulsory education and the senior secondary education) under the authoritarian regime. According to the results, none of the coefficients of interest are statistically significant. Second, this paper performs a sensitivity test by including more control variables, such as region-cohort fixed effects, to account for region-specific cross-cohort trends. However, the results remain unchanged.

This paper contributes to the literature on three grounds.

First, this study contributes to the literature about the long-lasting impacts of the institutions on individual outcomes. Alesina and Fuchs-Schündeln (2007) argue that the communism in Germany affects the individual's preferences in the long run. Fuchs-Schündeln and Masella (2016) find that the socialist education in East Germany could decrease the likelihood of getting a college diploma and also has an impact on labor market performance for males. In this study, we explore the causal relationship between the authoritarian regime experience and political outcomes. To the best of our knowledge, the long-term effects of the political institutions on individuals' political outcomes are insufficient. Pop-Eleches and Tucker (2014) investigate the effect of the communism on political attitude and conclude that more exposure to communism leads to more opposition to democracy and capitalism. However, they do not provide a clear mechanism about how the regime could affect the individual's political ideology. And we fill the gap by using the educational channel in this paper.

Second, this paper relates to the recent research on the effects of curriculum and textbooks changes. Clots-Figueras and Masella (2013) study the effect of a bilingual (Spanish and Catalan) education reform and find that respondents who have been exposed for a longer period to teaching in Catalan have stronger Catalan feelings. Chen et al. (2016) study the 1994 curriculum reform for junior high school social subjects that introduced vast amounts of Taiwan-related content and conclude that students exposed to the revised textbooks are more likely to hold a stronger Tai-

wanese identity. Cantoni et al. (2017) explore the causal effect of a new curriculum on students' ideology in China, and find that the new curriculum effectively shape students' beliefs and attitudes toward the Chinese Government.

This paper is close to the Cantoni et al. (2017), but the differences are as follows. Instead of a short-term effect (about four years in their paper) of curriculum changes, we focus on the long-term implications (about 23 years on average) of the variations in the curriculum. This contributes to the existing literature since limited study focus on the persistence of political preferences which last for more than 20 years. Second, we use a representative sample of the population, covering different educational backgrounds, while in Cantoni et al. (2017), the analysis is conducted on the sample of students from the best Chinese universities. The students are extremely good at studying. Given that selection issues might be present, it is unclear whether the results would hold if the subjects are students selected from non-elite universities (non-Project 211 universities).[4] For our study, however, this is not a concern. One significant advantage is that our dataset covers people with different educational backgrounds, which is much more representative compared to the elites sample used in the Cantoni et al. (2017).

Finally, this paper sheds light on historical reasons for the current Taiwanese attitude toward national identity and cross-strait reunification. If one wants to study the long-term impact of authoritarian regimes, there are many potential candidates, for example, countries in South America that experienced a wave of democratization in the late 20th century. So why we choose Taiwan? Because Taiwan is explicitly selected because one of our primary objectives is to find out the source of heterogeneity in national identity and reunification attitudes. It is well-known that historically Taiwan has been ruled by China since the Song Dynasty, and was under Chinese feudal rule for several centuries. But in 1949, the KMT lost the civil war, retreated to Taiwan and established a new regime opposed to that of the Communist Party of mainland China. From then, the two sides of the Strait have different histories. Especially since Taiwan transitioned to a democratic society, the difference in

[4]To raise the research standards, from1995 the Ministry of Education of China launched a project about establishing National Key Universities, which is Project 211. Now 116 higher education institutions are designated as Project 211.

political attitudes of the people in Taiwan and Mainland China is becoming more and more apparent. In recent years, a significant proportion of people in Taiwan identify as Taiwanese, rather than Chinese, and believe that Taiwan should be independent. The notion of national identity and political policy toward mainland China (independence or reunification) are still a quite crucial question in the light of Taiwanese leader elections. By exploiting a historical source of heterogeneity in national identity and cross-strait reunification attitude, this study makes a groundbreaking contribution to the literature on cross-strait relations as well.

The paper is structured as follows. A brief introduction of the institutional background and the educational system in Taiwan is given in section 2.2. In section 2.3 the data and the variables employed are discussed. The empirical analysis and results are presented in section 2.4. The paper concludes with section 2.5.

2.2 Institutional Background

2.2.1 Political Regime Transition in Taiwan

The point of the political regime transition in Taiwan is critical for this study, since estimating the impact of the political regime requires criteria to distinguish different regimes. Based on standard practices in the political economy literature and the discussion in Besley and Kudamatsu (2007), we use the Polity Score index from the Polity IV dataset as demarcation criterion (Marshall and Jaggers, 2005).[5]

Taiwan is a small island. Aborigines had inhabited it before the Han migrated to Taiwan. In 1895, the Qing army lost the Sino-Japanese War and ceded Taiwan to Japan. A half-century later, the Japanese surrendered to the Allies in World War II, and the KMT, led by Chiang Kai-shek, took control of Taiwan. Soon after the KMT occupied Taiwan, the corrupt conduct on the part of the KMT authorities and their massive mismanagement became evident to the local population. These contradictions led to an anti-KMT campaign around the whole island on February

[5]The "Polity Score" shows a regime authority spectrum ranging from the hereditary system (-10) to the full democracy (10). We could classify regimes in different categories: "autocracies" (from -10 to -6), "anocracies" (from-5 to 5), and "democracies" (from 6 to 10).

28, 1947. The uprising was violently put down with more than 3,000 Taiwanese elites being killed by the KMT military (this event is known as the "228 Incident"). Then the KMT enforced martial law to strengthen its control over the people after 1949, where the authority trend of Figure 2.1 starts.[6]

It is observed in Figure 2.1 that from 1949 to 1987, the authority has a value below -5, which indicates that Taiwan was in the authoritarian regime led by the KMT government. The autocracy in this period was featured by the fact that Chiang Kai-shek started one-party rule and announced the imposition of martial law in Taiwan. In 1975, after Chiang Kai-shek's death, his son Chiang Ching-kuo took over the power. Although Chiang Ching-kuo intended to democratize the political system gradually and to replace the older generation of the KMT with more politicians of Taiwanese origin, the martial law was still valid. People's rights and freedoms were largely limited. Finally, Chiang lifted martial law on July 14, 1987, and it was the end of the authoritarian regime in Taiwan. As evident from the graph, the value jumps dramatically from -5 to about -0.5 in that year. Between 1987 and 1992, Taiwan experienced a transitional period and people regained fundamental rights, such as freedom to establish political parties. On the 19th of December 1992, the election for the second Legislative Yuan was held in Taiwan, which was the first direct legislative election in Taiwan. At this very moment that Taiwan began its transition to an effective multi-party electoral system. From that year Taiwan enters the era of a democratic regime with a score higher than 6, and after 2004, it has a fully democratic government.

2.2.2 Political Indoctrination Changes in Primary and Junior High Schools

As we mentioned above, after the 228 Incident, to consolidate the stability of its rule, the KMT promulgated a series of relevant regulations to restrict civil rights, such as citizens lost freedom of speech, assembly, association, etc. In addition to imposing martial law, to ensure that the party's propaganda could have an active and lasting impact on students, the KMT emphasized control through education,

[6]I select the figure from Political Regime Characteristics and Transitions Project (1800-2013).

mainly primary and junior high schools. Related document (Tsai, 2007) confirms that the educational policy adopted by the KMT education administration committee explicitly pointed out the significance of party-oriented education:

> What we call "party-oriented education" is that under the guidance of the KMT, education should be in line with the party's values and ideology... When we start to design the party-oriented educational system, we should restructure the curriculum so as not to violate the party's principles... And we can give full play to the role of the party and its policies. We should quickly advance the reviewing and compiling of textbooks, which should be a combination of the party's ideology and educational purposes.[7]

All humanities, including Chinese and social science courses, were China-centered and KMT-oriented, strictly determined and audited by the KMT government. Textbooks, exams, degrees, and educational instructors were all controlled by the KMT officials. Additionally, the Ministry of Education established a new department called the "National Institute for Compilation and Translation," which invited many professors and educational specialists to participate in the textbooks and curriculum design and evaluation. Moreover, the government also arranged for "reliable" staffs, such as officials and experts who met the needs of the KMT's official ideology and decision-making, to audit teaching materials (Su, 2007).

Education under the KMT government had a unique feature not easy to perceive. The course "Citizenship and Morality" for elementary students is an illustrative example: when we examine the curriculum of this course, there is no significant difference between moral or civic classes in the KMT ruling period and those in the democratic regime or transition period. However, if we carefully examine the content of the textbooks, we can find it contains a lot of content related to authoritarianism, which is quite different from the description of the curriculum. The textbooks explain the concepts of citizenship and nations, as well as basic knowledge of laws

[7]Beginning in 1924, the Kuomintang had a plan to carry out a "party-oriented education" policy in mainland China. Two years later, the KMT government formulated a draft of this policy. When the KMT retreated from Mainland China in 1949, they began to implement this policy in Taiwan.

and regulations citizens should obey. However, the book includes little about the democratic regime and how it works. The purpose of education under the KMT government was to cultivate students to become persons who obey the law and government decisions, not to develop their ability to participate in that decision-making (Hsieh, 1988; Ou, 1990). Moreover, the KMT defined what constituted "citizenship and morality," not the laws and values accepted by an entire society, which is different from democratic countries' civics education. Anything relating to democratic regimes, citizen rights and Taiwanese identity was reduced to a minimum, or even forbidden to be introduced. Students were deprived of learning about democratic values and Taiwanese history.

Moreover, the KMT defined what constituted "citizenship and morality," not the laws and values accepted by an entire society, which was different from democratic countries' civics education. Anything relating to democratic regimes, citizen rights and Taiwanese identity was reduced to a minimum, or even forbidden to be introduced. Students who received the compulsory education were denied learning about democratic values and Taiwanese history.

The liberalization and democratization of Taiwan arose after the former leader Chiang Ching-kuo announced the lifting martial law. More and more people participated in political activities, and all circles of society began to criticize the regime. From an educational point of view, people were starting to express their dissatisfaction with the political indoctrination contents in the textbooks. Criticism was combined with various political movements, and a wave of textbooks reform followed (Zhang, 2005). More Taiwan-born people who disliked the indoctrination in education got the power the Ministry of Education and in the government, and they had a chance to change the "brainwashing education". From 1987, the Ministry of Education had revised the elementary and junior high school textbooks and published a revised version of the textbooks. There was much less political content in the revised version of the books,[8] as the teaching materials writing group reduced the content praising the KMT, authoritarian regimes, Chinese mainland culture, and a sense of

[8]Between 1987 and 1989, because there was not enough time to delete the indoctrination content mostly, the elementary and junior high schools still used the old textbooks with a minor change: the indoctrination content reduced, but not largely.

Chinese national identity, while incorporating more aboriginal cultures as well as western democracy knowledge.

Distinguishing the standard teaching content and political indoctrination in the textbooks is complicated. It is worth noting that from 1992 the pedagogical academia of Taiwan began to measure and compare the intensity of political indoctrination under authoritarian and transitional regimes, and published a series of pedagogical literature is considered very informative and reliable. Scholars have used various methods to identify political indoctrination and a large amount of pedagogical literature (Lin, 1993; Wang, 1996; Tsai, 2007; Lau, 2000) studied the political inculcation of compulsory education during martial law period, distinguishing standard and ideological parts by the phrases, sentences, statements, examples, and stories that appeared in the textbooks. In general, experts divide the texts into different paragraphs according to the content, then determine whether a paragraph contains ideological content. If yes, then scholars decide which type of political socialization the paragraph belongs to, and record the number of the occurrences.

According to Lin (2000), Lau (2000), Wang (1996) and Tsai (2007), by comparing the content of ideology in the social science textbooks of elementary school during martial law period (1978 edition, "the old textbooks") and the revised textbooks, the occurrences of topics related to patriotic sentiment, authoritarian structure, worship of political leaders, anti-mainland sentiments, spiritual constructions, authoritarian political values, and Chinese culture were significantly less after the lifting of martial law.[9] Figure 2.2 below shows that a vast majority of the ideological content was reduced in the revised textbooks.[10]

The vertical axis represents the frequency of ideological content in the old and revised textbooks, and the horizontal axis represents the type of ideological content. The blue histogram stands for the old textbooks, while the orange histogram represents the revised books. It is evident that all the types of political ideology content

[9]Other scholars use different political subject classification, but the conclusions are consistent with Liu et al. (2000), suggesting that the political ideology in the textbooks has decreased dramatically after 1987.

[10]There is no information on the variations happened in each grade, hence the bars show average changes of the political ideology content of primary school.

have reduced sharply in the revised textbooks.[11]

Of particular interest here is the content of the textbooks related to Taiwan: the revised textbooks for primary and junior high school students, the emphasis was still on China's geography and history. However, the material about the history of Taiwan and political and cultural aspects of both sides of the Taiwan Strait has been greatly increased. The newly added content on Taiwan may provide students a different understanding of national identity.[12] In summary, the dramatic changes to the curriculum along with the political transition allows us to observe a difference between treatment group and control group, as explained in the next section.

2.2.3 The Educational System in Taiwan

This section introduces the educational system briefly in Taiwan.

The government extended compulsory education from six to nine years after 1968, and all school-age children are required to attend and finish primary school (6 years) and junior high school (3 years). This compulsory rule enables us to keep almost all the observations in the surveys, given the fact that all individuals, both in the treatment and control groups went to school and were exposed to the authoritarian educational system. Eliminating the concern that some people in the treatment group were not in fact treated because they did not even go to school.

Another feature of compulsory schooling in Taiwan that relates to our study is that the new semester begins on the 1st of September. According to Spohr (2003) and Tsai (2010), children who are already six years old before September 1 of a given year must be enrolled in first grade in that year. Similarly, students who are 12 years old before or on August 31 must be enrolled in junior high school in September of

[11] Although scholars have not yet compared the changes in the political ideology of junior high school social subjects textbooks after martial law was abolished, several studies (Tsai, 2007; Xue, 2000) have analyzed the changes in the revised textbooks "Civil and Moral (Politics)" and find that in the 1989 version, the content of authorities and anti-mainland sentiments have drastically declined. Among them, there has been a drastic reduction in the content of the KMT leader worship.

[12] It should be noted here that the 1989 edition is a blueprint for the 1994 new curriculum and experimental teaching materials. The 1994 new curriculum contains a vast material understanding Taiwan, and has an even greater impact on students' national identity (Chen et al., 2016).

that year.

2.3 Dataset and Variables

The micro-level data used in this paper is from the Taiwan Social Change Survey (TSCS), which is a longitudinal survey that tracks political, socio-economic, and cultural changes in Taiwan. This interdisciplinary cross-sectional survey, as one of the largest survey series among general social studies in the world, began in 1985, enabling us to understand its socio-economic change from long-term perspectives. Main topics across surveys include political participation, national identity, family status, economic attitudes, social networks, physical and mental health, religion, etc. By 2016 the TSCS had accumulated 58 surveys with approximately 121,000 face-to-face interviews. To summarize, TSCS is considered one of the most authoritative datasets in Taiwan.

2.3.1 Outcome Variables

We select nine waves from the TSCS dataset, which contain the following outcome variables of interest, then choose a group of questions from the TSCS dataset, which asked respondents many questions about an individual's political attitude and national identity.

Preference for Democracy The main variable of interest is the preference for a democratic regime, which comes from the following survey question: "Which of the following three statements do you agree with? a) Democracy is preferable to other forms of government; b) Under some circumstance, an authoritarian government can be preferable to a democratic one; c) For people like me, it does not matter whether we have a democratic regime or not."

We create a dummy *Pref.Democracy* that equals one if a respondent answers a), and zero if a respondent answers b),[13] to measure whether an interviewee has

[13]Following Burm (2018), we also create a dummy that equals one if individual answers a), and zero if a resplendent answer b) and c) , the result does not alter much and remains significant.

a preference for democracy over any other kind of regimes. During the period that martial law was enforced, most of the content in Politics and Chinese textbooks complimented the benefit about authoritarian regimes (Tsai, 2007; Lau, 2000), for instance: "having one ruler or limited numbers of elites to decide on affairs regarding a country and without influences of others parties makes it easier for the ruler to make decisions in the interest of the nation. Moreover, in emergency cases where quick action is required, decisions can be carried out immediately". After martial law was lifted, the educational specialists who designed the textbooks gradually replaced such content advocating authoritarian regimes gradually with the introduction of democratic values.

Party Affiliation To measure the outcome of an individual's party affiliation, *SupportKMT* is a dummy created from the response to the question: Which political party do you support? The answer includes main political parties in Taiwan nowadays: the KMT, the Democratic Progressive Party (DPP), the Taiwan Solidarity Union (TSU), the People First Party (PFP) and so forth. The variable *SupportKMT* takes the value one if an interviewee self-reports as a KMT supporter, zero as a non-KMT supporter.[14] This variable is closely related to the ideological content before 1987: there was much pro-KMT content in the old textbooks, which required students to "unswervingly follow the absolute leadership of the KMT, follow to the directions set by the party and obey its command (Tsai, 2007; Chen, 2006)". The Ministry of Education deleted a large percentage of this pro-KMT content between 1987 and 1995.

Political Participation For political participation measures, the TSCS asks respondents whether they turned out and voted, exercising their fundamental rights. The value of the binary variable *Turnout* is based on the response to this question: Did you turn out to vote in the most recent Taiwanese leader election? If the answer is yes, we set the value of Turnout as one, otherwise it is zero. In the transition pe-

[14]Other related variables include "In the last Taiwanese leader / legislator election, which candidate(s)/party did you vote for?" We have created a series of dummies that equal to one if the respondent voted for candidate(s) from the KMT or gave their party vote to the KMT, zero for other party candidates. The results are reported in the appendix.

riod, the revised textbooks covered basic knowledge of elections, and students were taught to participate in class committee elections, but the old textbooks did not provide information about what is a free and fair election, how citizens can select government leaders, or the set of policies that the government will follow. The notion of democracy taught from the book was partial and limited (Xu, 2016). After lifting of martial law in 1987, knowledge about western democracy was introduced, such as respect for political rights and civil liberties, majority franchise.

National Identity and Attitude towards Unification The ordinal variable *TaiwaneseIdentity* is given the value of two when the respondent choose Taiwanese, 1 if one declares he/she is both Taiwanese and Chinese, and zero if they choose Chinese. In the old curriculum, politics, history, and geography courses focused more on China and taught students that they were Chinese citizens. It also claimed that the aborigines were all formerly residents of Mainland China. After 1987, the proportion of Chinese ideology and culture in the textbooks were replaced gradually by content about Taiwan, which included awareness of the ethnic and cultural diversity of aborigines.[15]

Similarly, the opinion about *Unification* comes from the following question: "Do you support unification with China?" $Unification$ takes the value two if the respondent chooses "Unification", value one for the option "Keeping the status quo,"and zero for "Independence." The old and revised textbooks compiled and published in the 1980s always highlighted the legality of the Republic of China (ROC) and illustrate that the KMT government has sovereignty over all of China, including both Taiwan and Mainland China. Moreover, before 2000, no information related to Taiwan's legal independence from China appeared in textbooks so there should be no variation between the old and revised textbooks.

[15]As mentioned earlier, the topic Knowing Taiwan was introduced not only by the revised textbooks but also by the subsequent 1994 edition, which also resulted in variations in self-declared national identity depending on the curriculum to which that individual was exposed.

2.3.2 Explanatory Variable

The variable of interest is constructed by exploiting the cut-off date of birth for school enrollment that results in variation in the length of exposure to the KMT's authoritarian regime within the same birth cohort, following the method by Fuchs-Schündeln and Masella (2016).

Children turning six before September 1 of a certain year are allowed to enroll in primary school by September 1 of that year. Hence, we define individuals born in year b before September 1 as the treatment group, and those born after September 1 as the control group. The difference between students in the treatment and control groups is that for any birth cohort b still in education at the time when martial law was lifted and subsequently strong political indoctrination was reduced in July 1987, treated interviewees are one year more advanced in school, and thus one year more exposed to an authoritarian educational system than non-treated respondents who enrolled in schools one year later. By comparing individuals born early and late in the same year and still in schools when the transition of regimes occurred, we compare groups affected differentially by the length of exposure to authoritarian education.

2.3.3 Data Statistics

We restrict our sample to respondents who were born between 1972 and 1980, inclusive. People who were born in or after 1981 did not experience any education under the authoritarian regime. Similarly, individuals who born in or before 1971 had their compulsory schooling entirely with political indoctrination. Only observations of individuals born in this time span allow us to exploit the variation of exposure within birth-cohorts.

After defining treatment and control groups, data statistics that describes the socio-demographic features of the survey respondents is presented in Panel A of Table 2.1. Columns 2-3 show the main characteristics of respondents in the treated and untreated respondents. We further check for the balance of observable demographic characteristics across these two groups. In Panel A, columns 4 and 5, we show the raw differences, and the p-values testing for the statistical significance of these raw

differences in features of interviewees who stayed in authoritarian education for one more or one less year in our sample. We can see the samples in two groups share very similar social economic and demographic characteristics.[16]

In Panel B of Table 2.1, the statistics of our outcome variables of interest are presented. It is worth emphasizing that for outcome variables there are statistically significant differences between the treatment group and the control group.[17] The results suggest that our main dependent variables (except for variable *Unification*) are significantly different among treated and control groups whose observable characteristics look identical, indicating that the difference is very likely to be driven by the treatment.

2.4 Empirical Analysis

Using the survey data, we estimate a model that includes birth cohort, region of origin and wave fixed effects, and examines the impact of the length of being exposed to the KMT-led education. The baseline specification is the following:

$$Y_{ibcw} = \alpha + \beta Treat_i + \eta_b + \delta_c + \phi_w + \gamma X_i + \epsilon_{ibcw}$$

where Y is the outcomes of the individual i, and $Treat_i$ is a dummy variable that equals 1 if the respondent was born before September 1. X_i represents a vector of demographic and socioeconomic characteristics of individuals listed in Table 2.1. Birth cohort fixed effects η_b capture any common characteristics held by all persons in the same birth year cell. These fixed effects account for cohort-specific beliefs and preferences: this would include, for example, cases in which younger cohorts may be "more democratic" than older cohorts. Birth cohort fixed effects capture gener-

[16] In columns 6 and 7 of Panel A we show differences between interviewees in the two groups, conditional on birth cohort and region fixed effects, and the p-values testing for the statistical significance of these conditional differences. Conditional on common features in the administrative division of origin (measured by the region they lived in at the age of 15) common characteristics of a cohort, those respondents are nearly identical on observable characteristics.

[17] Even when region and birth cohort fixed effects are controlled for, the differences between main outcome variables still remain significant, as shown in columns 6 and 7, Panel B. The findings from these two columns are quite impressive.

ational, slow-moving changes in beliefs and preferences. Still, this implies assuming cohort specificities are common across all regions. Likewise, region fixed effects δ_c account for common characteristics held by individuals from a given region, irrespective of the year in which they were born. These fixed effects account for differences across counties that do not change over time and that are common across all cohorts born in the region, which includes, for instance, differences in folk custom, religion, indigenous heritage, etc. ϕ_w stand for wave fixed effects, which capture the trend in beliefs and preferences of the whole island in the survey years. ϵ is an error term.

The baseline model allows addressing of a set of concerns on the identification of the causal effects of exposure to education in an authoritarian background. A first concern is a possibility of political preferences evolving and changing across cohorts independently of political regimes. As stated before, the inclusion of birth cohort fixed effects eases this concern, though it implies restricting the cohort coefficients to be the same across Taiwan. Likewise, region fixed effects account for structural differences across counties that do not alter over time and that are common across all cohorts born in the region. Still, there may be idiosyncratic shocks that affect only certain counties in given years, so the identification strategy is based on variation at the region-cohort cell: a shock that has an impact on a specific region in a specific year would threaten the identification strategy only if it has differential effects across cohorts. We use controls for region-specific, cross-cohort trends in the robustness checks.

A second concern arises from parents' political preferences, which means children's political attitudes do not come from schools, but also from or mainly from their parents' political preferences. We also included the dummies whether the parents are native Taiwanese (people who lived in Taiwan prior to the KMT's took over from Japanese rule, most of whom were dissatisfied with the policies of the KMT) or Mainlanders (people who moved from mainland China to Taiwan after 1945, most of whom are faithful supporters of the KMT) as control variables, which are good proxy variables for parental political preferences. After controlling for these variables, the result does not alter and remains significant. Hence parental preference factors do not threaten the identification strategy.

There are some challenges remaining, such as the unobservable effects from the

society that might have larger effects than the variation in the educational program. We will explain these in detail in the robustness checks section.

2.4.1 Baseline Results

We run probit and ordered-probit models instead of a linear probability model (LPM) on all outcomes for the following reasons. First, the specification is very likely to be non-linear: intuitively, the difference between 0 and 1 year of exposure might be lower than the difference between 8 and 9 years of exposure. Second, two dependent variables are ordinal variables, and using a linear probability model complicates the interpretation of the findings. All results stay robust when a LPM is estimated.

Tables 2.2 and 2.3 show the results of the main specification, controlling for the variables listed above. Instead of raw coefficients obtained from the regression, we report the marginal effects, namely dF/dx, which are directly interpretable, with the two-way cluster-robust standard errors in parentheses. Being exposed to one more year of authoritarian education is found to yield a marginally significant negative impact on an individual's preference for democracy, reducing the probability of reporting that democracy is preferable to other forms of the regime by 6%, as shown in columns (1)-(2) of Table 2.2. The estimation results presented in columns (3)-(4) indicate that being in the treatment group on average decreases the likelihood of turning out to vote in the last election by around 4%-5%, showing a strong negative effect on political participation, which is in line with the finding of reduction in satisfaction with democracy. In contrast to the first estimates, one additional year of authoritarian education increases the probability of supporting the KMT by 3%, indicating direct indoctrination in education at a young age does affect the formation of an individual's party affiliation. In Table 2.3, other estimates are shown with the order-probit model. From columns (1)-(3), we can see that the treated group is 3.8% less likely to identify as Taiwanese, while 3.2% more likely to self-report as dual identity, and 0.6% more likely to report as Chinese, compared with interviewees in the control group. From the results above, it is inferred that since the textbooks used during the authoritarian regime fostered the formation of a pro-ruler ideology and behavioral patterns, those who enrolled one year earlier in primary and junior

high schools were exposed to a more prolonged impact of indoctrinated education, which had a long-lasted impact on their political attitudes and preferences.

Concerning unification with China, there is actually no difference between the two groups. There were no changes in this notion before and after 1987 in schools. Even in the society, few people held the concept of independence, so from the results, we can see the coefficients of our interests are insignificant.

We also report the marginal effect of control variables to have a brief look at which variables correlate with the formation of one's political attitudes. It seems parental education does not play an important role, while parental origin does. Children of parents of Taiwanese origin are less likely to support KMT and instead are more likely to identify as Taiwanese and have a positive attitude towards independence. We safely infer that by controlling for Taiwan-born father (father is native Taiwanese) and Taiwan-born mother (mother is native Taiwanese), we capture the intergenerational transmission of political attitudes. For other control variables, such as age, gender, etc., the direction of their effects on outcomes fits our expectations.

When our observations are split into males and females, as shown in Table 2.4, the effect for men remains significant, but for women who were also in schools and exposed to the authoritarian content of education, the long-term effects disappear. According to the results, there are no apparent differences between women born in the same year in two both groups. The reason could potentially be attributed to the fact that women had been excluded from political participation since the beginning of feudal times in China, and this cultural heritage plays a long-term role. Besides unfavorable cultural biases, women are more likely to face practical barriers to entering politics, including lower levels of education, less access to political information, larger burden of family responsibilities, and a deprivation of rights that has left them with fewer opportunities to acquire political experience. Therefore, women who were born in the same year are not very sensitive to different intensity of treatment.

Another concern needs to be addressed here: after the lifting of martial law in 1987, the number of patriotic activities in schools reduced sharply, but the textbooks were not immediately revised. Instead, the revised version textbooks were fully implemented in 1989. However, someone might argue that the changes in the

educational program do not play an essential role between 1987 and 1989.

To rule out this threat, we provide quantitative evidence. We choose 1989, the year in which the revised textbooks were published, instead of 1987, as a cut-off. Then we select cohorts born from 1974 to 1980 to run the baseline regression.[18] The results are shown in Table 2.5. Only a few estimates remain significant, indicating that the reduction in political propaganda activities plays an important role. Moreover, the results also show that samples within the same birth cohort who received different treatment between 1987 and 1989 are very crucial to our findings. Excluding them leads to the insignificance of some estimates. Similarly, we choose those who were born between 1973 to 1981 (the year 1988 as a cut-off) to run the same regression, the outcomes of which are reported in column (3) and column (6) of Table 2.5. The result of the exercise shows that, though the introduction of the revised textbooks in 1989 has different effects on the two groups, the total impact is not as significant as suggested by the estimation using 1987 as a cutoff. To sum up, we can infer that from 1987 to 1989, although the textbooks have not mainly changed, the number of patriotic activities in schools indeed decreased: the elimination of a great deal patriotic and party propaganda activities indeed has the long-term effects on political outcomes.

There is also a concern about heterogeneous about the sample affected by changes in the educational program. Due to the limited number of observations in each grade, we classify two major grade groups (To choose those from first to fourth grade when the martial law was lifted in 1987 as a junior group; from the fourth to ninth grade as a senior group). From Table 2.6 we know that for the most variables: the impacts of the revised textbooks have significant effects on the senior group. These kids (10-15) were not only older than those in the junior group but also had a better understanding of the ideological content of the textbooks, which played a vital role in shaping students' worldview and personal ideology.

[18] In the baseline model, we choose the year 1987 as a cut-off so that we restrict our sample to respondents who were born between 1972 and 1980; now we use the year 1989 (two years later) as a cut-off, so now we choose the individuals who were born between 1974 and 1980. To approximate the number of observations as in the baseline model, we also select those who were born between 1974 and 1982.

2.4.2 Confounding Effects

Omitted variables may correlate with both our interested variable and political outcomes, such as human capital accumulation, and these factors may confound the effect of different intensity of treatment. People may argue that there exist other channels instead of the ideological content in textbooks that link enrollment age and outcomes of interest, and name several candidates: years of schooling/the level of educational attainment, income, and types of occupation. Nevertheless, we still can perform a sensitivity test by controlling these variables. Adding these variables, if the explanatory variables of interest lost their significance, it suggests that the effects are driven by the revised textbook which leads to better human capital achievement rather than the decreasing content of authoritarian regime. However, after controlling for these individual-level characteristics, the coefficients of our interest do not lose significance in Table 2.7 so that we can rule out these potential threats. Moreover, we also use years of schooling, earnings and job type as outcome variables. From Table 2.7, we could see that the coefficients of main interest are insignificant, and most of them are close to zero.[19]

2.4.3 Robustness Checks

In this part, we attempt to rule out other alternatives to our interpretation of estimates, and verify the robustness of the results to solve the issues addressed above: 1) Heterogeneous impacts of time-varying shocks; 2) The confounding variables.

Including region-specific cross-cohort trends One may still argue that time-varying elements/shocks which happened in different regions might have different effects within the same birth cohort. Moreover, even in the same region, the effects of exogenous shocks on different cohorts might be different. A vivid example of such a shock is the so-called "Henry Liu Case", which took place in 1984. Henry Liu was a vocal critic of the KMT. He was assassinated in the United States by

[19] Another argument is the age trends drive the effects above. People who were born at the beginning of the year (Jan.) are different from those who were born in December because they might have various age trends. Hence, we allow for age trends by including birth month information. After we put birth month gap as controls, the results do not alter (results are available upon request).

Bamboo Union members trained by the KMT military intelligence. After the news came to Taiwan, people were disgruntled by the assassination of Democrats. Due to imperfect informational transmission at that time, it is likely that students in Taipei were better informed about the event than students in Kaohsiung; and even in the same place, the shock might have different effects on various cohorts: senior students were likely to be more interested in the news than juniors. Independently from whether the region-cohort variation caused by exogenous shocks is sharp or smooth, they represent unobserved factors that may threaten the identification strategy. To explore this possibility, we estimate the model that contains a full set of region-birth cohort fixed effects, which control for these region-specific, cross-cohort trends in political outcomes across cohorts as shown in Table 2.8. We can see that controlling for region-specific cross-cohort trends does not qualitatively affect any estimates of the effects of changes in political indoctrination in the textbooks.

Unobservable omitted variables It might be the case that some confounding factors, such as changes in social environment have larger effects on political outcomes than changes in the revised textbooks. Trying to address this issue, we use a sample of those born between 1960 and 1968 as the placebo group. The samples in the placebo group were fully exposed to education under the authoritarian educational system and had already finished primary or junior high schools when martial law was lifted. Therefore, they were not affected differentially by the authoritarian education but would face similar factors driven by these confounding variables from society. According to the results of Table 2.9, none of the coefficients of interest are significant for the placebo groups.

2.5 Conclusion

This paper sheds light on the long-run impacts of political regimes on individuals' the political outcome, exploiting evidence from the rapid reduction in the party-oriented content of the educational system in Taiwan. The historical event provides a unique setting to understand how the political regime affects individuals through education. Based on the Taiwan Social Change Survey (TSCS) dataset, we find that

the features of the authoritarian education system to have a long-lasting influence on those students, who are now their middle age. Our results suggest that one more year of exposure to ideological education reduces preference for democracy and election turn-out rate while increasing the support for the KMT. Meanwhile, as the old program contain much more Chinese ideology, those who received an additional year with the old program would consider themselves to be Chinese rather than Taiwanese. Moreover, it turns out that attitudes towards unity and independence are, on average, the same among treated and control groups.

This paper is not only the first paper that uses quantitative methods to study Taiwan's authoritarian regime and its long-term influence but also is one of the first to provide potential historical reasons for the Taiwanese people's current attitude towards national identity and cross-strait reunification. We exploit within birth cohort estimation since the difference for people born in the same year is not large. A further advantage is that this paper employs many ways to address the possible threats to identification. For observable variables that may be correlated with both enrollment age and outcome variables, such as years of schooling, salary, and occupation, we find that these factors were not significantly related to the age of enrollment. For confounding factors, we run placebo tests, and find the effect is statistically insignificant.

Furthermore, our results offer new evidence to support a range of historical work arguing that education can play a crucial role in forming national identities (Weber, 1976; Clots-Figueras and Masella, 2013; Alesina and Reich, 2013). We believe that the attitude towards the national identity that individuals make regarding educational content deserve further study.

Bibliography

[1] Alesina, A., Fuchs-Schündeln. 2007. Good-Bye Lenin (or not?): The Effect of Communism on People's Preferences. American Economic Review, 97 (4), 1507-1528.

[2] Alesina, A and Reich B. 2013. Nation-building, Unpublished, Harvard University.

[3] Angrist, Joshua D., and Alan B. Krueger. 1991. Does Compulsory School Attendance Affect Schooling and Earnings? Quarterly Journal of Economics. 106 (4), 979-1014.

[4] Anderson, M. L. 2008. Multiple Inferences and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. Journal of the American Statistical Association, 103, 1481-1495.

[5] Apple, Michael W. 2004. Ideology and Curriculum. New York: Routledge.

[6] Apple, Michael W. 1982. Education and Power. Boston: Routledge.

[7] Bisin, Alberto and Thierry Verdier. 2001. The Economics of Cultural Transmission and the Dynamics of Preferences. Journal of Economic Theory, 97 (2), 298-319.

[8] Brim, O. G., Kagan, J. 1980. Constancy and Change in Human Development. Cambridge, MA: Harvard University Press.

[9] Brückner, M., Ciccone, A., Tesei, A. 2012. Oil Price Shocks, Income, and Democracy. Review of Economics and Statistics. 94 (2), 389-399.

[10] Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., Zhang, Y. J. 2017. Curriculum and Ideology, Journal of Political Economy. 125 (2), 338-392.

[11] Cantoni, D., and N., Yuchtman. 2013. The Political Economy of Educational Content and Development: Lessons from History. Journal of Development Economics, 104 (2), 233-244.

[12] Chen Yin-Hong. 2006. Jieyanhou guoxiao shehui jiaokeshuzhong guozu gainian zhi zhuanbian (The Change of Nation Identity Concept in Social Textbooks in Elementary Schools). Taiwan shifandaxue chuban.

[13] Clots-Figueras, Irma, and Paolo Masella. 2013. Education, Language and Identity. Economic Journal, 123 (570), 332-357.

[14] DellaVigna, S., Enikolopov, R., Mironova, V., Petrova, M., Zhuravskaya, E. 2011. Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia. Technical report, National Bureau of Economic Research.

[15] Dong Jin-Yu. 2014. High School Chinese Textbooks Serial. Roll House Press.

[16] Education in Taiwan. 1975. Taipei: Ministry of Education.

[17] Education in Taiwan. 1988. Taipei: Ministry of Education.

[18] Education in Taiwan. 1993. Taipei: Ministry of Education.

[19] Fuchs-Schündeln and Paolo Masella. 2016. Long-lasting Effects of Socialist Education. The Review of Economics and Statistics, 98 (3), 428-441.

[20] Giuliano, P., Spilimbergo, A. 2014. Growing Up in a Recession. The Review of Economic Studies, 81 (2), 787-817.

[21] Hu Yu-Reng. 2000. Guomin xiaoxue jiaokeshu zhi bentuhua yanjiu (An Analysis and Research on the Localization of Social Science Textbooks in Elementary Schools), Taiwan shifandaxue chuban.

[22] Lau Xin-Cheng. 2000. Taiwan zhengzhi minshenghua dui xiaoxue shehuikechen yingxiang zhi yanjiu (The Impact of Taiwanese Political Mogulization on Primary School Social Studies Curriculum). Taipei shifandaxue chuban.

[23] Lin Rui-Rong. 1994. Guomin xiaoxue zhengzhi huihua yanjiu (A Study of the Conciliation of Politics among Students in Primary Schools in Our Country). Taiwan shifandaxue chuban.

[24] Marshall, M. G. and Jaggers, K. 2005. Polity Iv Project, Political Regime Characteristics and Transitions, 18002003. Center for International Development & Conict Management.

[25] Ou Yung-Sheng. 1990. Guomin xiaoxue shehuikechen fenxi (An Analysis of Social Curriculum in National Primary School). Taiwan shifandaxue chuban.

[26] Routledge, Kegan Paul, Blum, W. 2003. Killing hope: US military and CIA interventions since The World War II. Zed Books.

[27] Rubinstein M. 1999. Taiwan: A New History, Boston: Routledge Press.

[28] Sacerdote, Bruce. 2001. Peer Effects with Random Assignment: Results for Dartmouth Roommates. Quarterly Journal of Economics, 116 (2), 681704.

[29] Shin-Yi Chou, Jin-Tan Liu, Michael Grossman and Ted Joyce. 2010. Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan. American Economic Journal: Applied Economics, 2 (1), 33-61.

[30] Spohr C. 2003. Formal Schooling and Workforce Participation in a Rapidly Developing Economy: Evidence from Compulsory Junior High School in Taiwan. Journal of Development Economics, 70 (2), 291-327.

[31] Su Ya-Chen. 2007. Ideological Representations of Taiwan's History: An Analysis of Elementary Social Studies Textbooks, 19781995. Curriculum Inquiry 37 (3), 205-237.

[32] Tsai Pei-E. 2007. Jieyanshiqi de guozu goujian (National Construction in the Textbook during the Martial Law). Taiwan Chengchidaxue chuban.

[33] Tsai, Wehn-Jyuan,Jin-Tan Liu, Shin-Yi Chou and Michael Grossman. 2011. Intergenerational Transmission of Human Capital: Results from a Natural Experiment in Taiwan. Working Paper.

[34] Tsai, Wehn-Jyuan. 2007. Does Educational Expansion Encourage Female Workforce Participation? A Study of the 1968 Reform in Taiwan. Ph.D. Dissertation. National Taiwan University.

[35] Tsai, Liu, Chou, Thornton. 2009. Does Educational Expansion Encourage Female Workforce Participation? A Study of the 1968 Reform in Taiwan. Economics of Education Review, 28 (6), 750-758.

[36] Voigtlnder, N., and H., Voth. 2012. Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany. Quarterly Journal of Economics, 127 (3), 13391392.

[37] Wen-lung Wang. 2007. An Imaginary Motherland in the Geographic Textbook of Taiwanese Middle School (1949-1999). National History Museum Academic Collection, 17, 201-251.

[38] Wang Hao-Bo. 1996. Guomin xiaoxue shehuikechenzhong zhi zhengzhixue tongzejieshi (An Analysis of the General Principles of Politics in the New Curricula in National Primary Schools. Study Information). Yanxi zixun, 13-3,51-54.

[39] Wang Hao-Bo. 2002. Guoxiao shehuike yu zhengzhishehuihua zhi yanjiu (1988-2000)(A Study on Social Science and Politics in Elementary Schools (1988-2000)). Jiaoyurenli yu zhuanyefazhan, 19-3,06-18.

[40] Weber, E, Peasants Into Frenchmen. 1976. The Modernization of Rural France, 18701914, Stanford, Calif.: Stanford University Press.

[41] Xie Qing-De. 1988. Ertong zhengzhi shehuihua zhi yanjiu- Youguan xiaoxue jiaokeshu de neirongfenxi (Political Socialization of a Primary School Textbook Content Analysis). Wenhuadaxue zhengzhi yanjiusuo chuban.

[42] Xue Yuan-Hua. 2000. Zhanhou Taiwan jiaoyu tizhi zhong guojia quanli wenti de lishi tantao (Historical Discussion on the Issue of State Power in Taiwan's Post-war Educational System). Proceedings of the Symposium on Social and Cultural Change in Taiwan, 393-407.

[43] Xu Ben. 2016. Tongzhi yu jiaoyu: cong gongmin dao guomin (Domination and Education: From Nationals to Citizens). Zhongyang bianyiju chubanshe.

Figure 2.1: Political Regimes Change in Taiwan (Polity IV)

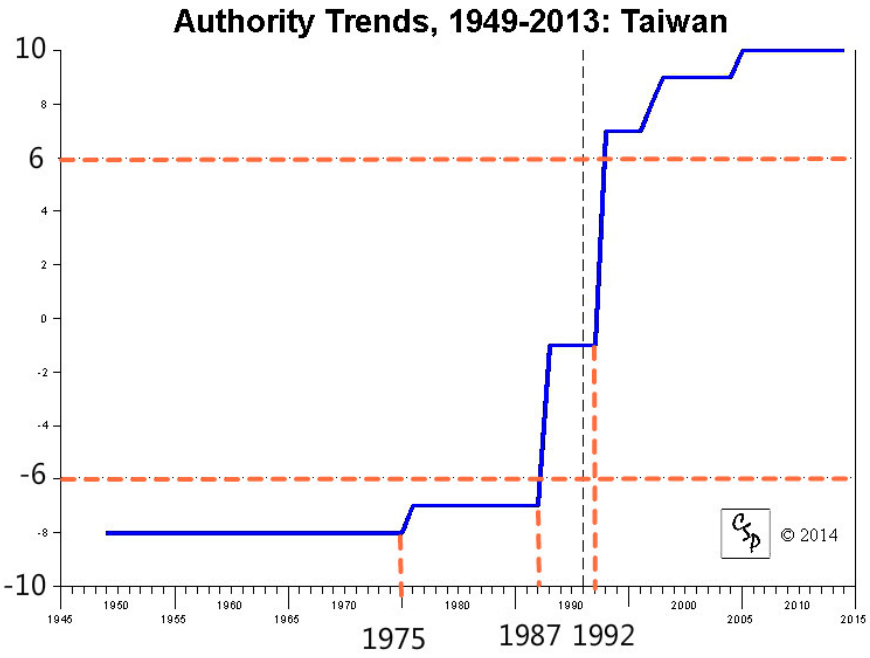

Figure 2.2: Teaching Hours Used for Political Indoctrination in the Curriculum

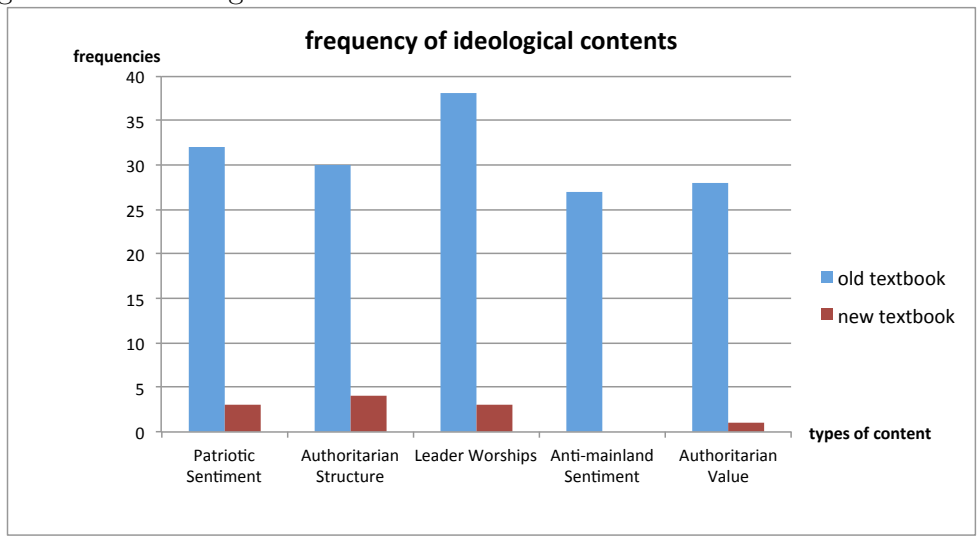

Table 2.1: Summary Statistics

	(1) All	(2) Control Sep.-Dec.	(3) Treated Jan.-Aug.	(4) Unconditional Diff.	(5) Unconditional p-value	(6) Conditional Diff.	(7) Conditional p-value
Panel A							
Age	32.13	32.33	32.01	-0.321	0.134	-0.168	0.369
	(5.411)	(5.519)	(5.347)				
Male	0.511	0.504	0.516	0.012	0.533	0.002	0.931
	(0.500)	(0.500)	(0.500)				
Taiwan-born father	0.905	0.912	0.901	-0.011	0.377	-0.015	0.204
	(0.293)	(0.284)	(0.298)				
Taiwan-born mother	0.958	0.953	0.961	0.008	0.275	0.007	0.372
	(0.200)	(0.212)	(0.192)				
Father edu. (years)	8.285	8.289	8.283	-0.006	0.972	-0.156	0.371
	(4.493)	(4.458)	(4.514)				
Mother edu. (years)	7.006	6.973	7.024	0.051	0.772	0.091	0.601
	(4.479)	(4.429)	(4.508)				
Religious	0.701	0.715	0.693	-0.022	0.213	-0.021	0.235
	(0.458)	(0.452)	(0.462)				
Urban	0.840	0.839	0.841	0.002	0.840	-0.002	0.888
	(0.366)	(0.368)	(0.365)				
Married	0.548	0.538	0.553	0.015	0.429	0.028	0.131
	(0.498)	(0.499)	(0.497)				
Treated	0.639	0	1	1	-	-	-
	(0.480)						
#obs.	2763	997	1766				
Panel B							
Pref. for democracy	0.580	0.615	0.560	-0.055	0.068	-0.053	0.084
	(0.494)	(0.487)	(0.497)				
Turn-out	0.779	0.807	0.762	-0.046	0.010	-0.042	0.023
	(0.415)	(0.395)	(0.426)				
Support for KMT	0.322	0.301	0.334	0.034	0.063	0.041	0.029
	(0.467)	(0.459)	(0.472)				
National identity (0-2)	1.561	1.601	1.537	-0.064	0.009	-0.055	0.027
	(0.560)	(0.543)	(0.569)				
Support unification (0-2)	0.916	0.895	0.929	0.034	0.218	0.036	0.196
	(0.577)	(0.599)	(0.564)				

[1] Columns 4 and 5 report raw (unconditional) differences in means across cohort, and the p-value for a t-test of differences in means. Columns 6 and 7 report differences conditional on birth cohort and region fixed effects.

[2] "Father edu." and "Mother edu." are measured by parental years of schooling. Other control variables are dummies.

[3] "Democracy"(#obs: 1646) appears only in wave 2000, 2004, 2005, 2010 and 2015; "Re-unification"(#obs: 2042) and "Taiwanese identity"(#obs: 1990) dummies show only in wave 2000, 2004, 2005, 2010, 2013, 2014, 2015; "Turnout"(#obs: 2701) and "Support KMT"(#obs: 2699) are dummies that exist in every wave.

Table 2.2: Baseline Results I

	Probit Model					
	Preference for democracy		Turn-out		Support KMT	
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-0.059*	-0.058*	-0.051***	-0.043**	0.032*	0.034*
	(0.032)	(0.030)	(0.017)	(0.018)	(0.018)	(0.019)
Age	-0.003	0.000	0.005	0.000	0.002	0.000
	(0.003)	(.)	(0.020)	(.)	(0.002)	(.)
Male	-0.006	-0.006	-0.031*	-0.026	-0.039**	-0.026
	(0.028)	(0.028)	(0.018)	(0.017)	(0.018)	(0.018)
Taiwan-born father	-0.059	-0.061	0.003	-0.004	-0.158***	-0.160***
	(0.059)	(0.059)	(0.034)	(0.033)	(0.034)	(0.033)
Taiwan-born mother	0.116	0.108	0.024	0.045	-0.065	-0.057
	(0.094)	(0.094)	(0.052)	(0.050)	(0.052)	(0.051)
Father edu.	0.000	0.001	0.008**	0.007*	0.005	0.004
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Mother edu.	-0.001	-0.001	-0.001	0.004	0.005	0.007*
	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)	(0.003)
Religious	0.017	0.015	0.047**	0.053***	0.033	0.026
	(0.032)	(0.032)	(0.020)	(0.019)	(0.020)	(0.020)
Married	-0.122***	-0.123***	-0.002	0.008	0.012	0.009
	(0.032)	(0.032)	(0.021)	(0.020)	(0.021)	(0.021)
Urban	0.041	0.041	0.023	0.017	0.043	0.032
	(0.045)	(0.046)	(0.029)	(0.028)	(0.030)	(0.030)
Birth cohort FE	Y	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y	Y
Wave FE	N	Y	N	Y	N	Y
N	1126	1126	2054	2054	2498	2498
Pseudo R2	0.1967	0.2049	0.1452	0.1804	0.1658	0.1735

Figures without brackets are marginal effects, namely dF/dx. Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.

Table 2.3: Baseline Results II

	Ordered Probit Model					
	National identity			Support for unification		
	Taiwanese	Both	Chinese	Unification	Status quo	Indep.
Treat	-0.038*	0.032*	0.006**	-0.013	-0.001	0.003
	(0.020)	(0.017)	(0.003)	(0.016)	(0.005)	(0.017)
Age	0.001	-0.001	-0.001	0.002	0.001	-0.002
	(0.003)	(0.003)	(0.001)	(0.002)	(0.001)	(0.003)
Male	-0.093***	0.080***	0.013***	0.041***	0.016***	-0.057***
	(0.019)	(0.016)	(0.030)	(0.012)	(0.005)	(0.016)
Taiwan-born father	0.144***	-0.124***	-0.020***	-0.084***	-0.032***	0.117***
	(0.033)	(0.029)	(0.005)	(0.022)	(0.010)	(0.031)
Taiwan-born mother	0.137***	-0.118***	-0.019**	-0.059*	-0.023*	0.081*
	(0.051)	(0.044)	(0.007)	(0.036)	(0.014)	(0.049)
Father edu.	-0.006*	0.005*	0.001	0.004**	0.002**	-0.006**
	(0.003)	(0.003)	(0.000)	(0.002)	(0.000)	(0.003)
Mother edu.	-0.003	0.002	0.000	-0.002	-0.001	0.003
	(0.004)	(0.003)	(0.000)	(0.002)	(0.001)	(0.003)
Religious	-0.037	0.032	0.005	-0.019	-0.007	0.026
	(0.023)	(0.020)	(0.003)	(0.015)	(0.006)	(0.021)
Married	-0.016	0.013	0.002	0.013	0.001	-0.002
	(0.023)	(0.021)	(0.003)	(0.013)	(0.001)	(0.018)
Urban	-0.087*	0.066*	0.010*	0.026	0.010	-0.036
	(0.039)	(0.034)	(0.006)	(0.023)	(0.009)	(0.031)
Birth cohort FE		Y			Y	
Region FE		Y			Y	
Controls & Wave FE		Y			Y	
N		1979			2009	
Pseudo R2		0.2021			0.1146	

Figures without brackets are marginal effects, namely dF/dx. Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.

Table 2.4: Results by Gender

Dep.Var.	Pref. for democracy		Turn-out		Support KMT	
	Male	Female	Male	Female	Male	Female
Treat	-0.071*	-0.038	-0.046*	-0.037	0.051*	0.002
	(0.042)	(0.043)	(0.027)	(0.026)	(0.028)	(0.028)
Birth cohort & region FE	Y	Y	Y	Y	Y	Y
Controls & wave FE	Y	Y	Y	Y	Y	Y
N	687	683	1346	1319	1342	1321
Pseudo R2	0.2774	0.3363	0.2567	0.2896	0.2398	0.2561
Dep.Var.	Taiwanese Identity		Support unification			
	Male	Female	Male	Female		
Treat	-0.076**	-0.016	0.065	-0.013		
	(0.027)	(0.027)	(0.043)	(0.040)		
Birth cohort & region FE	Y	Y	Y	Y		
Controls & Wave FE	Y	Y	Y	Y		
N	2237	836	1934	842		

Figures without brackets are marginal effects, namely dF/dx. Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.
The variable $TaiwaneseIdentity$, which is an ordinal variable, is given the value 2 when the respondent chooses Taiwanese, 1 if one declares he/she is both Taiwanese and Chinese, and 0 if one prefers Chinese. Here I report only the estimates of effects on one's Taiwanese identity for simplicity. The same for $Unification$

Table 2.5: Cut-offs Rather than 1987

Dep.Var.	Pref. for democracy			Turn-out		
	1989 as cut-off		1988 as cut-off	1989 as cut-off		1988 as cut-off
	1974-1980	1974-1982	1973-1981	1974-1980	1974-1982	1973-1981
Treat	-0.057*	-0.033	-0.054*	-0.044**	-0.043**	-0.044**
	(0.034)	(0.030)	(0.028)	(0.020)	(0.018)	(0.019)
Birth cohort & region FE	Y	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y	Y
N	1202	1430	1191	1927	2062	1928
Dep.Var.	Support KMT			Taiwanese Identity		
	1989 as cut-off		1988 as cut-off	1989 as cut-off		1988 as cut-off
	1974-1980	1974-1982	1973-1981	1974-1980	1974-1982	1973-1981
Treat	0.028	0.032*	0.030*	-0.029	-0.036*	-0.040**
	(0.020)	(0.019)	(0.018)	(0.021)	(0.019)	(0.019)
Birth cohort & region FE	Y	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y	Y
N	2361	2408	2375	1171	1279	1271

Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.
The variable $TaiwaneseIdentity$, which is an ordinal variable, is given the value 2 when the respondent chooses Taiwanese, 1 if one declares he/she is both Taiwanese and Chinese, and 0 if one prefers Chinese. Here I report only the estimates of effects on one's Taiwanese identity for simplicity.

Table 2.6: Heterogeneous Treatment Effects

	Turnout		Support KMT		Taiwanese Identity	
	Senior group	Junior group	Senior group	Junior group	Senior group	Junior group
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-0.050*	-0.004	0.051*	0.002	-0.070**	-0.012
	(0.028)	(0.006)	(0.028)	(0.006)	(0.026)	(0.015)
Birth cohort & region FE	Y	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y	Y
N	1200	983	1370	1125	1182	968

Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.
The variable *Taiwanese_identity*, which is an ordinal variable, is given the value 2 when the respondent chooses Taiwanese, 1 if one declares he/she is both Taiwanese and Chinese, and 0 if one prefers Chinese. Here I report only the estimates of effects on one's Taiwanese identity for simplicity.

Table 2.7: Confounding Factors: Education and Job Profiles

	Years of Schooling	Log Income	Job Type		
			Officials	Professionists	Workers
Treat	-0.140	0.002	0.003	0.011	-0.014
	(0.095)	(0.020)	(0.015)	(0.019)	(0.018)
Birth cohort FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
Controls & Wave FE	Y	Y	Y	Y	Y
N	2780	2702	927	625	1228

Job type is based on ISCO-08 occupations code, which has 9 major groups. Using these 9 binary dummies directly does not alter the conclusion.Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.

Table 2.8: Robustness Checks: Birth-cohort × Region FE

	Pref. democracy	Turn-out	Support KMT	Taiwanese Identity	Unification
Cutoff	-0.098**	-0.039*	0.069**	-0.014	0.029
	(0.040)	(0.023)	(0.036)	(0.040)	(0.028)
Controls & Wave FE	Y	Y	Y	Y	Y
Birth cohort & region FE	Y	Y	Y	Y	Y
Birth cohort × region FE	Y	Y	Y	Y	Y
N	1126	2054	2498	2045	1725

Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.

Table 2.9: Regressions for Placebo Cohort Groups

	Pref. for democracy 1960-1968	Turn-out 1960-1968	Support KMT 1960-1968
Treat	0.022	0.022	0.008
	(0.028)	(0.017)	(0.020)
Birth cohort & region FE	Y	Y	Y
Controls & Wave FE	Y	Y	Y
N	1382	2649	2471
	Taiwanese Identity 1960-1968	Unification 1960-1968	
Treat	0.028	0.004	
	(0.018)	(0.027)	
Birth cohort & region FE	Y	Y	
Controls & Wave FE	Y	Y	
N	737	562	

Two-way cluster-robust standard errors in parentheses. Significant at * 10%, ** 5%, *** 1%.
The variable $Taiwanese\_identity$, which is an ordinal variable, is given the value 2 when the respondent chooses Taiwanese, 1 if one declares he/she is both Taiwanese and Chinese, and 0 if one prefers Chinese. Here I report only the estimates of effects on one's Taiwanese identity for simplicity. The same idea for the variable $unification$, I only report the estimates of effects on one's unification attitude.

Chapter 3

Quantity-Quality Trade-off in Northeast China during the Qing Dynasty

3.1 Introduction

The possibility of a fertility inhibition model emerging in pre-modern China has generated widespread controversies among historical demographers. Malthus (1798) believed that the suppression of population growth in ancient China had been mainly caused by involuntary positive checks such as war, famine, and infectious diseases. While since the 1990s, younger-generation historians such as James Lee, William Lavely, and Cameron Campbell have reinterpreted the history of population growth since the Qing dynasty (1644-1911),[1] holding the view that the restrained births had occurred in China and internal fertility controls played a greater role in comparison to external inhibitors.[2] To tackle the long debate on the timing and origins behind

[1]These scholars have explored records of the Qing royal family, and household registers of villagers from Dao Yi, Liaoning.

[2]They suggested that Chinese families, influenced by an unintended consequence of children's competition for limited household resources, engaged in postnatal abortion and other contraceptive methods to control their fertility rates. However, Cao and Chen (2002; 2003) point out that the role of infanticide due to budget constraints cannot fully explain the dramatic decrease in male children during this period, since the boy preference in China made male infanticide minimal.

fertility controls in the male population in ancient China, this paper attempts to argue that the fertility restrictions happened and were based on rational parental decisions in accumulating human capitals. With high returns to education and an improved access to educational resources for ordinary people in the northeastern region, parents invested more in human capital by reducing the number of children they planned on having.

Our argument employs the quantity-quality trade-off theory, where parents limited their number of children to increase the quality of each of them, as portrayed by Becker (1960) and Becker and Lewis (1973). Being a key feature in the transition from Malthusian stagnation to modern economic growth as considered by long-term growth theories (Galor and Weil, 2000; Galor, 2011), the trade-off was accompanied by various underlying mechanisms, including the technological progress (Galor and Moav, 2002), income elasticity of preferences and demand for human capital (Fernihough, 2017). Although there is considerable debate on whether the quantity-quality relationship ever existed in industrialized countries in the midst of their demographic transition, the related literature on the Chinese social and economic history seems insufficient, mainly due to the lack of data on pre-modern China.

In the literature, using data from genealogies of individuals and households in the 13th-20th centuries Tong Cheng County of Anhui, Shiue (2017) is the first scholar to confirm that, since the 1600s, the fertility rates have shown a similar pattern to those of Western populations in the latter part of the industrial revolution, recording a substantial substitution of quantity by quality within the population. She finds that households with a smaller number of male children had a significantly higher chance of letting one of the children participate in the civil exam. However, Tong Cheng was a county with a well-developed economy, especially in business and trade, thus may not be the most representative one in Qing dynasty. This study focuses on the northeastern region of China. The fertility controls for educational purposes in this region, which was one of the most educationally underdeveloped areas (Liang, 2005), can imply that in other vast areas of China, where the Han-style schools were far more spread and popular, the trade-off may also exist and even emerge earlier.

In Northeast China, Han-style education began to develop in the late 17th century as a consequence of the inflow of exiled intellectuals and literary talents into the

northeast, and the adoption of the civil exam (*keju*). Then it rapidly spread and became popularized since the mid-18th century (i.e., around 1760-1770) when the number of Confucianism-oriented Academies (*shuyuan*) and private schools (*sishu*) dramatically increased and thus the children from a broader spectrum of households had a chance to access to education. Since then, an increase in the incentives to control births in response to education goals has occurred in the northeastern region among the households of all levels.

By investigating historical micro-data from the China Multi-Generational Panel Dataset, Liaoning (CMGPD-LN), we confirm that fertility restrictions due to human capital objectives were generated during the Qing era even in the underdeveloped northeast China. A negative relationship between fertility and education does not directly imply a rational choice of household; it may be unintended consequences due to limited household resources or behaviors taking the number of children as given. This concern is addressed by running regressions by household categories (elite vs. non-elite households[3]) and birth cohorts relating to different periods (pre- vs. post- mid-18th cohorts) in the process of a plebeianized education. The results of this paper support a historical link between human capital and reproductive rate for regular status households as a whole from the mid-18th to the end of the 19th century, corresponding to the period when a large number of *shuyuan* and *sishu* was established for people of common social classes in Liaoning.[4] For men from the non-elite households and born after the 1760s, one less brother leads to a 0.4% increase in the probability of obtaining an education; however, the negative relationship between reproduction of sons and human capital investment does not appear before, when the number of schools failed to cope with children from non-elite households. While among the elite households,the trade-off appears for both baseline (males born in the period of 1760-1880) and pre-baseline sample (men born before 1760 as pre-baseline cohorts), and the size of the sibship effect on education is always larger for those households and equals 3.2% in our baseline. The conclusion does not alter when the birth-order effect is cleared out with a two-stage estimation.

[3]Elite households refers to the households in which either the father or the grandfather had an administrative position or was educated, and the rest were non-elite households.

[4]This is not claiming that the resources of education were equally distributed across all social categories, and in fact the elite households still had a strong privilege in receiving an education.

In addition, this paper has a substantial advantage due to its dataset, enabling not only the counting of siblings but also the matching of twins. We explore the twins at last birth as the main instrument to control for unobservable determinants of fertility and parental choice. Other factors are considered in our robustness checks, such as household-level estimations, the inclusion of female siblings into the regressions, and the exclusion of infanticide, techniques of contraception, and natural disasters which may have influenced the sibship size unintentionally within a household.

Numerous studies of fertility-human capital trade-off in China are based on contemporary populations, mostly exploiting China's family planning policy and one-child policy as natural experiments. Wang, Zhao, and Zhao (2017) examines systematically the labor market outcomes of China's family planning policies, which limited the number of children. Liang and Gibson (2018) find an adverse effect of siblings on human capital accumulations and discover a stronger negative influence on girls than on boys, which reveals so-called "boy-preference" in Chinese society (Li, Yi, and Zhang, 2011; Zhang, 2017). Though this strand of literature could not quantitatively support for the transition occurred in pre-modern history, it enlightens our choice of empirical strategies.

This paper contributes to the literature by showing empirical evidence for the historical fertility transition from micro-level records collected in dynasty Qing and estimating unbiased results. Among the literature regarding the human capital-fertility relationships in history, aggregated data, which was firstly exploited, masks important heterogeneity determining their relationships (Guinnane, 2011). Recent and increasingly expanding literature employs more disaggregated information. Becker, Cinnirella, and Woessmann (2010, 2012), for instance, use county-level data to show that a quantity-quality tradeoff existed in 19th century Prussia. Scholars even explore individual data from genealogies and registers. Klemp and Weisdorf (2018) explore a genealogy of English individuals living between the 16th and the 19th centuries, and show that lower reproductive capacity exerted a positive influence on the educational achievements of offspring. The individual-level data we employ can also capture the essential individual and household characteristics in determining the human capital-fertility link amidst the fertility transition.

Another contribution is we generalize the origins of fertility transitions, by shift-

ing the focus from Europe to China. Generally, fertility decline has been associated with changes in child mortality, public schooling, child labor or other factors (Easterlin and Crimmins, 1985; Lee, 2003; Doepke, 2004). In contrast with Western developments during the period of interest, China experienced few socio-institutional changes involving fertility controls, perhaps even fewer in the northeastern region. Our findings emphasize that the characteristics to the Western experience were not a critical factors; however, the educational returns and costs were necessary for the quantity-quality trade-off to occur in a pre-industrial society. Finally, the current study not only challenges the traditional Malthusian model regarding ancient Chinese fertility and enriches the existing literature on the Chinese moral, cultural, institutional, and economic development, but it also deepens the understanding of the relationship between historical experience and economic growth.

The rest of the paper is organized as follows: Section 3.2 briefly introduces the historical background of northeastern education. Section 3.3 discusses the CMGPD-LN dataset and the variables. After that, we introduce the empirical methods used in the analysis in section 3.4. Section 3.5 presents the empirical results, as well as the robustness checks. Section 3.6 concludes the paper.

3.2 Historical Background of Northeastern Education during the Qing dynasty

The Jurchens were a federation of non-Chinese tribes living in the northeast of China (modern provinces of Liaoning, Jilin, and Heilongjiang). With a strong military organization of all tribes, they conquered northern China and founded the Jin dynasty (1115-1234). Their ethnic descendants, the Manchus established the Qing dynasty, later conquered the Ming dynasty and became the ruler of entire China between 1644 and 1911.

Before the Qing dynasty started to take over Inner China (i.e., China proper), Hong Taiji, a tribal leader of Manchus from 1626 to 1636, had realized the importance of education, had implemented regulation encouraging the elites' participation in the educational system and promoted Confucianism-oriented education. These efforts

fostered the development of northeastern education until 1644. However, due to the Qing dynasty's dominion over China proper which started in 1644, a substantial migration of educated people from the northeast to Inner China destroyed the base of education in this region (Gao, 1996; Liang, 2005).

3.2.1 The Emergence of Elite Education in the late 17th century

The Han-style education revived in northeast China after several decades due to two factors. The first factor was that the exiled intellectuals and literary talents lived and taught in this region. Though Han people's entry to the northeastern region was forbidden, the criminalized Han people were sent there by the Qing government for exile; many of them were political prisoners or intellectuals who had surrendered when the Qing army had attacked the Central Plains. These exiled people brought advanced knowledge of Inner China to the northeastern area and eventually fostered the development of this region as they were hired as teachers in local schools. Their roles in the spread of Confucianism, advanced culture, and Han Chinese language, made the natives regard them with the utmost respect: some highly-educated exiled people even had a high social status similar to that of military generals. As part of the development process of the northeastern education, the inflow of exiled people has founded the supply base of educational talents (Gao, 1996).

The second factor was the civil exam and high returns to education (Shi, 1994). To develop an effective form of governance and to consolidate the empire, the Qing dynasty adopted most of the institutions left from the Ming dynasty, including the civil exam. China's civil exam aimed to recruit talented citizens to govern the country and had three levels: the prefectural exam, the provincial exam, and the *jinshi* exam (Chen, Kung, and Ma, 2017). For students the first step was to pass the prefectural exam; the successful candidates could become *shengyuan* and were eligible to take the provincial exam. Only a few of them could pass this second exam and receive a *juren* title. At last, those with a *juren* qualification had a chance to take the final stage of the civil exam, which is also called the *jinshi* exam. Small parts of the *juren* qualification holders could achieve *jinshi* qualifications and were then guaranteed a

position in the upper levels of the government. These *jinshi* holders could become wealthy due to their high income and high social status. For the citizens with *shengyuan* or *juren* title, the return to education was not as high as for *jinshi* holders, but their income was much higher than the average income level.[5] Those who had received an education, but had not passed the *shengyuan* test, had a greater chance of finding a paid and non-physical job compared to illiterate citizens.[6] The roles of education were reinforced starting from the end of 17th century, when the Manchu people lost their privilege in the civil examination and governmental job applications (Shi, 1994; Zhang, 2004). Before 1687, Manchu people had a separated exam track and thus a relatively high chance to become state officials; therefore most of the Manchu officials were not required to have an equivalent educational level to that of their Han Chinese colleagues. However, in 1687 the Manchu people started being required to attend the same civil exam as the Han Chinese and to encounter direct competition from these highly-educated citizens; as a result of the policy, Manchu people had no more premiums during the application process for a governmental position. Due to this deregulation, Confucianism and Han Chinese-type education further expanded their influence among Manchu people as they directly determined the likelihood of becoming a state official (Ren, 2007; Xu, 2010).

Although the education began to develop and the returns to schooling remained high since the late 17th century, educational resources, such as schools, were so limited that they inevitably allowed only the entrance of the children from elite households, in which either father or grandfather was educated or had a governmental position. In the early period of the Qing Dynasty, the education in the Northeast was dominated by government-run schools, i.e., formal schools owned by the governments (*guanxue*), including state, prefecture or county-run schools (*fuxue*, *zhouxue* or *xianxue*), though the number of government-run schools was small with limited scale, and was mainly concentrated in Shengjing, Jilin, Heilongjiang and Dongsanmeng and aimed at preparing officials for the local government. Their admission

[5]Many of them could still work in educational institutions, such as teaching in public or private schools; meanwhile, they had the right to be exempt from taxation, and many of them received a subsidy from the state.

[6]Their level of knowledge and ability to use words skillfully made them qualified to work as a bookkeeper or a secretarial staff member.

procedure to their courses was strict and largely based on the administrative ranks (*pin*) of the candidates' father or grandfather; as a consequence, most of the admitted students were aristocratic bureaucrats (Zhao, 1928), only a quite small part of the students were plebeians with a qualification of *shengyuan*.

3.2.2 The Popularization of Northeastern Education since the mid-18th Century

Given that the supply of the government-run schools could not meet the educational demands of the ordinary citizens, the Qing government offered a considerable subsidy to Confucian Academies (*shuyuan*) as a way to encourage their operation and to satisfy the need for educational resources. However, the development of the Academies subsidized by the governments was still restricted, from the early Qing Dynasty to the eleventh year of Yongzheng era (1733). As documented by the "Chinese Ancient Academy System," - "Apart from several historically significant and influential ones, most *shuyuan* were shut down because the government could not provide enough funds for the functioning." In the early years of Qianlong era, with the support of the government, the *shuyuan* started to use the donations of urban industrial and commercial Confucian scholars to restore and allocate a large number of their services. In the middle of the Qianlong period (since the 1760s), the government also formulated stricter rules and regulations to ensure the rights of the *shuyuan* to their industries. Through these efforts, the *shuyuan* were able to develop and grow in both number and scales, and could be attributed to the limited admissions in the state or county-run schools (Xie, 2012): among the 18 *shuyuan* in Liaoning, 16 of them were established after the 1760s.

These formal schools were still insufficiently supplied, a large number of old-style private schools (*sishu*) thus emerged to provide training. Under the influence of the exiled intellectuals, the local scholars gradually grew up, and some people started the apprenticeship in private schools as a profession during the Qianlong era (Sun, 2010). As a result, also since the second half of the 18th century, these *sishu* could be found everywhere in Liaoning (Zhang, 2013). The objects of private education were quite extensive: children from the age of 5 or 6 up to the age of 25 or 26 years old,

regardless of their social background, had the opportunity to receive an education. They actively supplement the insufficiency of formal government-run schools and Confucian Academies, becoming the most important part of education in Northeast China, especially to children from non-elite origins.

To conclude, the exiled Han intellectuals, the establishment of an equitable civil exam fostered the birth of Han-style education in Northeast China, although the schools were only designed and accessible for aristocratic elites. Later the increased educational resources in this region, particularly the expansion of both private-fund *shuyuan* and *sishu* starting in the mid-18th century, more precisely 1760-1770, eventually generated incentives for the regular status citizens to invest in the human capital of their successors. We propose that high returns to education and the popularization of education, and thus lower costs of education,[7] were the keys generating the quantity-quality trade-off within the population based on the historical context examined.

3.3 Data and Variables

3.3.1 Dataset

In this study, the data used comes from the China Multi-Generational Panel Dataset-Liaoning (CMGPD-LN), a large-scale dataset which comprises triennial household register data covering the period of 1749-1909 and on individuals born between 1664 and 1909; it was set up by James Lee and Cameron Campbell, and its features have been discussed in Lee and Campbell (1997) and Lee and Wang (2009). The dataset contains 1.5 million records for approximately 260,000 people who lived in nearly 700 villages in the Liaoning province of northeast China. The spatial distribution of the villages involved is mapped in Figure 3.1. Compared with other datasets

[7]The negative relationship between accessibility of formal training and costs of education can be illustrated as follows: the tuition fees (*shuxiu*) in private schools (*sishu*) were decided by the supply and the demand sides, i.e., fewer *sishu*, higher fees; and if a plebeian one would like to go to public school, which enrolled students based on the classes of households, his father had to purchase a title/position to fit the requirements. Both two cases suggest that when the educational resources failed to meet the needs of non-elite households, the costs of education were high.

of China before the 20th century, CMGPD-LN provides the accurate sociological and demographical information required for this study, including important individual data on fertility and education, household factors such as basic information on family members, and village-level characteristics.

Arrangement of the Dataset: from Longitudinal to Cross-sectional data Given that the CMGPD-LN is a dataset collected specifically for Event History Analysis (EHA), data on the participating individuals was repeatedly recorded every three years. However, this paper only requires cross-sectional information on each individual (i.e., one record per individual), since we aim at disentangling the relationship among defining life decisions, such as human capital investments. To do so, we consider the birth year of a subject instead of their age, and select cross-sectional values at a certain point in the life of an individual as inputs; these are therefore regarded as fixed values in the analysis. For instance, variables such as the household size, the father's social status, occupation, education, retirement status at the time of his child's birth, the average grain prices and village population size when the child was of school-age (6-15 years old), along with some time-invariant factors such as individual educational experience, number of surviving brothers, and birth order, are all selected as main variables in the regression. By keeping and arranging all the needed information for one single record (i.e., observation), we make sure that each individual appears only once in our dataset.

Baseline Cohorts Figure 3.2 maps the pattern and timing of fertility controls in the northeastern region, where the values on the y-axis report the average size of surviving brothers for the birth cohorts 1680-1880.[8] It is observed from the figure that, although the reproductive rate had been in decline since the late 17th century, it was maintained at a low yet stable level, which signals an onset of birth controls which had taken place in these villages.[9] This implies is that the people in the northeastern

[8]Since the registers omit boys who died in their infancy or early childhood and most of the unmarried girls, by considering the production of surviving brothers, what we actually measure and study is reproductive rate rather than fertility rate.

[9]We noticed that the fertility controls in Manchuria (Figure 3.2) took place around 150 years before the transition occurred in Europe. We believe that all countries have different patterns of

region had by then accepted the idea of a quantity-quality trade-off: people were more willing to increase their investment in human capital by reducing their reproductive rate. This timing coincides with the historical statement concerning the growth of *shuyuan* and *sishu* in this region happened during 1760-1770. Hence we consider the males born in the period of 1760-1880 as a baseline to examine the linkage between fertility and education, and men born before 1760 as pre-baseline cohorts.[10]

Sample Limitations Since females in ancient China had neither political positions nor jobs and the data recording unmarried females is limited, we restrict the data to males who have survived childhood,[11] and whose family linkages could be identified so that their household conditions could be traced. The unique parental ID allows not only for the matching of parental education and social status, but also for the identification of twins. Individuals without a valid parental ID are excluded from the sample. By merging several files into one single dataset, household information (particularly household income from salaried positions, household size) and other data, such as average grain prices and the logged population during an individual's schooling time, are linked to personal data. Those whose household ID and village ID could not be traced were excluded from the study. Considering that there are possible mistakes in data collection, we also exclude some missing variables and observations with unreasonably extreme values.

These restrictions leave us with 19,490 records for 19,490 individuals who were born between 1760 and 1880 and came from approximately 699 different villages.

fertility change and do not necessarily follow the same pattern. For instance, although people deem that fertility transition in England happened in the nineteenth century, Cinnirella et al. (2017) show that the sampled families adjusted the timing of their births in accordance with the economic conditions as birth control in the three centuries that preceded historical fertility transition. During the Qing Dynasty, the popularity of imperial examinations and the accessibility to education led to changes in birth controls in some parts of ancient China since the mid-18th century. Besides, Shiue (2017) also confirms our point of view.

[10] Consider that individuals may go to school during their 5-15, we transfer this to a birth-year cutoff of around 1760. Using [1760-10, 1760+10] as cutoff between baseline and pre-baseline group does not alter the conclusion.

[11] We restrict our analytical sample to those who have survived into adulthood and therefore had exam opportunities to avoid detecting an artificial negative correlation between sibling number and individual education driven by higher early mortality risks of those having more siblings.

3.3.2 Variables

The subscript of the variable represents multiple variable dimensions: i, b, j, k stand for individual i of birth cohort b, who lived in household j of village k.

The dependent variable Edu_i is a dummy indicating if an individual i has ever passed one or more levels of the civil exam or was an official student (i.e., preparing for *shenyuan* exam), which we use to measure the individual educational achievement or his family's investment in human capital accumulation.[12] The main explanatory variables include Sib_i, the number of surviving brothers that each individual had, which is a suitable proxy of sibship size.[13] We generate this variable by sorting and counting total surviving brothers of every traceable individual.

Moreover, the siblings who share the same parental ID and born in the same month and year are considered as twins.[14] The twin variable is generated based on the all-male sample by excluding the females before the variable is defined, therefore includes only male twins.[15] We assign a dummy $Twin_i$ equals one to the individual belonging to the household with twins as the last birth.

Our set of covariates includes individual and cohort characteristics. $BirthOrder_i$ indicates individual's birth order, while $Ethnic_i$ suggests if the individual was an ethnic minority (ethnic Manchu, Mongol or Korean) or not; $BirthYear_b$ approximates the fixed effects of birth cohort b, which take into account the secular trends and patterns of education. Another birth cohort-specific variable is the average prices

[12]This information is obtained from the variables called *EXAMINATION* and *GUAN_XUE_SHENG* (official students) in the dataset. As the introduction and background sections suggest, there were several levels of education attained. However, the CMGPD-LN dataset does not contain the information about the levels of education, we thus can not use different levels of education instead of a 0/1 education dummy.

[13]Our paper does not count and include the female siblings in the baseline results due to the omission of girls in the dataset.

[14]Considering the recording errors of the variable named BAZI_MONTH (bazi birth month), we checked each record of the individual birth information, excluding the siblings born in the same month and year, but one (or both) of them has inconsistent BIRTHYEAR or BAZI_MONTH. We lost 70 pairs, but we thus make sure that all the remaining individuals in 545 twin-pairs have consistent recorded BIRTHYEAR and BAZI_MONTH.

[15]For the case that the male twin from a twin of opposite sexes left in the sample, we switch our focus to the overall sample, record down their IDs, and exclude these 283 observations from the male sample.

of main grains during birth-cohort b's school ages (between 6 and 15 years old in this paper), denoted $RicePrice_b$, and $SoyPrice_b$. These prices are measured by taels (liang) of silver per dan.

The parental and household-level covariates include: $F\_low\_status_j$ (if the father of household j belonged to a low status administrative population:[16] 1 if yes, 0 if no), $F\_position_j$ (if the father held salaried official positions or titles, such as soldiers, artisans and clerks: 1 if yes, 0 if no), $F\_education_j$ (if the father had received education: 1 if yes, 0 if no), F_ptitle_j (if the father held a purchased title: 1 if yes, 0 if no), F_retire_j (if the father was retired: 1 if yes, 0 if not), F_order_j (the father's birth order); as well as HH_salary_j (household income from salaried positions), HH_size_j (household size). We consider the values at each individual's birth for analysis.

A series of village-level control variables are considered, including $Region_k$ the birthplace dummies (North, Central, South-Central, South), $Logpop_{bk}$ (logged average population of village k at one's school ages).

3.3.3 Descriptive Statistics

In Table 3.1 the characteristics of the independent variables, as well as the individual-, household- and village-level control variables are summarized. The data shows that the average number of surviving brothers was 1.807. Around 2.5% individuals had a pair of twin brothers in the family. The average birth order in the sample was 2.084, and 4% of the included individuals were ethnic minorities. Very few fathers had a premium administrative status, with only 6.7% occupying an official position, and 1.4% having an academic degree. A summary of the other variables is also included in Table 3.1.

Moreover, columns (2) and (3) of Table 3.1 report summary statistics by educated status. Educated individuals had a smaller number of male siblings than not educated ones, and the difference in the number of male siblings of educated and non-educated was around 0.39, consistent with the theory of quality-quantity trade-off. These were also signs of heterogeneity of resource across cohorts: ethnic minorities had better

[16]This information is obtained from the variable in the codebook named *POPULATION_CATEGORY*, which identifies whether an individual belonged to a regular, specialized, or low status administrative population.

access to the education in the northeastern region; educated men tended to have more educated and high-status fathers as well as more wealthier households. All these factors are explicitly controlled for in the following regressions.

3.4 Empirical Methods

After deriving a simple theoretical framework in line with Galor and Moav (2002) which predicts the rational choice between human capital and fertility based on the relative returns and costs of education in Appendix C.1, we introduce the empirical methods to test whether our data supports the theoretical predictions.

3.4.1 Baseline Model

Given that the outcome is a binary variable, we apply the logit model and the linear probability model.[17] The outcome of the equation describing the human capital investment can be expressed as:

$$Edu_i = \beta_0 + \beta_1 Sib_i + \beta_2 \mathbf{X}_{ib} + \beta_3 \mathbf{W}_j + \beta_4 \mathbf{Z}_k + \epsilon_i \tag{1}$$

where Edu_i regards as the probability that individual i was educated and had formal degree; Sib_i is the number of surviving brothers that individual i had, which represented the reproductive rates; $\mathbf{X}_{ib}$ is a vector of covariates including individual features and cohort indicators; $\mathbf{W}_j$ contains parental specific factors and household characteristics of household j where i lived. $\mathbf{Z}_k$ further considers a series of village k effects; ϵ_i is an error term. Robust standard errors are clustered at the father level to account for the fact that siblings born to the same father may be correlated.[18]

Considering the coefficient β_1, as our Figure 3.3 below shows, larger sibship size is associated with a lower likelihood of being educated, which is in line with the

[17] Though the linear approximation may be inconsistent with probabilities close to zero or one; we still exploit linear regressions to ease the interpretation of the results. All results are robust regardless of the choice of specifications. The magnitude of coefficients estimated by logit is usually slightly larger than those by linear models.

[18] Using the standard errors based on clustering at other levels, such as mother ID, household ID does not alter the results.

findings of the literature.

3.4.2 Opportunity Costs of Education

Though we consider a comparison between baseline cohorts and pre-baseline cohorts inferred from the historical literature, we have neither a long enough across-dynastic cohorts series to show changes in the quantity-quality relationship in a long run nor precise information on yearly tuition fees and earnings, thus we cannot estimate how the quantity-quality relationship alters as in response to an unitary change in the educational returns and costs. Besides, some types of schools were government-or-society-subsidized, it thus seemed the improvement of quality would not crowd out the quantity. In light of these subsidies, would we expect to find a trade-off?

To solve these two problems, we take care of the opportunity costs (i.e. loss of labour income) associated with becoming a student. The real costs of human capital investments were not only the direct educational costs, such as tuition fees,[19] but also the indirect costs, such as opportunity costs, namely the earnings that were lost by choosing to receive an education instead of working during one's childhood. Pre-industrial northeast China was an agricultural society, the majority of households in this region made a living on farming,[20] therefore, the loss of labour income was mostly associated with the loss of agricultural outputs. The CMGPD-LN contains yearly information of main grain price in this region, which offers enough temporal variations and allows us to measure the opportunity costs of education by using the average grain prices during birth cohort b's school-age period (age 6-15 years) as a proxy and to examine how it changes quantity-quality relationship.

The profits gained from farming can reliably approximate the opportunity costs of education among farm households because the children (specifically for males) in agricultural households usually were required to be involved in farming activities during their entire childhood, while individuals pursuing education participated less in farming activities; some of them did not need to take such physical labor because

[19] In government-or-society-subsidized schools, these fees were almost equal to zero.
[20] In our sample, 83% of the households were pure farm households.

their parents hoped them to concentrate on their studies. We collect information on annual variations in prices of two main grains (rice and soy) and restrict the sample to men from farm households. The model with variations in the opportunity costs of education during one's childhood can be written as follows:

$$Edu_i = \alpha_0 + \alpha_1 Sib_i + \alpha_2 Price_b + \alpha_3 Price_b \times Sib_i + \alpha_4 \mathbf{X}_i + \alpha_5 \mathbf{W}_j + \alpha_6 \mathbf{Z}_k + e_i \quad (2)$$

Where α_2 captures an exogenous direct effect of grain prices on the probability of receiving education.[21] The interaction term captures the change in the child quantity-quality relationship when the opportunity costs vary. It means that, the marginal impact of the sibship size on the likelihood of receiving the education, is associated with grain prices. We expect when the opportunity costs were high (i.e., high grain prices), by contracting the sibship size, the human investment will increase less when compared to the baseline value; while as a consequence of a negative shock on the main grain prices (i.e., lower opportunity costs), the quantity-quality relationship should be statistically stronger among agricultural households.[22]

3.4.3 Netting out the Birth-order Effects

There are still two potential interpretations behind households' intended behaviors: first, in order to increase average education, one had to cut down the number of children that could be raised within certain budget constraints; second, higher educational attainment was typically associated with lower birth order, so constraining sibship sizes increased average education. Previous empirical studies ignore that sibship size and birth order are jointly determined, i.e., a household cannot change the family size, holding the average birth orders within the family constant. Particularly, if parents have a predisposition to favor some birth orders, the coefficient on sibship size in a regression of human capital investment on birth order and sibship

[21] The coefficient on grain price in the regression could be interpreted as the overall effect of price shock on education, including both substitution effect and income effect: the sign of overall effect indicates which effect dominates another.

[22] Our expectation is supported by the theoretical model in Appendix C.1

size provides a biased estimate of the quantity-quality trade-off (Bagger, Birchenall, and Mansour, 2013).

The two-steps empirical strategy from Bagger et al. (2013) resolves the issue. In the first step, controlling for family fixed effects (i.e., father fixed effects), birth-order effects are estimated. In the second step, the ultimate family size effect, is captured by "netting out" the estimated birth-order effect from the educational outcome $\hat{NE}_i = Edu_i - \sum_{q=1}^{Q} \hat{\gamma_{1q}} BirthOrder_{ijq}$.[23] Under this strategy, we get a clear impact of family size on average children's quality:

$$Edu_i = \gamma_0 + \sum_{q=1}^{Q} \gamma_{1q} BirthOrder_{ijq} + \gamma_2 \mathbf{X}_{ib} + \psi_j + u_i \tag{3.1}$$

$$\hat{NE}_i = \beta_0' + \beta_1' Sib_i + \beta_2' \mathbf{X}_{ib} + \beta_3' \mathbf{W}_j + \beta_4' \mathbf{Z}_k + v_i \tag{3.2}$$

Where u_j represents household fixed effects. Since the dependent variable in the second stage is generated by a regression, standard errors are corrected by bootstrapping across both stages. With the two-stage model, we are assessing the effect of sibship size on firstborns by looking at the outcomes of firstborns in households of different size, i.e., it is not feasible to look at the outcome of fifth-borns when the sibship size changes from three to four, as fifth born children are only observed in larger families.

3.4.4 Instrumental Variable

The coefficients for the sibship size as estimated by the OLS are likely to be biased even after controlling for a series of observable factors, as the fertility rates may be endogenous with respect to education. The likelihood of sending some children to school is also related to one's fertility choices. We thus employ an instrumental variable methodology and explore an exogenous fertility-determinant variable generated by the birth of twins. The use of twins as an instrument for estimating sibship size is pioneered by Rosenzweig and Wolpin (1980a), and has been applied by Bronars and Grogger (1994), and Angrist and Evans (1998), who use twin births

[23] The subscript q indicates the birth order q=1,..Q where Q is the maximum number of children.

as an instrumental variable for both male and female twins. Twin births have also been previously investigated in attempts to empirically support the Becker and Lewis (1973) quantity-quality model that predicts a negative impact of sibship size on a child's present-day outcomes (Rosenzweig and Wolpin, 1980b; Black and Salvanes, 2005; Caceres and Delpiano, 2006; Angrist, Lavy, and Schlosser, 2010).

Our IV strategy explicitly uses the occurrence of twins at last birth to instrument for sibship size. Using twins at last birth ensures that desired fertility is, on average, the same for parents with singletons and for parents with a twin birth. A key concern raised throughout the literature lies on the validity of the instrument. For our identifications to be valid, the instrument should satisfy the following three conditions: relevance, homogeneity, and the exclusion restriction assumption.

The first condition is satisfied, given that the main advantage of using twin births is that they provide an unexpected shock to fertility. Therefore, it is a particularly relevant instrument, as it increases the number of children by nearly one.

The occurrence of twins is exogenous to other personal or parental characteristics, the correlations among which could be rejected at a 10% significance level; the few existing exceptions, such as an household size, become important controls in the regression model. The IV approach requires the instrument to be random and conditional on observable covariates. This assumption can be violated if parents choose not to give birth to twins through medical treatments or abortions. These issues are unlikely to arise in our case: ancient China had no such technologies, therefore the sex of one's child, the propensity for twin births, and the birth order were all unpredictable. A possible concern to be addressed here is that of post-natal fertility controls. However, infanticide mostly affected girls and they have already been excluded from the sample.

In order for our instrument to be valid, it must also satisfy the exclusion restriction assumption related to the effect of having a twin brother on the probability of receiving an education. Using twin birth at last birth ensures that the family size changes without also changing the birth orders of the non-twin children.[24] Meanwhile, in the education-focused equation, we include variables that act as confounding factors and are affected by this fertility shock while having a direct effect on an individual's

[24]While twins had the same birth order in the registers

human capital formation, such as the household size.

3.5 Results

3.5.1 Sibship Size Effects

First, we examine the *sibship size* effect. We begin with the logit estimates as our benchmark results in Table 3.2, where the dependent variable is the probability of being educated Edu_i. Column (1) of Table 3.2 reports the results for the logit model including solely sibship size. The significant coefficient indicates that, without controlling for any heterogeneity, on average one less male sibling is correlated with a 1.2% increase in the likelihood of human capital formation. Though the effect seems small in absolute value, it roughly represents a large raise in human capital accumulations at the sample average. Conditional on individual characteristics and cohort indicators, we estimate the impact of brother size on an individual's education, as shown in column (2). Regression in column (3) further includes parental and household-level elements, which account for the potential endogeneity of fertility choice due to parental preferences and family background. In column (4), we estimate the full specification of the regression by adding the village population size, and the region fixed effects, which account for some unobservable factors such as regional development and other demographic elements. The results are very similar, showing that one less brother raises the probability of being educated by 0.86% for the sample as a whole, a 15% (0.86/5.74=0.15) increase when we evaluate at the mean of the outcome variable.

After reporting the estimate from a linear full-specification in column (1) of Table 3.3, whose absolute value is slightly smaller than the one obtained from the logit model, we divide individuals into several subsamples. The first comparison is between the elite and non-elite households in the baseline sample. In columns (2) and (3), the coefficients for the sibship size are significant in both regression. As expected, the size of the coefficient of interest is larger in column (3), and exhibits that one less male sibling for those individuals from elite households, who were able to acquire human capital easier, is associated with a greater chance of receiving an education

of 3.2%. In columns (4)-(7), we perform the same regression on pre-1760 cohorts as placebo.[25] From the historical documentation in section 3.2, during when the period they were raised, there were not sufficient Han Chinese-type *shuyuan* or *sishu* to meet their needs, resulting in a too high threshold to receive education. The fertility control for the human capital purpose was thus not popular among the pre-baseline sample as a whole. While when we restrict the sample to the children from elite households, the coefficient of *sibship size* becomes significant and equals -0.016. This result indicates that since education was profitable because of the adoption of civil exam, the households who had privilege in sending their children to *guanxue*, suppressed births to invest in education.

3.5.2 Opportunity Costs of Education

We then analyze the quantity-quality trade-off separately for those individuals from full-time farm households, where all the elder household members of the individual made their own clothing, tools, used farming to grow their own food and maintained a self-sufficient living, from those men in salaried household, where at least one elder household member earned salaries specified in archival documents for the positions or titles recorded in the register such as head of the household groups, artisan, clerks, and soldiers, etc. In column (1) of Table 3.4, the coefficient of interest for the farming households is significant but much closer to zero, relative to that of salaried household in column (4), whose probability of human capital investment raises by 2.7% with a unit decrease in sibling size, constituting evidence for a much more strengthened quantity-quality trade-off for the households with salaries. Meanwhile we expect the mechanism (i.e., both direct and indirect effect) of the average grain prices during one's six- to fifteen-year-old, defined as opportunity costs of going

[25] We choose individuals born between 1700 and 1760 as pre-baseline sample. The CMGPD-LN registers started from 1749, with limited coverage at the beginning. The calculated sibship size for those born in 1700-1760 may thus biased toward large and wealthy families with old fathers and adult brothers who managed to survive to 1749. To deal with this selection problem, we further restrict the pre-baseline sample to individuals born between 1730 and 1760, so that even the eldest cohorts just entered their adulthood in the first register, and thus just qualified to be compared with other males who had been alive until their adulthood in the regressions (see "Sample Limitations" of section 3.3.1)

to school, work only for farm households; and the coefficient of grain (rice) price, which suggests the total effect of our setting of educational costs, is only significant for farm households.

Columns (2) and (3) report that the opportunity costs played a role in determining the level of human capital accumulation and the quantity-quality correlation among households living on farming activities, no matter what kinds of grain considered. The coefficients of the interaction terms indicate the difference in education between one more and one less brother is shortened, capturing a negative effect of opportunity costs on the quality-quantity relationship explored. The results in these columns confirm the importance of opportunity costs on receiving an education as well as on the effect of reproductive rates for the households living on farming activities, and in general quality elasticity is substantially more amplified. In columns (5) and (6), we find that the direct and indirect effects of opportunity costs of education on the optimal educational decision are statistically zero for the households who live on the salary from positions, rather than agricultural activities. Individuals from this type of households were likely not required to work on land when they were at their school-age, thus the grain price is not their opportunity costs of education. Therefore, we do not observe any effect.

3.5.3 Two-step Estimation Results

Column (1) of Table 3.5 shows estimate with a linear specification of birth order, while column (2) allows for a more flexible specification by putting specific birth order indicators instead. The within-family estimates show that an individual with one additional unit of birth rank is 1.4% less likely to be educated, and this effect is highly significant. When the first-borns are used as the reference group, the second-born, third-born, fourth-born children, and the children with a higher birth rank than that are less and less likely to receive education. These estimates are negative and highly significant.

Columns (3)-(8) report the results in the second step of our two-step estimation. When we apply the linear regression of "netted education", which nets out the birth-order effect estimated in the first step, we find the coefficients of the *sibship size*

remain highly significant but with a smaller magnitude of 0.5%, as reported in column (3).

When we apply the same regression to the low-status or special duty population subpopulations (i.e., Booi Aha[26]), who according to historical documentation, did not have any chance or just a teeny tiny chance to take civil exams. As expected, the result reject a quantity-quality trade-off among these households, with an effect of *sibship size* close to zero as shown in column (4). Other population apart from this is considered as regular population, which is further divided into elite and non-elite households. Results from columns (5)-(7) are similar as our previous findings, while with smaller magnitudes.

3.5.4 Instrumented Results

As outlined above, the coefficients of sibship size estimated with the OLS are upwardly biased because of the endogeneity problem. Thus, we use an IV approach and explore twins at last birth as a plausible instrument. Table 3.6 presents the two-stage least squares (2SLS) estimate after implementing a linear specification in both stages. In columns (1) and (2), without controlling for any covariate, we simply instrument the sibship size with the incidence of twins at last birth, while in columns (3) and (4) we use the full specification of the model. Throughout all these specifications, the results of the first stage show a statistically significant and positive correlation between twins and sibship size. In particular, any incidence of twins at last birth increases the number of siblings by 70-92%.

According to the main results in second stage, the coefficients for instrumented *sibship size* are always negative and significant. We find that the OLS estimates are biased upwards, which fits our expectations. We can, therefore, draw strong conclusions for our study, as the calculated F-statistics of our instrument is greater than 10. Estimate of the full specification from the linear model in column (3) suggests that one less brother leads to a 2% increase in the likelihood of receiving formal schooling. In column (4), when the "netted education" is the outcome variable

[26]More introduction and regression results on the low-status population is addressed in Appendix C.2

instead of the education dummy, the size of effect on the human capital investment drops to 1.7%, but remain huge with respect to outcome mean and significant.

3.5.5 Robustness Checks

Inclusion of Female Siblings Although females were excluded on the ground that they were largely omitted from the household registers and typically not educated at that time, surviving girls would still carry some weight in the household budget.[27] We thus employ a sensitivity check where the number of female siblings in the same household are accounted for in addition to male siblings.

The analysis in Table 3.7 further includes the gender composition of the siblings in the covariates. We apply the previous exercises including two-step estimation (in column (2)), IV estimation (in column (3)), regressions by administrative classes (low-status population in column (4), comparison between elites and non-elites in columns (5) and (6)) and cohorts (baseline vs. pre-baseline cohorts in columns (1) and (7)). The results do not alter; and all estimates move upwards when the number of female siblings is taken into account.

Household-level estimation In this part, the education equation is estimated with household as the unit of analysis instead of the child, in order to check the robustness of the baseline findings by changing the estimation unit and sample. Focusing on the ratio of educated children as a function of total fertility in the household, it is not required to include variables at individual or cohort dimension, such as birth-order effects. We estimate the following specification:

$$RatioEducated_j = \zeta_0 + \zeta_1 Child_j + \zeta_2 \mathbf{W}_j + \zeta_3 \mathbf{Z}_k + \nu_j \tag{4}$$

Where the outcome is the ratio of educated children in the household j and the explanatory variable is $Child_j$, i.e., the number of children, of this household. The coefficient ζ_1 approximates the change in the ratio correlated with a unitary increase in the total number of children. Likewise, $\mathbf{W}_j$ includes household level characteristics

[27] For instance, a household could be huge if it consisted of two males and eight females; the study would make this like a relatively small family with a sibship size equal to two rather than ten.

such as the household income from salaried positions, household size, and parental characteristics such as the father's position, title, etc. And $\mathbf{Z}_k$ the village level characteristics, including logged population and regional dummies; ν_j is an error term. This model is estimated with both OLS and IV.

Results are shown in Table 3.8. OLS and IV estimates in columns (1)-(2) support that as a consequence of a unitary increase in the number of children, the ratio on average decreases about 1.5-2% for the full sample. Other results on different subsamples in columns (3)-(7) show a similar pattern to Tables 3.3 and 3.5 regardless of the approach explored, indicating the deliberate fertility control for human capital objectives was not homogenous to all households and start among the general population from the end of 18th century.

3.5.6 Alternative Explanations of Fertility Declines

In this part, we extensively rule out other possible influential factors, such as natural disasters, infanticide, and improved techniques of contraception, which may have influenced the sibship size unintentionally within a household.

Natural Disasters During Qing dynasty, natural disaster area concentrated on Hebei and Shandong (including flooding, windstorm, and drought). The northeastern region has rarely been a disaster-affected area due to its altitude and land fertility. According to the *Qing lichao shilu* (Record of the Qing Dynasty), a large-scale drought and hunger influenced many regions in Inner China in 1876, especially the North China Plain. Although we find no evidence that this drought also affected the northeastern region, we exclude the households with any child born after 1876 and find no substantial changes in our estimation after the exclusion.

Infanticide Lee and Campbell (1997) point out that in the northeastern region a substantial increase in food prices raised the potential threat of hunger, which in turn increased the likelihood to commit infanticide in low-income households. There are two reasons why infanticide does not affect our conclusion. First, the majority of our observations were male and thus unaffected by infanticide, given that this

practice had no impact on the male population in China (Cao and Chen, 2002). Second, the rate of infanticide could be overestimated because the northeastern farmers considered perinatal mortality to be a type of infanticide and thus misreported it.

Contraception Started during the Ming dynasty, the technology of contraception existed in Jiangsu and Zhejiang. However, this sort of technique was unreliable and unsuccessful in contraception. Cao and Chen (2003) highlight that the method of contraception plays no importance in fertility before the 20th century. As this technique never prevailed in the northeastern region during the period studied, the considerable reduction in Liaoning's male population could not be attributed to the contraception when facing limited household resources.

3.6 Conclusion

This paper provides empirical support for one of the central mechanisms of the unified growth theories which attempt to explain demographic transitions as well as economic growth over a long period of time (Galor, 2011; 2012). These theories typically are applied to ascribe the past fertility decline in Western Europe to the raising needs for education stimulated by the process of industrialization, and indicate that technological progress is the element which has encouraged both a demographic and a historical transition from Malthusian stagnation to modern growth.

In China, there was evidence of fertility control motivated by educational objectives starting as early as the 17th century in some highly-populated and developed counties (Shiue, 2017). This paper focuses on one other region in China, the northeastern region. Using quantitative and rigorously gathered evidence, we find evidence that the quantity-quality relationship held for a large part of China's population since the 18th century; as such, a smaller sibship size would lead to a higher likelihood of investing in human capital accumulations. This occurred along with increased returns on investments in education (i.e., more salaried positions) and greater accessibility to education (i.e., increased Han-style schools), providing incentives for people to reduce their family size and focus on human capital objectives.

We attempt to rule out other two interpretations of this negative linkage other

than rational behaviors in restricting fertility: unintended consequences due to limited household resources; and birth-order effects, that a constrained sibship size limited the individual with lower birth-order, therefore, increased the average education. The historical background as well as the heterogeneity across subpopulations strongly support the deliberate fertility restriction for incentives to receive education; while the birth-order effects are netted out by a two-step estimation. When IV procedures were employed and we instrument the sibship size by using twins at the last birth, the instrumented results also suggest that an additional sibling leads to a decrease in the probability of receiving education. All results remain robust when we change the estimation unit and sample.

To conclude, the trade-off between fertility and education was not necessarily the direct consequence of industrialization, which had not yet occurred in the region, but was caused by rational behaviors within households in response to regional development in education. Further studies on other economies within and outside of China are required to determine whether these findings can be generalized. These implications could help social scientists to better understand specific historical behaviors, but also lend support to growth patterns in ancient China.

Most literature on how quantity-quality relationship holds in third-world developing countries pays attention to the educational outcome of a birth control policy. While from our policy point of view, it is necessary to understand the extent to which a policy formulated to popularize education, such as the introduction of public education system (Li, Zhang, Zhu, 2008), the enrollment expansion of schools and universities (Duraisamy et al., 1999), etc., incentivizes the household's internal fertility controls for human capital investment and shapes a higher quality population for economic development.

Bibliography

[1] Acemoglu, Daron. 2009. Introduction to Modern Economic Growth. Princeton: Princeton University Press.

[2] Angrist, J. and W. Evans, 1998. Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. American Economic Review, 88 (33), 450-477.

[3] Angrist, J., V. Lavy, and A. Schlosser. 2010. Multiple Experiments for the Causal Link between the Quantity and Quality of Children. Journal of Labor Economics, 4 (28), 773-824.

[4] Bagger, J., J. A. Birchenall and H. Mansour. 2013. Education, Birth Order, and Family Size. NBER Working Papers 19111, National Bureau of Economic Research.

[5] Becker, Gary S. An Economic Analysis of Fertility. 1960. Demographic and Economic Change in Developed Countries. Princeton: Princeton University Press.

[6] Becker, Gary S. and H. Gregg Lewis. 1973. On the Interaction between the Quantity and Quality of Children. The Journal of Political Economy, 81 (2), 279-288.

[7] Becker, Sascha O., F. Cinnirella, and L. Woessmann. 2010. The Trade-Off between Fertility and Education: Evidence from Before the Demographic Transition. Journal of Economic Growth, 15 (3), 177-204.

[8] Becker, Sascha O., F. Cinnirella, and L. Woessmann. 2012. The Effect of Investment in Children's Education on Fertility in 1816 Prussia. Cliometrica, 6 (1), 29-44.

[9] Black, S. E., P. J. Devereux and K. G. Salvanes. 2005. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. The Quarterly Journal of Economics, 120 (2), 669-700.

[10] Bronars, Stephen G. and Jeff Grogger. 1994. The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment. American Economic Review, 84 (5), 1141-1156.

[11] Cáceres-Delpiano, J. 2006. The Impacts of Family Size on Investment in Child Quality. Journal of Human Resources, 41 (4), 738-754.

[12] Cao, Shuji and Chen Yixin. 2002. Malthusian Theory and the Chinese Population since the Qing Dynasty: A Review of American Scholars' Recent Studies (in Chinese), Historical Studies, 1, 41-54.

[13] Cao, Shuji and Chen Yixin. 2003. Respects the Reality of China's Population History (in Chinese), Academics, 3, 116-132.

[14] Chen, Ting and Kung, James Kai-Sing and Ma, Chicheng. 2017. Long Live Keju! The Persistent Effects of China's Imperial Examination System. Working paper.

[15] Cinnirella, F., M. Klemp, and J. Weisdorf. 2017. Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England. Demography, 54 (2), 413-436.

[16] Doepke, Matthias. 2004. Accounting for Fertility Decline during the Transition to Growth. Journal of Economic Growth, 9 (3), 347-383.

[17] Duraisamy, P.James, Estelle Lane, Julia Jee-Peng Tan. 1998. Is There a Quantity-Quality Tradeoff as Enrollments Increase? Evidence from Tamil Nadu, India. Policy, Research working paper; No. WPS 1768. Washington, DC: World Bank.

[18] Easterlin, Richard A. and E.M. Crimmins. 1985. The Fertility Revolution: A Supply-Demand Analysis. University of Chicago Press.

[19] Elvin, Mark. 1973. The Patter of the Chinese Past: A Social and Economic Interpretation. Stanford University Press.

[20] Fernihough, A. 2017. Human Capital and the Quantity-Quality Trade-Off during the Demographic Transition. Journal of Economic Growth, 22 (1), 35-65.

[21] Fiore, S., Bratti, M. and M. Mendola. 2015. Family Size, Siblings Rivalry and Migration: Evidence from Mexico. Working Paper.

[22] Franck, Raphael and Oded Galor. 2015. Industrialization and the Fertility Decline. Working Paper.

[23] Galor, Oded. 2011. Unified Growth Theory. Princeton: Princeton University Press.

[24] Galor, Oded and Omer Moav. 2002. Natural Selection and the Origin of Economic Growth. The Quarterly Journal of Economics, 117 (4), 1133-1191.

[25] Galor, Oded and D. N. Weil. 2000. Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond. The American Economic Review, 90 (4), 806-828.

[26] Gao, Rong. 1996. The Basic Characteristics of Education Development in Northeast China in the Early Qing Dynasty (in Chinese). Journal of Social Sciences of Jilin University, 6, 69-71.

[27] Greif, Avner and Guido Tabellini. 2010. Cultural and Institutional Bifurcation: China and Europe Compared. American Economic Review, 100 (2), 135-140.

[28] Greif, Avner and Guido Tabellini. 2012. The Clan and the City: Sustaining Cooperation in China and Europe. SSRN 2101460.

[29] Guinnane, Timonthy. 2011. The Historical Fertility Transition: A Guide for Economists. Journal of Economic Literature, 49 (3), 589-614.

[30] Klemp, Marc and J. Weisdorf. 2018. Fecundity, Fertility and the Formation of Human Capital. Economic Journal, forthcoming.

[31] Lee, James and Cameron Campbell. 1997. Fate and Fortune in Rural China: Social Organization and Population Behavior in Liaoning 1774-1873. New York: Cambridge University Press.

[32] Lee, James, Campbell Campbell, Chen S. 2010. China Multigenerational Panel Dataset, Liaoning (CMGPD-LN) 1749-1909 User Guide. Inter-university Consortium for Political and Social Research; Ann Arbor.

[33] Lee, James and Feng Wang. 2009. One Quarter of Humanity, Malthusian Mythology and Chinese Realities, 1700-2000. Cambridge: Harvard University Press.

[34] Lee, Ronald. 2003. The Demographic Transition: Three Centuries of Fundamental Change. Journal of Economic Perspectives, 17 (4), 167-190.

[35] Liang, Yuduo. 2005. The Characteristics and Historical Role of Northeastern Education before the Opium War in the Qing Dynasty (in Chinese). Blackwater Culture Proceedings, 3, 290-294.

[36] Li, Hongbin, Junjian Yi, and Junsen Zhang. 2011. Estimating the Effect of the One-Child Policy on the Sex Ratio Imbalance in China: Identification based on the Difference-in-Differences. Demography, 48 (4), 1535-1557.

[37] Li, Hongbin, Junsen Zhang, and Yi Zhu. 2008. The Quantity-Quality Trade-Off of Children in a Developing Country: Identification Using Chinese Twins. Demography, 45 (1), 223-243.

[38] Liang, Yun, and John Gibson. 2018. Do Siblings Take Your Food Away? Using China's One-Child Policy to Test for Child Quantity-Quality Trade-Offs. China Economic Review, 48, 14-26.

[39] Liu, Tsuijung. 1992. Chinese Genealogies as a Source for the Study of Historical Demography. Studies and Essays in Commemoration of the Golden Jubilee of Academia Sinica. Taipei: Academia Sinica.

[40] Ren, Yinyan. 2007. Research on the Imperial Examination System of Kangxi Dynasty. Northwest Normal University.

[41] Ren, Yuxue. 2003. The Analysis of Shengjing House Office Account (1644 - 1796) (in Chinese). The Qing History Research, 2, 16-26.

[42] Rosenzweig, M. R. and K. I. Wolpin. 1980a. Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models. Journal of Political Economy, 88 (2), 328-348.

[43] Rosenzweig, M. R. and K. I. Wolpin. 1980b. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. Econometrica, 48 (1), 227-240.

[44] Shi, Zhongwen. 1994. The Whole History of China-The History of Chinese Qing Dynasty Education (in Chinese). People's Publishing House.

[45] Shiue, Carol H. 2017. Human Capital and Fertility in Chinese Clans before Modern Growth. Journal of Economic Growth, 22 (4), 351-396.

[46] Tabellini, Guido. 2010. Culture and Institutions: Economic Development in the Regions of Europe. Journal of the European Economic Association, 8 (4), 677-716.

[47] Wang, Fei, Liqiu Zhao, and Zhong Zhao. 2017. China's Family Planning Policies and Their Labor Market Consequences. Journal of Population Economics, 30 (1), 31-68.

[48] Xu, Yang. 2010. Eight Banners Imperial Examinations: 1644-1795–Comparing and Studying with Jin and Yuan Dynasty, Northeast Normal University.

[49] Zhang, Jie. 2004. The Multiple Influences of the Imperial Examination System in the Qing Dynasty on the Development of Manchu Culture (in Chinese). Study and Exploration, 4, 129-133.

[50] Zhang, Junsen. 2017. The Evolution of China's One-Child Policy and Its Effects on Family Outcomes. Journal of Economic Perspectives, 31 (1), 141-160.

[51] Zhao, Dexiang. 1995. The History of Liaoning Education in the Early Qing Dynasty (in Chinese). Journal of Liaoning Normal University, 5, 80-84.

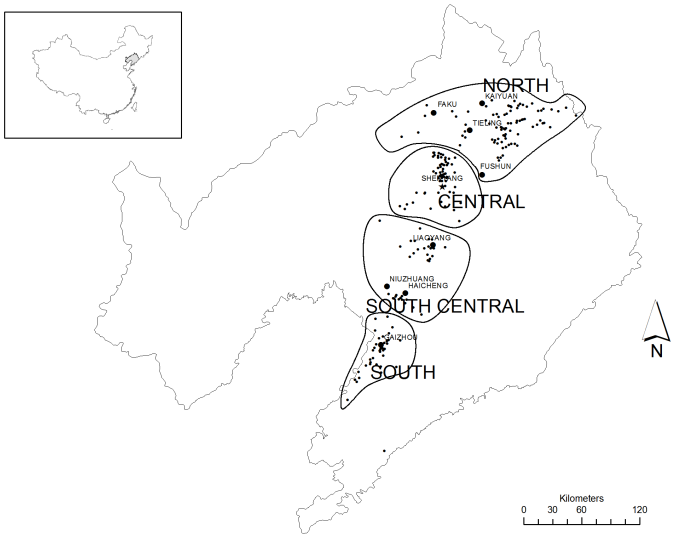

Figure 3.1: Located CMGPD-LN villages. Source: Lee et al. (2010)

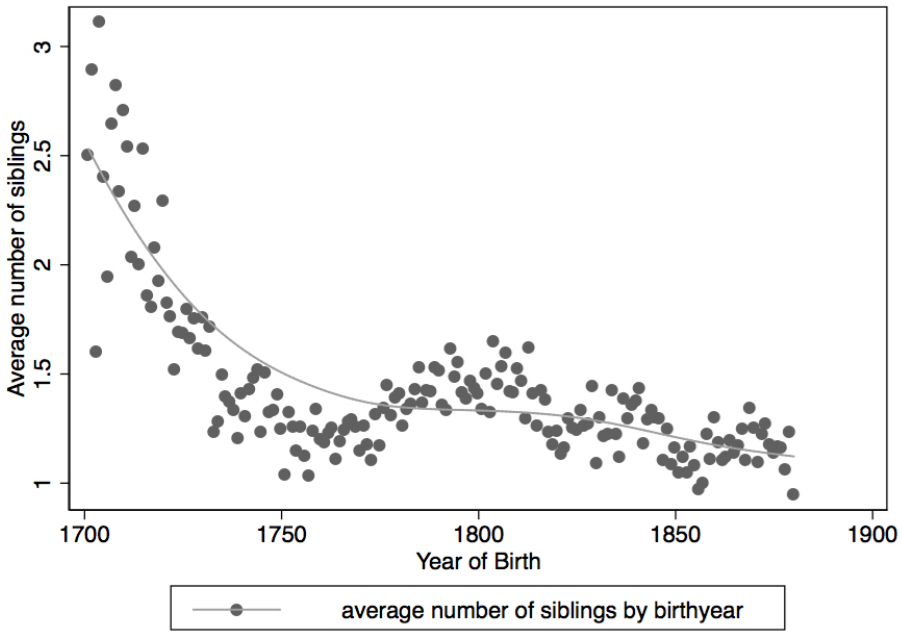

Figure 3.2: Fertility Transitions

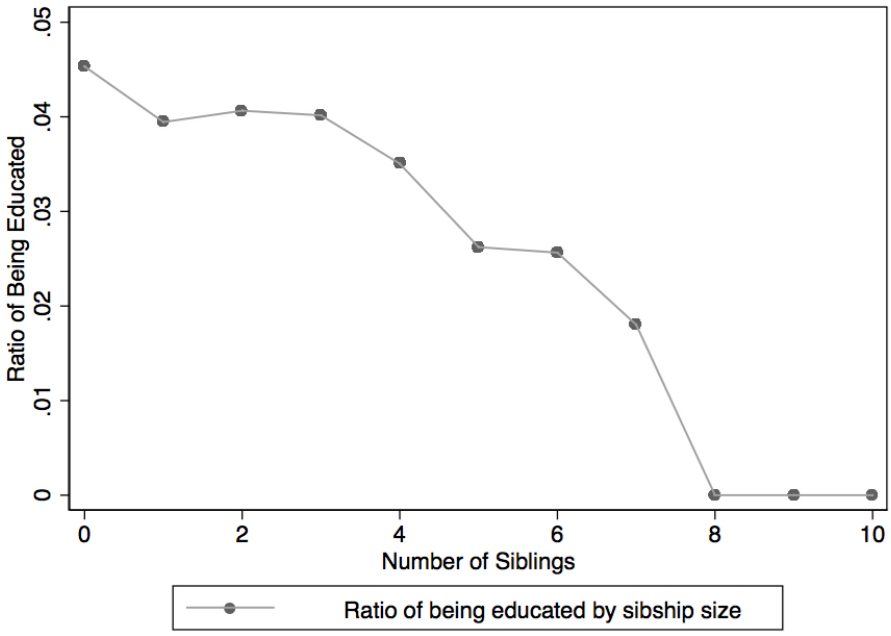

Figure 3.3: Fertility-Education Trade-off

	All	Non-educated	Educated	Difference
Explanatory & Instrumental Variables				
Sibship size	1.807	1.830	1.439	0.390***
	(1.437)	(1.448)	(1.182)	
Twins in family	0.0247	0.0255	0.0116	0.0139***
	(0.155)	(0.158)	(0.107)	
Individual level				
Birth order	2.084	2.084	2.084	-0.000
	(1.373)	(1.378)	(1.409)	
Ethnic minority	0.0419	0.0396	0.0804	-0.0408***
	(0.200)	(0.195)	(0.272)	
Cohort level				
Birth year	1833.6	1834.4	1820.4	14.02***
	(36.67)	(36.76)	(32.51)	
Rice price	1.831	1.834	1.765	0.0708***
	(0.417)	(0.411)	(0.507)	
Soy price	0.606	0.608	0.587	0.0205***
	(0.131)	(0.130)	(0.154)	
Household level				
Father's low status	0.0692	0.0719	0.0241	0.0478***
	(0.254)	(0.258)	(0.153)	
Father's position	0.0669	0.0499	0.345	-0.295***
	(0.250)	(0.218)	(0.475)	
Father's education	0.0140	0.0099	0.0813	-0.0709***
	(0.118)	(0.099)	(0.273)	
Father's purchased title	0.0074	0.0049	0.0482	-0.0433***
	(0.0856)	(0.0698)	(0.214)	
Father is retired	0.0078	0.0079	0.0054	0.0025
	(0.0877)	(0.0885)	(0.0730)	
Father's birth order	2.131	2.121	2.303	-0.182***
	(1.542)	(1.526)	(1.777)	
Household salary	6.338	4.627	34.40	-29.78***
	(26.04)	(22.70)	(50.15)	
Household size	20.48	20.66	17.51	3.143***
	(14.39)	(14.50)	(12.13)	
Village level				
Logpop	5.837	5.820	6.121	-0.301***
	(1.305)	(1.301)	(1.346)	
#obs.	19,490	18,370	1,120	

[1] Column 4 report raw differences in means across cohort.
[2] The average grain prices have a different number of observations: 18,918.

Table 3.1: Summary Statistics

	Education (1=Yes; 0=No)			
	(1)	(2)	(3)	(4)
Sibship size	-0.220***	-0.247***	-0.201***	-0.202***
	(0.028)	(0.032)	(0.033)	(0.033)
Birth order		0.066**	-0.109***	-0.123***
		(0.028)	(0.031)	(0.031)
Ethnic		0.568***	0.084	0.082
		(0.126)	(0.149)	(0.145)
Father's low status			-1.025***	-0.923***
			(0.225)	(0.220)
Father's position			1.766***	1.575***
			(0.195)	(0.162)
Father's education			0.972***	1.111***
			(0.295)	(0.302)
Father's purchased title			1.237***	1.645***
			(0.372)	(0.381)
Father is retired			-0.778	-0.729
			(0.548)	(0.528)
Father's birth order			0.011	-0.021
			(0.027)	(0.027)
Household salary			0.019***	0.014***
			(0.005)	(0.004)
Household size			-0.039***	-0.034***
			(0.006)	(0.006)
Village population (logged)				0.096***
				(0.031)
Marginal Effects				
Sibship size	-0.0119***	-0.0128***	-0.0088***	-0.0086***
	(0.0015)	(0.0016)	(0.0014)	(0.0014)
Birth cohort FE	No	Yes	Yes	Yes
Region FE	No	No	No	Yes
N	19490	19490	19490	19490

[1] Coefficients are estimated from the logit model.
[2] Robust standard errors are reported between parentheses and are clustered at the father level. Significant at * 10%, ** 5%, *** 1%.

Table 3.2: Sibship Size Effects

	Education (1=Yes; 0=No)						
	Baseline sample			Pre-baseline sample			
	All sample	Elite	Non-elite	1700-1760	1730-1760	Elite	Non-elite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sibship size	-0.0076***	-0.0319***	-0.0042***	0.0021	0.0039	-0.0165*	0.0038
	(0.0013)	(0.0061)	(0.0011)	(0.0026)	(0.0031)	(0.0098)	(0.0025)
Birth order	-0.0059***	-0.0140***	-0.0041***	-0.0048*	-0.0064*	-0.0291**	-0.0009
	(0.0014)	(0.0045)	(0.0013)	(0.0028)	(0.0033)	(0.0117)	(0.0027)
Ethnic	0.0128	0.0088	0.0132	0.0040	0.0038	-0.0081	0.0042
	(0.0102)	(0.0314)	(0.0093)	(0.0088)	(0.0095)	(0.0501)	(0.0079)
Father's low status	-0.0278***	-0.0965**	-0.0212***	0.0192	0.0200	0.0000	0.0212
	(0.0046)	(0.0397)	(0.0045)	(0.0198)	(0.0199)	(.)	(0.0199)
Father's position	0.1697***	0.1336***	0.0000	0.0246	0.0307	0.1154***	0.0000
	(0.0170)	(0.0184)	(.)	(0.0196)	(0.0206)	(0.0438)	(.)
Father's education	0.1121***	0.1072***	0.0000	0.2294	0.2410	0.3558**	0.0000
	(0.0355)	(0.0362)	(.)	(0.1603)	(0.1771)	(0.1778)	(.)
Father's purchased title	0.1398***	0.1465***	0.0000	0.3320**	0.3280**	0.2612	0.0000
	(0.0512)	(0.0514)	(.)	(0.1466)	(0.1564)	(0.1596)	(.)
Father is retired	-0.0333*	-0.1718	-0.0118	0.0043	-0.0014	-0.0857	0.0133*
	(0.0187)	(0.1060)	(0.0150)	(0.0090)	(0.0109)	(0.0529)	(0.0080)
Father's birth order	-0.0016	-0.0020	-0.0025**	-0.0001	0.0001	-0.0180	0.0011
	(0.0013)	(0.0042)	(0.0011)	(0.0021)	(0.0023)	(0.0152)	(0.0019)
Household salary	0.0014***	0.0015***	0.0013***	0.0030***	0.0029***	0.0033***	0.0030***
	(0.0003)	(0.0004)	(0.0005)	(0.0003)	(0.0003)	(0.0005)	(0.0004)
Household size	-0.0010***	-0.0044***	-0.0003***	0.0001	0.0002	-0.0011	0.0002
	(0.0001)	(0.0006)	(0.0001)	(0.0002)	(0.0002)	(0.0011)	(0.0001)
Village population (logged)	0.0080***	0.0244***	0.0041***	-0.0013	-0.0017	0.0141	-0.0021*
	(0.0014)	(0.0072)	(0.0013)	(0.0013)	(0.0015)	(0.0088)	(0.0012)
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	19490	2438	17052	6348	4982	478	5870
Mean DV (%)	5.75	21.29	3.52	3.5	3.75	17.36	2.37

[1] The baseline sample means males born in the period of 1760-1880 and pre-baseline sample means men born between 1700 and 1760.
[2] Coefficients are estimated from the linear probability model.
[3] Robust standard errors are reported between parentheses and are clustered at the father level. Significant at * 10%, ** 5%, *** 1%.

Table 3.3: Sibship Size Effects by Subsamples

	Education (1=Yes; 0=No)					
	Farm households			Salaried households		
	(1)	(2)	(3)	(4)	(5)	(6)
Siblings	-0.0043***	-0.0170***	-0.0167***	-0.0275***	-0.0119	-0.0164
	(0.0009)	(0.0045)	(0.0047)	(0.0060)	(0.0204)	(0.0239)
Rice price	-0.0214***	-0.0339***		0.0122	0.0282	
	(0.0040)	(0.0065)		(0.0181)	(0.0289)	
Siblings × Rice price		0.0070***			-0.0086	
		(0.0023)			(0.0110)	
Soy price			-0.1094***			0.0579
			(0.0208)			(0.0955)
Siblings × Soy price			0.0205***			-0.0184
			(0.0073)			(0.0381)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
N	16231	16231	16231	2687	2687	2687

[1] Coefficients are estimated from the linear probability model. The marginal effect of siblings on education is a function of grain prices; this function we assume to be linear.

[2] Robust standard errors are reported between parentheses and are clustered at the father level. Significant at * 10%, ** 5%, *** 1%.

[3] Controls include birth order, ethnicity, father's low status, position, education, purchased title, retirement status and birth order, household salary, household size, and village population. These regressions do not contain cohort indicators; they absorb entirely the effect of opportunity costs which vary only across cohort.

Table 3.4: Opportunity Cost of Education: Grain Prices and Farming Households

	First step		Second step				
	Education (1=Yes; 0=No)		Netted Education ($Edu_i - \sum_{q=1}^{Q} \hat{\gamma_{1q}} BirthOrder_{ijq}$)				
			All	Low status	Regular status		
					All regular	Elite	Non-elite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sibship size			-0.0049***	0.0014	-0.0054***	-0.0330***	-0.0012*
			(0.0011)	(0.0027)	(0.0011)	(0.0054)	(0.0007)
Birth order	-0.0137***						
	(0.0019)						
Second-born		-0.0228***					
		(0.0037)					
Third-born		-0.0323***					
		(0.0054)					
Fourth-born		-0.0396***					
		(0.0085)					
≥Fifth-born		-0.0598***					
		(0.0102)					
Father fixed effects	Yes	Yes	-	-	-	-	-
Controls	-	-	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	-	-	Yes	Yes	Yes	Yes	Yes
Region FE	-	-	Yes	Yes	Yes	Yes	Yes
N	19490	19490	19490	1348	18142	2438	15761

[1] Coefficients are estimated from the linear probability model.

[2] Robust standard errors are reported between parentheses and are clustered at the father level. Significant at * 10%, ** 5%, *** 1%.

[3] Covariates include ethnic minority, father's low status, position, education, purchased title, retirement, age at child's birth, household salary, household size, and logged village population.

Table 3.5: Birth-order Effects

	(1)	(2)	(3)	(4)
First Stage	Without covariates		With all covariates	
Twin births	0.922***		0.698***	
	(0.124)		(0.093)	
F statistics	55.55		18.69	
Second Stage	Education (1=Yes; 0=No)	Netted educ	Education (1=Yes; 0=No)	Netted educ
Siblings	-0.0338***	-0.0228**	-0.0221**	-0.0174*
	(0.0103)	(0.0106)	(0.0103)	(0.0104)
Birth order			0.0010	-
			(0.0051)	
Ethnic			0.0135	0.0118
			(0.0103)	(0.0095)
Father's low status			-0.0288***	-0.0282***
			(0.0046)	(0.0045)
Father's position			0.1687***	0.1665***
			(0.0172)	(0.0172)
Father's education			0.1081***	0.1082***
			(0.0357)	(0.0358)
Father's purchased title			0.1433***	0.1426***
			(0.0516)	(0.0515)
Father is retired			-0.0300	-0.0295
			(0.0189)	(0.0188)
Father's birth order			-0.0024	-0.0023
			(0.0015)	(0.0015)
Household salary			0.0014***	0.0015***
			(0.0003)	(0.0003)
Household size			-0.0009***	-0.0008***
			(0.0002)	(0.0002)
Village population (logged)			0.0083***	0.0082***
			(0.0014)	(0.0014)
N	19490	19490	19490	19490

[1] Coefficients in first stage are estimated by OLS; coefficients in second stage are estimated from the linear probability model.

[2] Robust standard errors are reported between parentheses and are clustered at the father level. Significant at * 10%, ** 5%, *** 1%.

Table 3.6: Instrumented Results

	All sample			Low status	Regular status		Pre-baseline sample	
	OLS	Two-step est.	IV		Elite	Non-elite	All sample	Elite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Sibship size	-0.0054***	-0.0031***	-0.0197*	-0.0004	-0.0203***	-0.0030***	0.0028	-0.0151*
(*female siblings included*)	(0.0013)	(0.0011)	(0.0102)	(0.0035)	(0.0055)	(0.0011)	(0.0026)	(0.0085)
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	19490	19490	19490	1348	2438	15761	6348	478

[1] Coefficients are estimated from the linear probability model.
[2] Besides the covariates indicated above, we further add the gender composition in this analysis.
[3] Robust standard errors are reported between parentheses and are clustered at the father level. Significant at * 10%, ** 5%, *** 1%.

Table 3.7: Robustness Checks: Inclusion of Female Siblings

	All sample		Low status	Regular status		Pre-period	
	OLS	IV		Elite	Non-elite	All	Elite
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of children	-0.0198***	-0.0153*	-0.0037	-0.0490***	-0.0146***	0.0007	-0.0311***
	(0.0021)	(0.0088)	(0.0053)	(0.0101)	(0.0019)	(0.0021)	(0.0096)
Household charac.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10243	10243	715	1260	8268	5111	399

[1] Robust standard errors are reported between parentheses and are clustered at the household level. Significant at * 10%, ** 5%, *** 1%.
[2] In this table, the baseline period is set to be 1780-1880 instead of 1760-1880 as before, since for household sample, we use survey year instead of birth year.

Table 3.8: Robustness Checks: Household-level estimation

C.1: Theoretical Predictions

We consider a standard household utility maximization framework (Galor and Moav, 2002; Shiue, 2017), that the utility of household depends on the consumption level, the number of children, more specific, the number of surviving sons, and the education level of those sons. To examine the trade-off between fertility choice and education, we maximize a log-linear utility function of a given household:

$$U(c, q, h) = (1 - \gamma) \ln c + \gamma(\ln q + \beta \ln h)$$

$$s.t. \qquad c + p_q q + p_h q h \leq Y$$

$$c \geq \tilde{c}$$

where the parameters $0 < \gamma < 1$ and $0 < \beta < 1$ indicates the substitution effects among c, q and h. All income is spent on either consumption at level c, raising a certain number q of children or human capital investment h measured by the proportion of educated sons over total sons. Let the relative price of household consumption be 1, p_q the relative price of household's spending in raising a child, and p_h the relative price of the household's expenditures required for increasing one quality-adjusted child. In this paper, p_h includes the opportunity cost of education, such as profits that the sons would have gained from labor work if they were not working for a degree.

The condition $c \geq \tilde{c}$ indicates a certain consumption level to maintain basic livings. Y denotes the exogenously given household income. The optimal fertility and human capital decisions are obtained from first order conditions:

$$q* = \frac{\gamma}{p_q + p_h h*}$$

$$h* = \frac{\gamma\beta}{p_h q*} \qquad (A1)$$

or

$$h* = \frac{\beta p_q}{(1 - \beta) p_h} \qquad (A2)$$

Assumption A: Equations (A1) suggests a quantity-quality trade-off from $\frac{\partial h*}{\partial q*} <$

0. A given increase in the optimal fertility $q*$ leads to a drop in the ratio of educated kids by decreasing human capital investments. Not surprisingly, that quality and quantity of children are substitutive under a given budget constraint.

Assumption B: Then we allow an increase in the relative returns to education and/or a decrease in the direct or indirect costs of education, which leads to a lower p_h, denoted p'_h. Comparing the ratios of educated sons concerning the fertility before and after such change, we can see that $\frac{\partial h'*}{\partial q'*} < \frac{\partial h*}{\partial q*} < 0$ from equation (A1), which implies a greater trade-off. The increase in the net returns on human capital formation will increase the optimal level of human capital investment, as equation (A2) suggests, and thus leads to a lower fertility rate.

C.2. Sub-population: Booi Aha

In CMGPD-LN, the low-status sub-population, called Booi Aha in Manchu, refers to hereditarily servile people. Booi Aha mainly came from the enslaved Chinese during wartime, criminalized Manchu people (or people from the Eight Banner military system) and their family. In the early Qing era, Booi Aha could be regarded as a slave of their benefactor, so that no freedom of movement, marriage, and working was given to them and their descendant. Given that this hereditary system was perilously blocking social mobility, Qianlong Emperor reformed the Booi Aha's system after 1736: Booi Aha's political status was improved to a large extent. Booi Aha-Manchu people's marriage was legalized and allowed to participate in the civil exam or to become an official (Ren, 2003). CMGPD-LN claims that Booi Aha was totally prevented from attending in the civil exam, but it is contrary to this evidence.

However, this does not mean the studied population is homogenous in access to education or examination. We thus control for father's Booi Aha status and apply subgroup analysis for this reason.