Five Essays on Labor and Public Economics

Five Essays on Labor and Public Economics

A dissertation presented

by

Wei Huang

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

May 2016

Dissertation Advisors:
Professor David Culter
Professor Richard Freeman

Author:
Wei Huang

Five Essays on Labor and Public Economics

Abstract

It is important to understand the individual behavioral responses to public policies and the corresponding social consequences because they are key parameters to evaluate and design *efficient* social policies. In this dissertation, I examine the effects of a series of public policies in China by investigating the policy-induced individual behaviors and social consequences in different stages over lifetime. Chapter 1 examines the impact of fertility policies on the education investment in girls; Chapter 2 shows how the ethnic-specific terms in the OCP distorted individual behaviors and equilibrium outcomes in marriage market; Chapter 3 examines an unintended outcome of birth control policies - more reported twins; Chapter 4 uses the compulsory schooling laws (CSL) as exogenous shocks to estimate the causal effects of education on health at prime ages; Chapter 5 examines the effects of new social pension provision on the outcomes of the elderly, including income, expenditure health and mortality. I conclude that the public policies have significant and remarkable effects on the behaviors and welfare outcomes of individuals. In addition, these lessons from China may shed lights on the some important and general interested questions in economics.

Contents

List of Tables

List of Figures

Acknowledgments

Above all, I would like to express my special appreciation and thanks to my dissertation committee: Professor David Cutler, Professor Richard Freeman and Professor Adriana Lleras-Muney. I would like to thank them for encouraging my research and for allowing me to grow as an economist.

I met Richard at NBER-CCER conference in Beijing. I was preparing for my PhD application, and he was really nice and agreed to write a recommendation letter for me. It is Richard's email that changed my life ever since: "*Hi Wei, I don't know long it takes for the official notification but the news will be very good for you so relax, enjoy, and get ready for a trip to US and an exciting intellectual time...* " So I came to Harvard! Since I reached Cambridge in July 2011, Richard has been consistently providing his great guidance to me for my whole PhD period. He has a great mind and always provides brilliant ideas. Just after a couple of months I arrived here, Richard had an idea on ethnic co-authorship in the US and offered me the opportunity to coauthor with him. We talked at least once a week during the co-authorship period, and he taught me, step by step, how to think of a great idea, how to do rigorous analysis, and how to write a good paper, and even how to get a catchy topic, etc. I benefited tremendously from his guidance.

I knew David's name far before I came here. When I was a master student in China Center for Economic Research at Peking University, "Understanding differences in health behaviors by education" was among the first few economics papers I read in my life. This paper was just written by David and Adriana. Afterwards, I read a number of papers by David and became very interested in health economics. That is the reason I decided to register his Health Economics course when I was just a first year student here. This was the most important decision I made during my PhD. Since then health economics has become one of my major interested fields and over half of my research papers in my PhD period were about health issues. In the first few years, I usually talked with David about my ideas and thoughts. Although some of them are really preliminary and even naive (as I see them now), David always provided insightful suggestions and encouraged me to think more deeply. During thesis writing period, for every paper I wrote, he always provided detailed comments, constructive suggestions, and profound insights. After reading my papers, David usually asked me some enlightening questions. Among these questions, two affected me profoundly. The first one is "Why this question is important?" and the other is "Why do we need to learn from Chinese experience?" Hopefully I answered these questions well in this dissertation.

It is a great luck for me to have Adriana (i.e., the other author of "Understanding differences in health behaviors by education") in my dissertation committee. I never thought that I would coauthor with David and Adriana when I was reading their papers in the library in Peking University. In my third year of PhD, however, I did! Since then, I talked with Adriana frequently via telephone or Skype, and we usually talked for one hour or two on our coauthor projects and my research. She provided very detailed suggestions for every paper I wrote in the past few years. Our interests are very close and I benefited a lot from our discussion and conversations. I still remember the encouraging words Adriana told me when I went to UCLA to attend a conference. She said to me: "Wei, there are many questions to answer but I think you should do something important." I tried to do so in last few years and will go on in the future.

I would also thank Professor Amitabh Chandra, Professor Raj Chetty, Professor Edward Glaeser, Professor Claudia Goldin and Professor Lawrence Katz. During my PhD period, I also talked with them for many times and they also provided really good guidance and helpful suggestions on my research. Many other professors also helped me a lot. I cannot list all names here, but they are always in my mind. In addition, I want to thank my coauthors in the dissertation. They are Xiaoyan Lei, Ang Sun, Chuanchuan Zhang, Yaohui Zhao, and Yi Zhou. Thank them for their understanding and encouragement in my many many moments of crisis.

Finally, I would thank my family. They are most important people in my life. I would thank my parents first. They are both over 50 years old now but I cannot see them frequently. I can always get consistent support from. I also thank my younger brother because not only

he helped to take care of our parents but also it is him who let me know what is "One-Child Policy" in China. Our family paid the fertility penalties! I would also thank my grandfather and grandmother. I was living with them for 16 years and that is why I am so interested in topics about health and aging. When I was writing up the fifth chapter, I even asked my grandfather about his pension program when he was alive. In the end, I would like to thank my wife, Mi Luo. She is also an economics PhD and has provided many valuable comments to the papers. More importantly, she is always a strong supporter when I am not so confident.

This thesis is only a beginning of my journey.

To my parents

Introduction

The five essays in this dissertation investigate the impacts of public policies on individual behaviors at different stages in life. The five essays are ordered based on the stage in life cycle when the effects or consequences are in present. Generally, the findings in this dissertation suggest that the policies have significant and remarkable effects on the behaviors and welfare outcomes.

The first chapter investigates how the fertility policies affect the education investment in girls. Education investment is an important factor that contributes to economic growth in the long run and happens in the early stage of life cycle. Using the temporal and regional variation in financial penalties of One-Child Policy in China, I find that higher fertility fines during teenage years raised women's education, which explained 30 percent of the increase in 1945-1980 birth cohorts. The impacts of the policies on girls' education are significantly associated with those on female labor force participation, non-manual occupation, delayed marriage and childbearing, and attitudes regarding children and gender equality.

The second chapter examines how the ethnic-specific terms in the OCP distorted individual behaviors in marriage market. Marriage, which usually happens in early adulthood, is an important source of happiness and plays an important role in generating and redistributing welfare among individuals. Results in the second chapter provide new evidence for the transferable utility model by showing China's One-Child Policy induced a significantly higher unmarried rate among the population and more interethnic marriages. The results show that he welfare loss caused by the "taxes on children" is from not just the distortion in fertilities but also the policy-induced distortion in marriage market. Specifically, the total welfare loss, caused by reduced fertility and marriage distortion, is around 4.9 percent of annual household income, with marriage distortion contributing 17 percent of this. The findings in this chapter shows unintended behavioral with respect to the public policies may matter significantly when conducting relevant evaluations.

Since fertility generally happens later than marriage, the third chapter continues to examine how the "one-birth" restriction of the birth control policies induced people to give birth to twins in China. The results show that the One-Child Policy has accounted for more than one-third of the increase in the reported births of twins since the 1970s. The results suggest that one in six reported twins in the late 1990s are made intentionally to bypass the fertility policy regulations. Investigation using birth space with prior births and height difference within twins suggests that the increase in births of twins is partly due to parents reporting regularly-spaced children as twins to avoid the policy violation punishment. These findings highlight the unintended behavioral responses to public policies and corresponding social consequences.

The fourth chapter examines the effects of education on health at prime ages by investigating the exogenous variation in the implementation of "nine-year" Compulsory Schooling Laws initiated in 1986 in China. I find that compulsory schooling laws in China increased the education of eligible-cohorts by 1.1 years and significantly reduced the rates of reported fair or poor health, underweight, and smoking. Investigating the mechanisms finds that cognition and income only explain 15 percent and 7 percent of the effects on self-reported health. Spillovers from increased education of other people in the local region could explain over 25 percent. These findings present new evidence for the causal effects of education on health and help to reconcile the mixed findings in the literature.

The last chapter examines the effects of new pension scheme provision on outcomes of the elderly. Analysis using historical data from 10 countries shows that mortality of the age-eligible elderly reduced by 1.7-2.2 percent just after the social pension provision. Using the institutional variation of a new social pension provision in rural China, we further find that, among the pension-eligible people, the scheme increased their household income and food expenditure by 17.6 and 9.6 percent, and reduced labor supply and health insurance participation by 6.2 and 5.7 percent. In addition, it also significantly improved their health status in terms of less reported disability, underweight and lower mortality rate. These findings suggest that the effects of pension policies would have large and significant effects on the lives of the elderly.

Most of the policies examined and the data used in this dissertation are about China. The fertility policy in the first three chapters refers to the "One-Child" Policy (OCP) started in 1979, and the other two are the "nine-year" Compulsory Schooling Laws initiated in 1986 in China, and the New Rural Pension Scheme that took place in 2009. I use these policies and Chinese data not just because my family and I experienced them and thus knew them well. Actually I found that China provided unique or natural settings to shed some lights on some important questions that are interesting to all economists, which I will explain specifically in each chapter.

Chapter 1

When Fewer Means More: Impact of One-Child Policy on Education of Girls[1]

1.1 Introduction

The rise in the educational attainment of women has been one of the most significant changes during the past century in many countries (Goldin *et al.*, 2006; Goldin and Katz, 2009; Duflo, 2012).[2] It was accompanied by an increased involvement of women in the economy and an enhancement of their socioeconomic status both in society and in the family. These changes, named "Quiet Revolution" by Goldin (2006), have brought hot discussion over the potential impacts on the economic development.

Although the phenomenon has been well documented, the forces responsible for the "Revolution" are not as clear. The various explanations in the literature indicate no "silver

[1]Co-authored with Xiaoyan Lei and Ang Sun.

[2]For another, Goldin and Katz (2009) showed in their Figure 1.2 (pp. 17) that the male-to-female ratio in secondary school enrollment is less than one for most countries. This is especially true for countries with a real GDP per capita greater than $3000 (PPP 2000 $).

bullet" answer to the question across nations.[3] For example, Goldin and Katz (2002) and Bailey (2006, 2010) found that the oral contraceptives pills contributed much to women's increased education and rising role in the labor market in the United States. However, the pills' power may not apply in East Asian countries like China because a much smaller proportion of women take contraceptive pills: only 1.7 percent of women of reproductive age (15 to 49 years) in China take contraceptive pills, compared to 15.6 percent in the United States (UN, 2006).[4]

This paper examines the impact of the implementation of the birth control policies during women's teenage years (i.e., 10-19 years of age) on their educational achievement.[5] The birth control policies during the teenage years may matter to the educational attainment of girls in several aspects. First, since stricter fertility policy indicates higher cost of giving birth to more children when the girls grow up and enter fertility ages, the girls themselves (or their parents) would re-optimize the timing of their marriage and fertility. They will most likely delay their marriage and childbearing, which lowers the costs for a longer duration of high school education (Field and Ambrus, 2008). Second, with fewer children to care for, women are expected to be more involved in labor market activities. Third, the expectation of being supported by fewer children during old age is likely to generate an incentive to acquire more education, land a better job, and save for retirement. Finally, these birth control policies may, on the whole, increase women's socioeconomic status including education (Chiappori and Oreffice, 2008). Therefore, the fertility policies increased the motivation of educational investment in girls. In addition, the asymmetric fertility-related costs between genders are worthy of note: women bear the major share of the costs of childbirth and are

[3]For other explanations, Jayachandran and Lleras-Muney (2009) found that the sudden reduction in maternal mortality rate as a result of health policies in Sri Lanka led to a convergence in the education level of boys and girls because parents expect that investment in girls will be relatively more valuable. Chiappori *et al.* (2009) emphasized the role of the labor market and the marriage market in investment in education for women. Jensen (2012) pointed out that (future) labor market opportunities for women influence educational investment in India.

[4]In addition, the declined maternal mortality may not be an explanation in China either, because the maternal mortality rate in China in 1990 was less than 0.097 percent, which is much lower than that quoted by Jayachandran and Lleras-Muney (2009).

[5]For simplicity, we use "teenage years" and "10-19 years of age" interchangeably in this paper.

generally the primary caregivers. This implies no or smaller effects of the fertility policies for men, suggesting that the fertility policies could presumably explain the narrowing of the gender gap in education.

The birth control policies restricted the fertility of millions of couples in a number of countries for many decades (Miller and Babiarz, 2014) and an established literature has investigated the quantity-quality trade-off phenomenon, but the potential effects of policy implementation at adolescent ages or teenage years on women's education received relatively less attention from researchers.[6] One exception, Miller (2010), found that the policy introduction during fertility ages decreases fertility and increases the women's socioeconomic status in Colombia. However, Miller (2010) did not focus on education because most women finished their education before fertility ages, and he did not provide evidence for the possible mechanisms, either. This study aims to fill this gap.

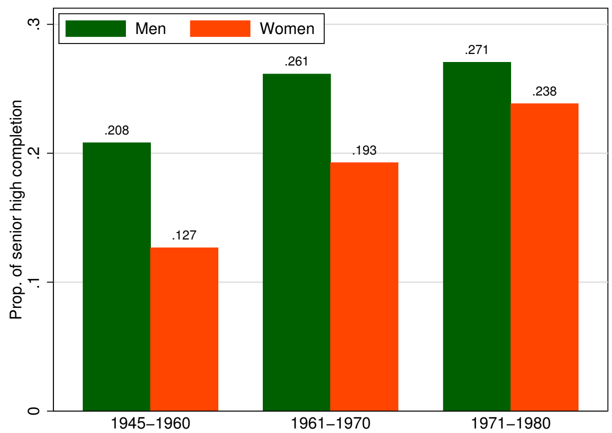
China provides a *unique* opportunity to investigate the relationship between women's rising education and fertility policies. For one thing, China experienced a rapid increase in female education and a convergence of the gender gap in the late 20th century (Rosenzweig and Zhang, 2009). Figure 1.1a shows that the rate of senior high school completion for women almost doubled from 13 percent among those born between 1945 and 1960 to 24 percent for those born in the 1970s. Meanwhile, the rate of senior high school completion for men only increased from 21 to 28 percent; thus, the gender gap narrowed by half from 8.1 to 3.3 percentage points. Figure 1.1b shows a similar pattern for college completion.

In addition, China formally initiated birth control policies in 1979, also known as the "One-Child Policy" (OCP) (Greenhalgh, 1986). Unlike birth control policies in many other countries, the OCP *compulsorily* assigned a "one-birth" quota to each couple in general, and there was great variation in its implantation across regions and ethnicities and over time. This policy led to the involvement of hundreds of millions of couples in China in a strict

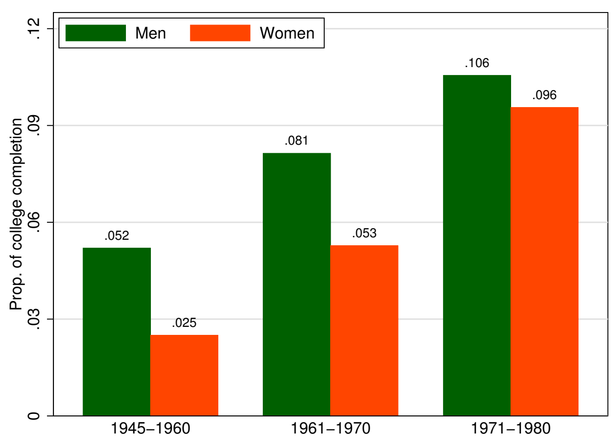
[6] Specifically, the literature on quantity-quality trade-off mainly focused on the effects of number of siblings on the education attainment but this paper mainly examined the effects of the OCP implementation *during age 10-19* on the girls' education attainment with controlling for the dummies of number of the siblings. In addition, the previous literature using the OCP (e.g., Qian, 2009; Rosenzweig and Zhang, 2009) found weak or no evidence on quantity-quality trade-off effects.

Figure 1.1: *Senior high school and college completion over birth cohorts 1945-1980, by gender*

(a) *Senior high school completion rate over birth cohorts*



(b) *College completion rate over birth cohorts*



Notes: Data source is Censuses 2000 and 2005. Mean rates of senior high school (or college) are calculated in the birth cohort groups and labeled above the bars.

family-planning program that lasted more than 30 years. Accordingly, the total fertility rate dropped from 2.81 in 1979 to 1.51 in 2000 (World Bank). It can also be seen in Figures 1.1a and 1.1b that the gender gap in education narrowed with the 1960s and consequent birth cohorts. The 1960s birth cohorts were in their teenage years when the policy was initiated.

Our analysis shows that the above finding is not a mere coincidence. Using both the regional and temporal variation in the average financial penalties for one unauthorized birth in the province-year panel from 1979 to 2000 (Ebenstein, 2010; Wei and Zhang, 2011; Huang *et al.*, Forthcoming). The fertility penalties are formulated in multiples of local household annual income, and have a large variation across regions over time. We find that an increase in OCP fines by 1 year of local household income during teenage years predicts an increase of over 2 percentage points in the rate of senior high school completion among women of Han ethnicity. The magnitude of the estimates is large, indicating that the fertility policy explains 30 percent of the increase in women's education and 50 percent of the narrowing of the gender gap in the birth cohorts from 1945 to 1980.

The identification assumption of *exogenous* variation of the fertility penalties should not be taken for granted. We thus conduct a series of empirical tests and all of them provide consistent evidence for the exogeneity of fertility fines used in this paper. First, we control for local population growth and regional assignment of educational resources and find robust results. We also investigate the regional heterogeneous effects in urban and rural areas in which the strictness of policy implementation differed and consistently find greater effects in regions with stricter implementation of the policy. Second, we also investigate the effects of the fines on two plausible comparison groups: (1) women of minority groups, because they are less subject to the fertility policy and (2) men of Han ethnicity, who, as noted earlier, take less responsibility for childcare and presumably less affected by the fertility policy. We consistently find no effects of the fines in either group. Finally, we conduct a placebo test and find that the education level of local women has no predictive power on the future fertility fine rate changes.

Investigating China General Social Survey (CGSS) data with more family background information, we further find greater effects of the OCP when parents have a higher level of education, work in the public sector, or have a higher administrative rank. These findings are consistent with the implementation of the policy: individuals who work in the public sectors or have a higher level of socioeconomic status face stricter enforcement and strengthened regulations.

We further investigate the possible mechanisms by exploring the correlations between the effects of fertility fines on girls' education and the effects of the fertility fines on later outcomes of the girls. We consistently find that the greater positive effects of the policy fines on girls' education are associated with greater positive effects on late marriage, as measured by marriage by 25 years of age; late fertility, as determined by whether the woman ever had a child; labor force participation and employment in a professional (or white collar) job. Parallel analysis in CGSS also finds consistent evidence for attitudes on gender equality and dependence on children. These associations provide some *suggestive* evidence for the explanations or mechanisms for policy-induced higher education of the women.

This study builds up several literatures. First, it contributes to the literature on women's empowerment (e.g., Goldin and Katz 2002; Goldin *et al.* 2006; Goldin 2006; Goldin and Katz 2009; Duflo 2012) by providing a new explanation from the policy side that the implementation of the OCP during the teenage years increases women's educational attainment. Second, it also contributes to the literature on education investment (e.g., Chiappori *et al.*, 2009; Jensen, 2000, 2010, 2012; Bursztyn and Jensen, 2015) by examining the effects of fertility policies. Finally, these findings are also relevant to the effects of expectations for the future on current behaviors (e.g., Manski, 2004; Jayachandran and Lleras-Muney, 2009; Oster *et al.*, 2013).

The paper is organized as follows. The next section introduces the background and our measurement for the OCP. Section III introduces the micro level data used in the analysis, including the Population Census 2000, the 2005 One-Percent Population Survey, and the CGSS. Section IV shows the empirical results for the effects of the OCP on educational attainment, and Section V provides additional results for possible mechanisms, and Section

VI concludes.

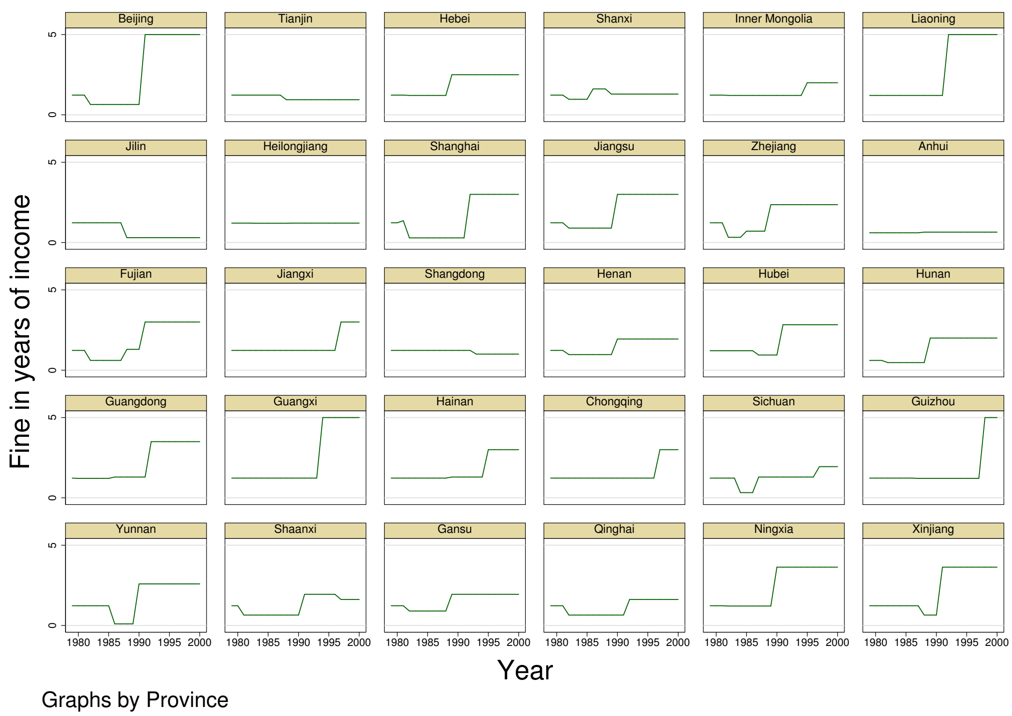
1.2 Background and One-Child Policy

In the 1970s, after two decades of explicit encouragement of population growth, policymakers in China enacted a series of measures to curb population growth, especially within the Han ethnicity. The OCP was formally conceived in 1979, and enforcement gradually tightened until it was firmly established across the country in 1980 Banister (1991). This was the first time that family planning policy formally became one article in the laws of China. Enforcement varies across regions. Studies have found that the strictness of enforcement is well reflected by the monetary penalties and subsidies that have been implemented since 1979. The measure used for the OCP in this study is the average monetary penalty rate for one unauthorized birth in the province–year panel from 1979 to 2000, which is formulated in multiples of local household annual income. It is called a "social child-raising fee" in China, and this paper uses "policy fine" or "fertility fine" for the sake of brevity. Figure 1.2 shows the pattern of policy fines from 1980 to 2000 in each province.[7]

The OCP in principle restricts a couple to having only one child. However, the de facto regulations vary among different regions and ethnicities. In 1984, the Central Party Committee issued "Document 7" in recognition of the diversity of demographic and socioeconomic conditions across China. The document stipulated that regulations regarding birth control were to be made in accordance with local conditions and to be approved by the provincial Standing Committee of the People's Congress and provincial-level governments, which devolved responsibility from the central government to the local and provincial governments (Baochang *et al.*, 2007). On the one hand, implementation of the OCP differs for Han and for minorities. It is mainly focused on the Han ethnicity, which is the largest ethnic group

[7]The amount of the policy fine collected were not made public until recent years: the total was about 20 billion RMB yuan (3.3 billion US$) among 24 provinces that reported in 2012. For example, Guangdong, one of the richest provinces in China, collected 1.5 billion. In comparison, the total local government expenditure on compulsory schooling was 10.5 billion, suggesting that the policy fines composed a sizable part of the local fiscal revenue.

Figure 1.2: *OCP fertility fine rate over year, by province*



Notes: Data source is Ebenstein (2010). The unit of penalties is times of local household annual income.

in China and makes up 92 percent of the population. In all provinces, most minorities are allowed to give birth to another child or have no restrictions. On the other hand, the regulations of the OCP also differ in urban and rural areas. The policy is strictly enforced in urban areas but less so in rural areas. For people with urban *hukou*, the policy that one couple is only allowed to have one child is strictly enforced; whereas couples with rural *hukou* are allowed to have a second child if the first is a girl, which is also named the "One-and-Half Children Policy." Besides the monetary penalties, we also exploit the time-invariant variation in the policy implementation by ethnicity and type of *hukou* to testify to the consistency of the results.

The public expectation for the number of births can be formed only under clear awareness and harsh enforcement. The fertility policies in China meet these requirements. Population and Family Planning Commissions (PFPC) were set up at every level of government to raise awareness and to carry out registration and inspection work. A large-scale public campaign about the law was launched during the 1980s, and an effective curb on population growth became the highest priority for local officials. To ensure the fine's enforcement upon violation, those who have an "illegal" birth but do not pay the fine can be sued by the local PFPC and the fine will be collected compulsorily. The provincial governments also set up detailed regulations to ensure the effective collection of policy fines. For example, the illegal birth child cannot be registered to the *hukou* system if the fines are not paid; thus, the child cannot go to school because the *hukou* is a requirement at school entrance. Property can also be confiscated for nonpayment of policy fines.[8]

[8] In an extreme case in Shaoyang City, Hunan Province, the "illegal" birth children themselves were confiscated and settled in the social welfare institution because the policy fine was not paid. (Source in Chinese: http://baike.baidu.com/subview/5708887/5757115.htm)

1.3 Data

1.3.1 China Population Census 2000 and 2005 One-Percent Population Survey

The main data used in this study are taken from the 2000 Population Census and the 2005 One-Percent Population Survey (Censuses 2000 and 2005, hereafter). Both datasets contain gender, education level, year and month of birth, type of *hukou* (urban/rural), *hukou* province, ethnicity (Han/Minorities), marital status, number of siblings, and relation to the head of household.[9] For married respondents, the data also provide the year and month of the first marriage. For ones older than 16 years of age, the data also provide information about labor market participation, including working status, occupation, and the number of days of work in the past week. Because the sampling rate differs in the two datasets, sampling weights are applied throughout this analysis.

For the analysis of education, we keep respondents aged between 25 and 55; thus the earliest birth cohort in our analysis is 1945 and the latest 1980. We keep those above 25 years of age who were born no later than 1980 because (1) most of the respondents have completed their formal education before the survey and (2) we want to avoid the identified effect of the OCP originating from a smaller number of siblings.[10]

Education is divided into six levels: illiterate, primary school, junior high school, senior high school, college, and graduate or above. We combine the highest three categories into "senior high or above" and choose it as our main outcome of interest for at least three reasons. First, the number of dropouts after finishing junior high (i.e., just before senior high) is the largest at any stage of education, implying that this is likely to be the most responsive margin regarding educational investment. In particular, the dropout rate is 68 percent, whereas that just after primary school is only 36 percent; second, the compulsory

[9] The number of siblings is provided for those below 30 years of age in 2005, i.e., those born later than 1975 in 2005 data. The results are nearly the same regardless of whether the number of siblings is controlled.

[10] We further include the number of siblings where available to control for the effect of the OCP on education through the Quantity-Quality trade-off channel. We also compare the results using the 2000 census when no sibling information is provided and those from the 2005 population study survey with the number of siblings controlled and find that the two sets of results are nearly the same. These suggest that the results should not be driven by the channel of number of siblings.

schooling laws initiated in 1986 require 9 years of school (i.e., junior high school) in China, and we wanted to prevent our estimates from being contaminated by the laws' effect; third, the college completion rate is fairly low for these cohorts (i.e., less than 6 percent) and thus may not have large variation for identification. But we still explore college completion and find the results are consistent, though we do not report them here.

1.3.2 China General Social Survey (CGSS)

The CGSS, launched in 2003, gathered longitudinal and nationally representative data on social trends in the mainland of China, with the targeted population being civilian adults 18 years of age and above. The CGSS provides subjective evaluation on certain aspects of life of interest to our study. Four survey questions are relevant to our analysis and are consistently measured in 2010 and 2012 waves.[11] The first two questions concern attitudes toward children: "Do you agree that the happiest thing is watching children growing up?" and "Do you agree that adult children are important support for old people?" The answers are 1 = "strongly disagree," 2 = "disagree," 3 = "neither agree nor disagree," 4 = "agree," and 5 = "strongly agree". Similarly, the respondents are also asked "Do you agree that men and women have equal housework duty?" We use the answers coded from 1 to 5 directly. In addition, there is another question: "What is the optimal arrangement in the presence of a preschool child at home?" We code the answer as 1 if the respondent chose "Women at home and men work outside" and 0 for other answers that indicate that men share the responsibilities of child rearing.

To make the analysis consistent, we confine the sample to those born between 1945 and 1980. The rate of senior high school completion is 24 percent for women and 31 percent for men. The difference between the CGSS sample and the Census is likely to originate from different sampling frames. Caution should be used when comparing the results of CGSS to those of the censuses.

[11]Since 2003, the CGSS has three different sampling designs and has used three sets of sampling frames: 2003 to 2006, 2008, and 2010 to the present. Thus the two waves 2010 and 2012 used in this study are consistently surveyed under the same sampling frame.

1.4 Effect of the OCP During the Teenage Years on Education Attainment

1.4.1 Educational Attainment Response to Fertility Policy Fines Changes

We begin the analysis by investigating how educational attainment responds to the variations in the fertility fine. We show how the changes in the senior high school completion rate are correlated with the changes in the fine rate during the teenage years. The changes are calculated using two consecutive birth cohorts with the same *hukou* province and the same type of *hukou* (urban/rural).

Figures 1.3a-1.3c report the nonparametric estimation when plotting the changes in the senior high school completion rate against those in the fine rate, in three separate samples—Han women, Han men, and minority women, respectively. All the results are weighted by the population in each birth cohort–province–*hukou*. Table 1.1 reports the corresponding OLS results.[12]

Both Figure 1.3a and the first column in Table 1.1 show a significant positive correlation between an increase in the fine rate during the teenage years and an increase in the rate of senior high school completion. The estimation suggests that a 1-year–household–income increase in the fine rate during the teenage years predicts a 2-percentage-point increase in the rate of senior high school completion. In contrast, Figures 1.3b and 1.3c and the remaining columns of Table 1.1 suggest no correlation either among Han men or among minority women.

1.4.2 Effect of the OCP on Education Attainment

To obtain a more precise estimation, we control for potential individual heterogeneity and conduct the following regression:

[12] In section 4.1, we only retain the ones born no earlier than 1965 because the change in the fine rate during the teenage years for these earlier cohorts is almost zero. Therefore, the number of observations is 900, as there are 30 provinces, 2 types of *hukou*, and 15 birth cohorts, except in the case of minority women, for whom some observations are missing. In the later sections, we use the full sample so that we can control for the pre-trends in each province.

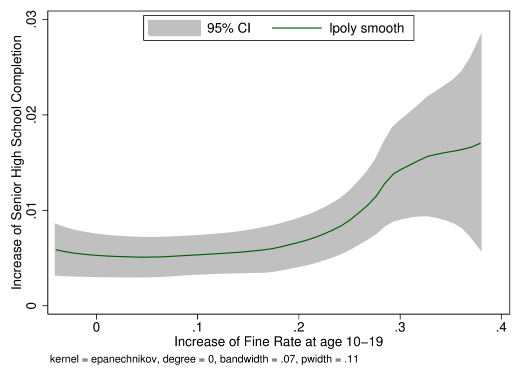
Table 1.1: *Education Responds to the Fertility Fine Changes, by Gender and Ethnicity*

	(1)	(2)	(3)
Dependent variable	Increase in Senior High School Completion Rate		
Sample	Women of Han	Men of Han	Women of Minorities
Change in Fine Rate	0.0211***	-0.00330	0.0112
at Age 10-19	(0.00732)	(0.00687)	(0.0121)
Observations	900	900	832
R-squared	0.004	0.000	0.000

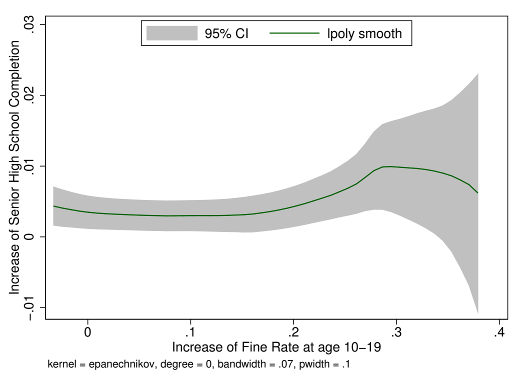
Notes: Sample is from Censuses 2000 and 2005. Birth cohorts from 1965 to 1980 are kept. The independent variable is the change in the senior high school completion rate in two continuous birth cohorts in the same *hukou* province-type of *hukou* group. The dependent variable is the corresponding change in the fertility policy fine rate at 10 to 19 years of age in the cell. The ideal observation is 30 × 2 × 15 = 900. Some observations are missing in the sample of women from minorities. Huber-White robust standard errors that allow for clustering by *hukou* province to account for possible serial dependence in errors across cohorts within *hukou* province are presented in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1.3: *Associations between Increase of policy fine rate and increase of senior high school completion, by gender and ethnicity*

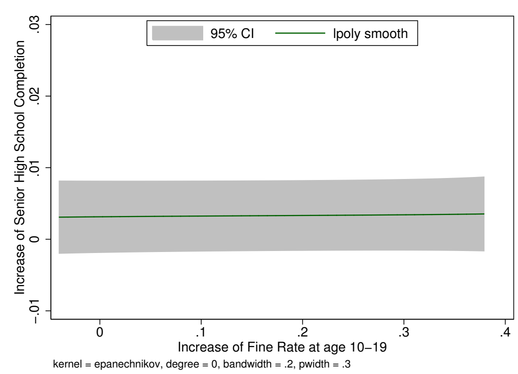
(a) *Women of Han*



(b) *Men of Han*



(c) *Women of Minorities*



Notes: Data source is China Censuses 2000 and 2005. Birth cohorts from 1965 to 1980 are kept. The figures plot the change in the senior high school completion rate in two continuous birth cohorts in the same *hukou* province-type of *hukou* group against the change in the fertility fine rate in the corresponding two continuous birth cohorts. Kernel-weighted local polynomial smoothing is applied.

$$Senior_{ijbt} = \beta_0 + \beta_1 Fine_{jb}^{10-19} + \beta_2 X_i + \delta_b + \delta_t + \delta_{bt} + \delta_j + \theta_j Prov_j \times T + \varepsilon_i \tag{1.1}$$

where subscript i denotes individuals, j *hukou* provinces, b the year of birth, and t the year of the survey. The dependent variable $Senior_{ijbt}$ is an indicator for individual i born in year b in j *hukou* province in survey year t who completed senior high school.

$Fine_{jb}^{10-19}$ denotes the mean value of the fine rate in *hukou* province j in the years in which birth cohort b was between 10 and 19 years of age. We use the fine rate in the province in which an individual registered her residency (*hukou*) because the respondents are more likely to have obtained education in their *hukou* province than in the province in which they currently lived. Inter-province migration could be a potential issue because it is possible that the *hukou* will be moved to another province. The census 2005 data provide information on the birth province, which shows over 95 percent of people have the same *hukou* province as the province where they were born and that over 93 percent of individuals are living in the same province in which they were born. It is noteworthy that the results are consistent if we use the birth province to match.

We use the fine rate at 10-19 years of age because it is when households (or girls) make decisions for their children's (or their own) education - senior high school starts at 15-17 years of age and college at around 18-21. The section 4.3 examines the effects of policy implementation at different ages and finds consistent results.

The set of control variables, X_i, includes the logarithm of the birth cohort size and indicators for the type of *hukou* (urban/rural) and the number of siblings. δ_b, δ_t, and δ_{bt} denote the indicators for the birth cohorts, the year of the survey, and the corresponding interactions between the two. We control for the interactions to avoid any systematic sampling difference in a specific birth cohort between the 2000 and the 2005 survey.[13] δ_j denotes the dummies for *hukou* province that capture the time-invariant heterogeneity across provinces, and $Prov_j \times T$ are the province-specific birth linear time trends to control for any potential trend in feminist views or attitudes toward women's education and careers in each

[13] The results are consistent if we drop the interaction terms.

province.

The coefficient β_1 is of our main interest because it presents the impacts of the OCP fine rate during the teenage years on the likelihood of senior high school completion. Panels A and B of Table 1.2 present the OLS point estimation for β_1 among women and men of Han ethnicity, respectively. The estimates in column 1 indicate that the OCP fines for an additional illegal birth during the teenage years is positively associated with a statistically significant increase in the likelihood of completing senior high school for women, but not for men. The coefficient is very close to that in Table 1.1. In addition, the magnitude is also economically significant: it implies that the increase in the OCP fine (i.e., 1.4 times of local household annual income) contributes a 2.8-percentage-point increase in the senior high school completion rate of women for the birth cohorts from the late 1940s to the 1970s, suggesting that the OCP explains about 30 percent of the educational increase for the women born during that time.

Based on the benchmark analysis, we include a series of control variables to further address the potential issue that the fertility fine might be correlated with other factors that drive female education attainment. According to the documentation of the OCP, the fertility policy fine may change according to the population size and its potential increase, which may be correlated with women's educational level due to competition and other resources allocated by the central government to the province. Therefore, in column 2, we also control for the logarithm of the population sizes in the local *hukou* province when the respondent was 10 and 19 years of age. The estimates are very close to the first column for both men and women, suggesting that the incentive to raise the policy fine caused by a potential population increase is not an important confounding factor.

To control for the potential correlation between the fertility fine and government support of education, we include variables that capture the education supply, including (the logarithm of) the number of primary schools and junior high schools. The results in column 3 show that the effect of the policy fine is quite robust, suggesting that the fertility policy–induced education should mainly originate from the forces from the education demand

Table 1.2: *Effects of the OCP Fertility Fines on Education (Senior High or above)*

	(1)	(2)	(3)
Dependent variable	Senior high or above (Yes = 1)		
Panel A. Female sample	*Mean of Dep. Var. = 0.181*		
Fertility fine rate at age 10-19	0.0206***	0.0207***	0.0222***
	(0.00394)	(0.00395)	(0.00395)
	[0.00997]	[0.00979]	[0.00960]
Observations	810,994	810,994	810,994
R-squared	0.336	0.336	0.336
Panel B. Male sample:	*Mean of Dep. Var. = 0.244*		
Fertility fine rate at age 10-19	0.00391	0.00426	0.00358
	(0.00406)	(0.00408)	(0.00403)
	[0.00439]	[0.00444]	[0.00470]
Observations	792,075	792,075	792,075
R-squared	0.289	0.289	0.289
Covariates controlled for in both panels			
Basic control	Yes	Yes	Yes
Population size and increase	No	Yes	No
Education supply condition	No	No	Yes

Notes: Sample is from Censuses 2000 and 2005. Basic control includes indicators for type of living residence, number of siblings, *hukou* province, year of birth, survey year, interactions between year of birth and survey year, and provincial year of birth (YoB) linear trends. Population size and increase includes the logarithm of the population sizes in the local *hukou* province when the respondent was 10 and 19 years of age. The education supply condition includes the number of schools in middle high school in the local province when the respondent was 15 years of age. Huber-White robust standard errors in parentheses allow for clustering of errors within *hukou* province–year of birth cells. More conservative standard errors that allow for clustering by *hukou* province to account for possible serial dependence in errors across cohorts within *hukou* province are presented in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

side, such as the higher policy-induced incentive in education among girls themselves or their parents.[14]

Because men bear smaller costs for childbirth and are generally the secondary caregivers, we expect smaller or no effects for men, and, therefore, men compose a natural comparison group. Because our identification is based on geographical and temporal variation in implementation of fertility policies, the possibility of region-birth cohort level omitted variables (e.g., education policies or economic development) also motivates this identification strategy of using males as comparison group. Smaller and insignificant estimates for men could serve as a piece of evidence for the exogeneity of fertility fines because similar effects for men should be identified if the effects of the fertility fines on women's education had come from other confounding factors. The estimates in Panel B shows the results for Han men, which is fairly consistent with the expectation. Specifically, fertility fine rate at 10-19 years of age is insignificantly associated with men's senior high school completion rate. The magnitude is much smaller as well, which is only one fifth of that among women. Comparing the effects for women and men, the OCP explains about 50 percent of the narrowing of gender gap in education shown in Figure 1.1a.

1.4.3 Effects of Fertility Policies at Different Ages

The above analysis established the relationship between the policy restriction during the teenage years and girls' education, and thus it is natural to ask whether policy changes at other ages matter. Because the decision to go to senior high school is made when children are teenagers, it is reasonable that policy changes that occur at these ages yield largest effects. Investigation of the relationship between policy changes at even younger ages also

[14] All of these results are robust when using logit estimation. We also interact the survey year dummy (equals 1 if the respondent is surveyed in 2005) with the policy fine at 10 to 19 years of age to test for a difference between the sample from the 2000 census data and that from the 2005 population survey. The coefficient on the interaction term is small and insignificant, indicating roughly similar effects of OCP fines on women's educational attainment in the two survey years. For the other, we interact the policy fine from 10 to 19 years of age with birth cohorts no earlier than 1969 to test whether the effect of the OCP is heterogeneous across different birth cohorts, with the concern that changes in social norms across the birth cohorts may be correlated with both the increase in the fine rate and the education of women. The results show no evidence for this concern. These results are not provided here but are available upon request.

shed light on the question of whether the effect of the OCP is time-accumulative or whether changes at certain ages matter most.

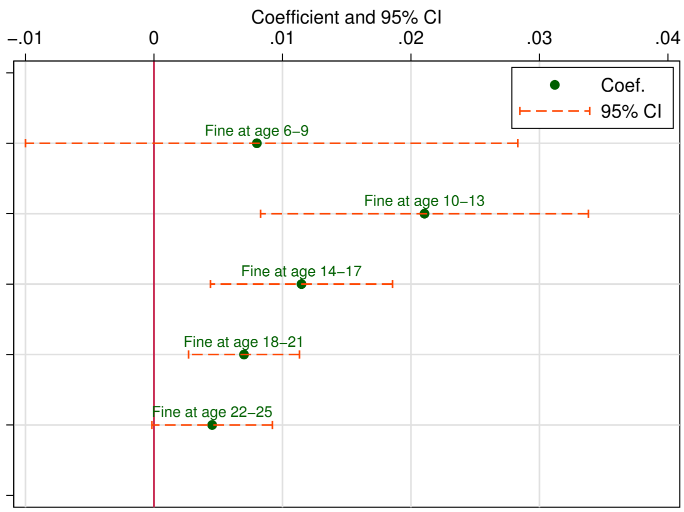
To answer the question above, we replace the fertility fine at 10 to 19 years of age with a series of fertility fines at different ages (i.e., fertility fines at 6-9, 10-13, 14-17, 18-21 and 22-25 years of age) in the regression. Figure 1.4 reports the results. The coefficient on fertility fines at 6-9 years of age is 0.007 and insignificant. That on fines at 10-13 years of age is three times of it and very significant and the effects gradually diminish as age increases. There is no evidence that the fertility fines at ages 22-25 has effects on the senior high school completion. These results are consistent with the patten of schooling ages in China: children generally begin senior high school at around 15-17 years of age and begin college at around 18-21 years old. These results suggest that a change in the policy fine rate matters mostly when it happens during the teenage years, the most critical time when households (or girls) make decisions for their children's (or their own) education. The small and insignificant coefficient of fines at 6-9 years of age suggest the effects of OCP not time-accumulative.

1.4.4 Heterogeneity: By Ethnicity, Type of *Hukou*, and Policy Implementation

Because of political reasons, most minorities are exempted from the restrictions of the policy (Baochang *et al.*, 2007). This paper follows the ideology from prior research to use minority women as another comparison group for the effects of OCP (Li *et al.*, 2011; Huang *et al.*, Forthcoming). Figure 1.5 shows the senior high school completion rate for women in the birth cohorts between 1945 and 1980 among Han and minority women, respectively. The senior high school completion rate among minority women is lower in general. More importantly, the gap between these two groups expanded in the post-1960 cohorts, i.e., the first cohorts to experience the OCP during their teenage years, suggesting that the OCP during the teenage years is likely to be an important factor contributing to female education improvement.

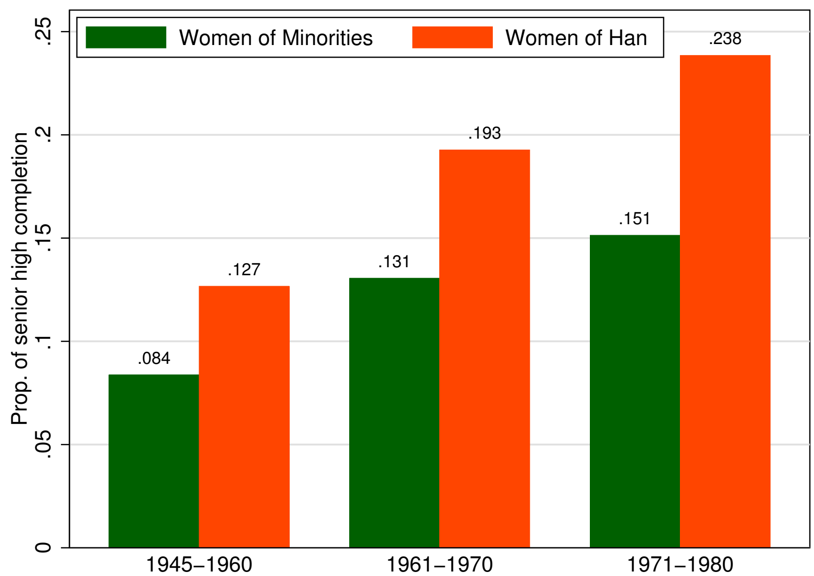
Figure 1.6a, Panel A shows that the coefficient of the benchmark analysis using the sample composed of Han women (0.021) is positive and significant but that the coefficient

Figure 1.4: *Effects of fertility fines at different ages on senior high school completion*



Notes: Sample is from Census 2000 and 2005. We conduct the regression by replacing the fertility fine at 10 to 19 years of age with a series of fertility fines at different ages (i.e., fertility fines at 6-9, 10-13, 14-17, 18-21 and 22-25 years of age) in equation 1.1. The covariates are the same as those in column 1 of Table 1.2. Both coefficients and confidence intervals are reported. The confidence intervals are calculated based on the standard errors clustered at province-YoB level.

Figure 1.5: *Women's senior high school completion over birth cohorts 1945-1980, by ethnicity*



Notes: Data source is China Censuses 2000 and 2005. In women of Han and minorities, mean rates of senior high school are calculated specifically in the birth cohort groups and are labeled above the bars.

using the sample of minorities (−0.007) is negative, smaller in magnitude, and insignificant. Panel B reports the estimates for Han men and for minority men. The coefficient for Han (0.004) and that for minorities (−0.006) are both statistically insignificant and much smaller in magnitude than that estimated for the Han women. Thus, we provide evidence that the effect of the OCP on education exists only among Han women, but not among Han men or minority women.

As discussed above in Section II, the strictness differs by urban and rural areas. Therefore, we keep only women of Han ethnicity, divide the sample by the type of *hukou*, and estimate Equation 1.1 in each separate sample. The results are reported in Panel A of Figure 1.6b. We then further divide the rural sample according to whether the province implements the "One-and-Half Children Policy" and conduct the same analysis in the two subsamples in Panel B of Figure 1.6b. Panel A of Figure 1.6b shows that the estimate for urban residents (0.031) is about two times larger than that for rural residents (0.016). This finding is consistent with the looser enforcement and more lenient birth allowance in rural areas. Panel B shows that the effect in the rural areas with a "One-and-Half Children" Policy (0.011) is also smaller than that in rural regions without it (0.020). Consistent with our expectation, all of the results show that the effect of the OCP during the teenage years is greater for regions with stricter policy implementation.

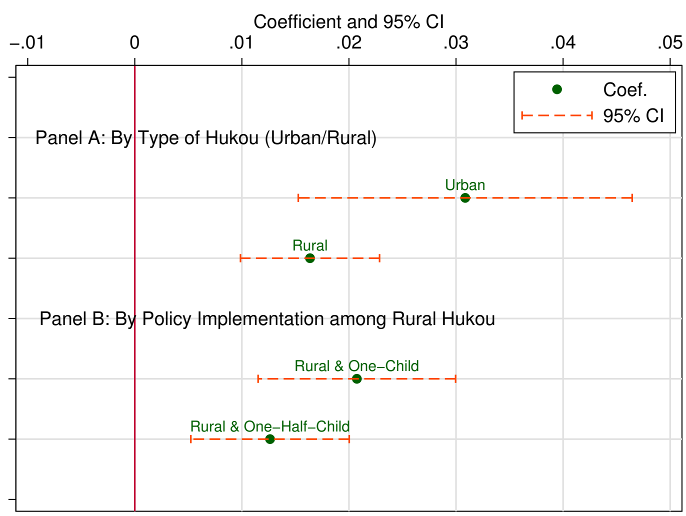
1.4.5 Placebo Test: Can Education In Earlier Cohorts Predict the Fertility Fine Later?

By adding controls such as population size and number of schools established in local regions, analyzing the effects of OCP implementation at different ages, and conducting regressions in two natural control groups - men of Han and women of minorities, above analysis provides evidence for the positive effects of OCP implementation at teenage years on the educational attainment of Chinese women. The findings suggest the potential endogeneity issue of the fertility fines caused by contemporaneous changes of other policies or economy development should not be the first-order driving factor for the estimates.

Figure 1.6: *Effects of fertility fines at 10-19 years of age on senior high school completion, by ethnicity, type of* hukou*, and policy implementation*

(a) *By gender and ethnicity*

(b) *By type of* hukou *and policy implementation*



Notes: Data source is China Censuses 2000 and 2005. The covariates are the same as those in column 1 of Table 1.2. OLS regressions are conducted in each corresponding subsample. The confidence intervals are calculated based on the standard errors clustered at province-YoB level.

Table 1.3: *Placebo Test - Prediction of Education on Fertility Policy Fines in the Future*

	(1)	(2)	(3)	(4)
Dependent variable	Policy Fine rate at age 24-26			
Sample	All Women	Women of Han	All Men	Men of Han
Senior High School Completion	0.00294	0.000895	0.00398	0.00373
(Yes = 1)	(0.00338)	(0.00333)	(0.00297)	(0.00293)
Observations	822,379	751,752	812,979	741,248
R-squared	0.865	0.871	0.862	0.867
Basic control	Yes	Yes	Yes	Yes

Notes: Sample is from Censuses 2000 and 2005. Those from individuals between 25 and 55 years of age are kept. Basic control is the same with that in Table 1.2. Robust standard errors in parentheses are clustered at province-YoB level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

However, another possible concern, which may not be addressed by these empirical results, is that the possibility that local officials change the fine rate in response to a change in women's education or factors related with education.[15] If this possibility holds, we should be able to observe some correlation between the educational attainment and the fertility fine implemented after the teenage years of these birth cohorts.

To test this, we estimate how much the fine rate would change according to the respective pre-existing educational attainments for women and men. In practice, we estimate the correlation between the local education level of the birth cohorts and the fine rate when these cohorts were 24 to 26 years of age, because almost all individuals have completed their education at 24. We regress the fine rate at 24 to 26 years of age on their senior high school completion in our sample, with all the covariates the same as in equation 1.1. By doing so, we actually examine whether officials will consider the education level of those 24 to 26 years of age as a factor in the design of the current policy fines. Table 1.3 reports the results.

[15] In particular, if the spatial and temporal variation in the OCP fines is affected by the local fertility rate, it may correlate with the preexisting pattern of education of women or the gender difference in education. If this is true, the above results may be driven by this correlation.

The first two columns show no evidence for the correlation of the preexisting educational level of women with the fertility fines when using all observations and when using the sample composed of the Han majority. The next two columns provide similar results for men. Note that all of the coefficient entries in Table 1.3 are fairly small: even if the senior high school completion rate increases from 0 to 1, the predicted fine rate increase would be smaller than 0.004 for both men and women. The similar coefficients for men and women also suggest that a change in the fine rate may not be correlated with the earlier gender difference in education. The findings provide further evidence for the exogeneity of the fertility fines used in this paper.

1.4.6 Consistent Evidence from CGSS and Heterogeneous Effects by Family Background

The CGSS provides detailed parental socioeconomic status when the respondents were 14 years of age. We exploit such information to investigate the heterogeneity of the effect across households with different family backgrounds. Because the sample size of minorities is insufficient to conduct any valid regression, we use only the Han sample in the CGSS to double-check the above results by re-estimating equation 1.1.[16]

The first two columns of Table 1.5 show the respective results for men and women. Consistent with the analysis using population census, we find that the OCP at 10-19 years of age increases education only for women. In column 3, we further control for the parental characteristics when the respondents were 14 years of age, including education (i.e., some formal education), political status (i.e., a party member), administrative rank (i.e., a rank of subsection chief or higher), and working status (i.e., whether for a state-owned or collective enterprise). These measures for parental socioeconomic status are likely to be strongly correlated with the respondents' number of siblings. The coefficient, however, only changes from 0.042 to 0.040, suggesting that the family background or the number of siblings should

[16] All of the covariates remain the same except that the CGSS does not have information about the number of siblings, which, as we will show later, is less likely to be a critical issue.

Table 1.4: *Means and Standard Deviations of the Variables in Census and CGSS for Women of Han*

Variables	(1)
Panel A: Census Data	
Married before 25 (Yes = 1)	0.77
	(0.42)
Ever had Child (Yes = 1)	0.94
	(0.23)
Work Now (Yes = 1)	0.75
	(0.44)
Having professional jobs (Yes = 1)	0.13
	(0.34)
Observations	810,994
Panel B: CGSS data	
Agreement with the statement (1-5, higher for agreement)	
"The happiest thing is to look at children growing up"	3.37
	(0.57)
"Adult children are important support for old people"	2.91
	(0.85)
"Men and women have the equal duty for housework"	3.94
	(1.00)
Agree with "Mother should stay at home and father work outside if having a pre-school child at home"(Yes =1)	0.55
Observations	6901

Notes: Standard deviations in parentheses. The sample in Panel A are the same with that in Panel A of Table 1.2. The sample in Panel B is the same with that in column 2 of Table 1.5.

not bring much bias when ignored.

Another motivation to investigate the heterogeneous effects by family background is that the strictness of enforcement varies according to family background. For example, individuals employed in state-owned or collective enterprises could risk their jobs and bonded securities and benefits with an illegal birth; party members who violate the regulation face more severe punishment, such as losing their membership, which represents higher social status in China. Thus, these households are more likely to take the fertility control policies more seriously than other households and to form a stronger expectation that their daughters would have low fertility in the future. We include the interaction term of these family background measures with the policy fines in the regression to investigate whether the effects of the OCP are greater for those with higher status, party members, and those who work for state-owned or collective enterprises. The remaining columns of Table 1.5 provide consistent evidence of this conjecture: all of the coefficients on the interactions are positive and in large magnitude.

However, the heterogeneous effects have at least one other interpretation: the parents with higher socioeconomic status may be more patient and farsighted, so they would respond more to the OCP even if they faced the same policy enforcement as other people. It is also possible that parents with higher socioeconomic status are more able to pay the tuition for their children's senior high school education. Therefore, the results provide consistent evidence that the girls, whose family or parents are in higher SES or facing stricter punishment (other than financial penalties) if violating the fertility policies, would have larger increase in educational attainment when the fertility fines at ages 10-19 increased.

1.5 Possible Mechanisms: Effects on Education and Those on Later Outcomes

1.5.1 Econometric Framework

We further explore the later outcomes in the marriage and labor markets and the subjective attitudes toward gender equality and children to examine whether the effect of the fertility policies on the latter outcomes is associated with that on education.[17] It provides some suggestive evidence for the possible explanations for fertility policy-induced higher education under the assumption that the later outcome is likely to reflect the change of the *ex ante* expectations.[18]

[17]We thank Professor Lawrence Katz for providing great help and guidance on the methods.

[18]Ideally, we could examine the role of the policy-induced expectations and subjective attitudes in the fertility policy-education nexus if we had information on them for the same periods as the policy fines. Unfortunately, to the best of our knowledge, no public data regarding individual expectations during the 1980s and 1990s in China are available. The earliest survey year is 2000, when the sample that faced a stricter OCP during their teenage years were at least 20 years of age.

In practice, we keep the Han women, divide the sample by *hukou* province, the type of *hukou*, and the survey year and then conduct the following regressions in each sample:

$$Senior_i^s = \alpha_0^s + \alpha_1^s Fine_{jb}^{10-19} + \alpha_2^s X_i^s + \delta_b^s + \varepsilon_i$$

$$Y_i^s = \gamma_0^s + \gamma_1^s Fine_{jb}^{10-19} + \gamma_2^s X_i^s + \delta_b^s + \varepsilon_i$$

where the superscript s denotes the subsample s. The term X_i^s includes the male proportion of the birth cohort of the *hukou* province and the number of siblings of individual i in sample s, and δ_b^s is a set of birth cohort group dummies (i.e., 1945 to 1960, 1961 to 1980). Note that we cannot control for the dummies of each specific birth cohort because within each *hukou* province the identification is based on the time-series variation in the fine rate within each sample s. We then test whether the effects of the fine on the outcome tend to be larger if the effect on education is larger by plotting γ_1^s against α_1^s, weighted by the representative population in each sample s. If we find a significant correlation between the coefficient γ_1^s and α_1^s, we can then conclude that the effects on education are associated with those on the later outcomes. Please note that this methodology *only* examines how the effects of fertility fines on later outcomes change according to the variation in the effects on education, and it is possible that the *average* effects of fertility fines on later outcomes could be positive or negative.[19]

1.5.2 Empirical Results

Following the method, we examine the following four *ex post* outcomes: marriage age (i.e., whether married at 25 years of age), fertility (ever having a child), labor market participation (current employment status), and occupation (holding a non-manual job).[20] Panel A in Table

[19]For example, it is possible that the *main* effects of fertility fines on women's labor supply may be negative because of bargaining power within marriage or positive because of fewer children. When people expect *relatively* more labor supply in the future (i.e., the coefficients are less negative/more positive) and thus invest in education more, we may still identify a positive correlation between the effects on education and those on the labor supply by using this methodology.

[20]A non-manual job is defined as a party leader, firm manager, administrative staff, and highly skilled workers including science researchers, engineers, and professors.

1.4 presents the means and standard deviations for the variables. In particular, 77 percent of women are married before 25 years of age, 94 percent have had a child, 75 percent are working, and 13 percent hold non-manual jobs.

Figure 1.7a and 1.7b show a negative correlation between the effects on education and those on the likelihood of being married at 25 years of age and of currently having children, respectively. Such correlations imply that women with a higher policy-induced educational level are also likely to be unmarried or currently have no child due to the policy. Similarly, Figure 1.7c and Figure 1.7d also provide consistent evidence by showing the positive correlations of the effects on education with those on working and professional job occupation. All of the correlations in the four figures are significant at a 1 percent level.

We then conduct a parallel analysis of the attitudes toward children and gender equality in the CGSS.[21] Panel B of Table 1.4 presents the means and standard deviations for the variables. Specifically, the women have almost neutral attitudes toward the importance of children; the mean values of the two measures for children are around 3 when subjects were asked whether they agree that "the happiest thing is to watch children growing up" and that "adult children are important support for old people." Regarding gender equality, they tend to agree that men should have equal duty regarding household work with women, with a mean value of almost 4 ("Agree"), but more than half of the women think that women should stay at home and that the father should work outside the home when there is a pre–school-age child at home. Figures 1.8a-1.8d show the respective correlations between the effect of the OCP on each measure for the attitudes and those on education.

Figure 1.8a and 1.8b show that the greater positive effect of the fertility policy on education is correlated with a larger negative (or smaller positive) effect on agreement with the statements that "the happiest thing is to watch children growing up" and that "adult children are important support for old people." Similarly, Figure 1.8c and 1.8d consistently show that those with more gender-equal views caused by the fertility policy are more likely

[21] Because the sample is much smaller for the CGSS, we only divide the sample by the province and type of *hukou*.

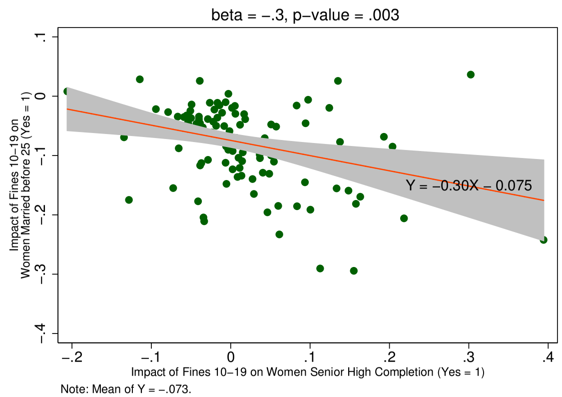
Table 1.5: *Impact of OCP on Education in CGSS, Birth Cohorts 1945 - 1980*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	Senior high or above (Yes = 1)						
Sample	Men	Women					
Mean of dependent var.	0.293	0.222					
Fertility fine rate at age 10-19	-0.0114	0.0423**	0.0396**	0.0172	0.0340	0.0360*	0.0267
	(0.0237)	(0.0199)	(0.0199)	(0.0251)	(0.0211)	(0.0197)	(0.0219)
Parents Condition or SES when respondents were 14 years old							
Either parent had some education (Yes = 1)			0.0595***	0.0841***			
			(0.00863)	(0.0119)			
Either some Education × Fine				0.0290*			
				(0.0165)			
Either parent was party member (Yes = 1)			0.0502***		0.102***		
			(0.0129)		(0.0171)		
Either party member × Fine					0.0306*		
					(0.0176)		
Either parent had higher administrative rank (Yes = 1)			0.0882***			0.169***	
			(0.0264)			(0.0338)	
Either higher Rank × Fine						0.0649**	
						(0.0282)	
Either parent in state-owned or collective enterprise (Yes = 1)			0.154***				0.180***
			(0.0151)				(0.0178)
Either in state-owned or collective enterprise × Fine							0.0293*
							(0.0172)
Observations	6,901	6,901	6,901	6,901	6,901	6,901	6,901
R-squared	0.314	0.314	0.354	0.325	0.326	0.327	0.345
Basic control	Yes	Yes	Yes	Yes	Yes	Yes	Yes

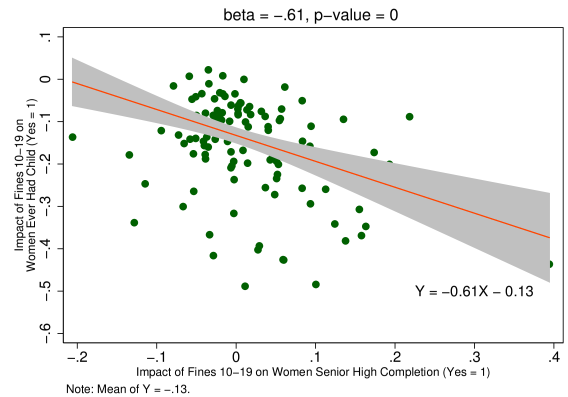
Notes: Sample is from CGSS 2010 and 2012. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1.7: *Correlation between effects of policy fine on education and those on marriage and labor market outcomes, Birth Cohorts 1945-1980 in Census*

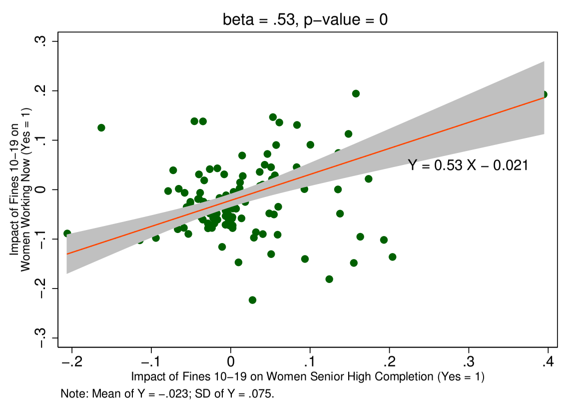
(a) *Married before 25*



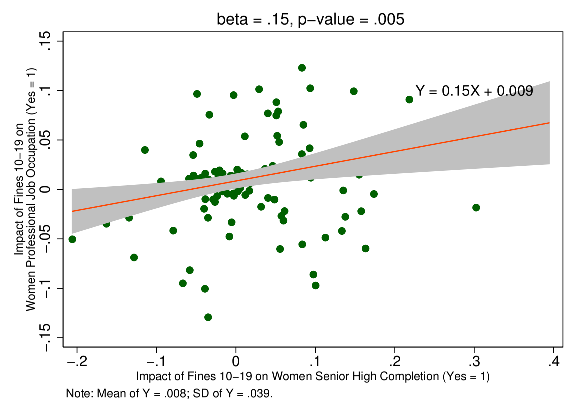
(b) *Ever had a child*



(c) *Working status*



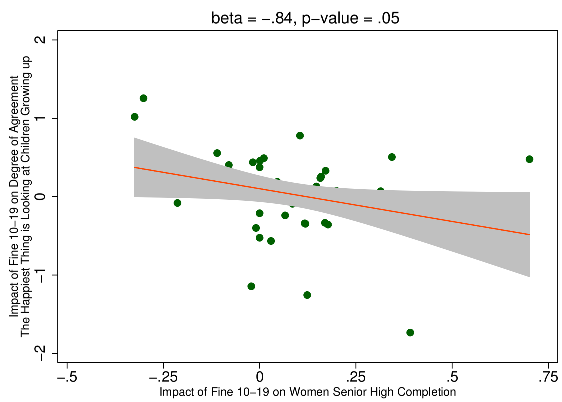
(d) *Higher-level occupation*



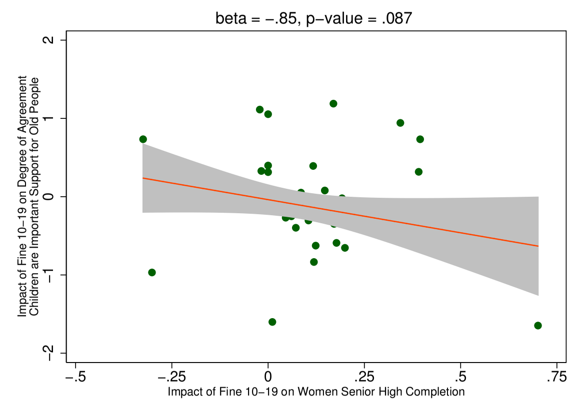
Notes: Data source is China Censuses 2000 and 2005. The four figures are plotted for different later outcomes, respectively. Each point is plotted based on the effects of fertility fines on senior high school completion (X-axis) and the effects on the later outcomes (Y-axis) in the *same* subsample. The slopes and the P-values are shown on the top of each figure.

Figure 1.8: *Correlation between effects of policy fine on education and those on attitudes regarding children and gender equality, Birth Cohorts 1945-1980 in CGSS*

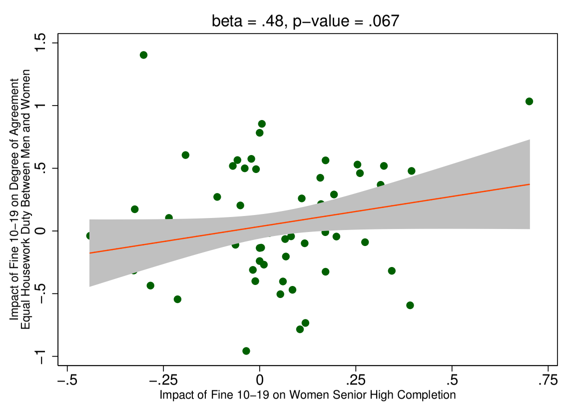
(a) *"The happiest thing is looking at children growing up"*



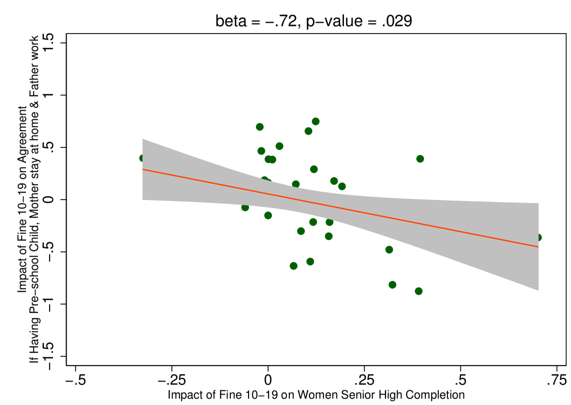
(b) *"Adult children are important support for old people"*



(c) *"Equal housework duty between men and women"*



(d) *"Mother stay home & father work if having a preschool child"*



Notes: Sample is from CGSS 2010 and 2012. The four figures are plotted for different later outcomes, respectively. Each point is plotted based on the effects of fertility fines on senior high school completion (X-axis) and the effects on the later outcomes (Y-axis) in the *same* subsample. The slopes and the P-values are shown on the top of each figure.

to have a higher policy-induced educational level. In other words, the more fertility policies increased on the gender-equality and self-reliability attitude, the more the policies increased girls' education, although the correlations are less significant (only at a 10 percent level) due to the much smaller sample size.

To sum up, we find significant associations between the effects of the fertility policies on education and those on outcomes in the labor and marriage markets as well as some subjective attitudes. These results suggest that the OCP-induced expectations for the future could induce a greater motivation for education investment. However, we should bear in mind that these associations are *not causal or determinant.* The effects of the policies on later outcomes may also be consequences of the policy-induced increased education. We look forward to studies in the future may shed some light on this.

1.6 Conclusions and Discussion

The past century witnessed a rise in educational attainment for women all over the world and the emergence of fertility policies in many countries. This paper examines the relationship of the two phenomena by investigating the effects of fertility policy implementation during the teenage years on women's education in China.

Using the temporal and regional variation in the financial penalties of OCP, we show a positive effect of the policy fines at 10 to 19 years of age on the likelihood of completing senior high school in women of Han ethnicity. The magnitude, (i.e., 1-year-household-income increase in the fine rate predicts a 2-percentage-point increase in the rate of senior high school completion), explains 30 percent of the increase in women's education in the birth cohorts from 1945 to 1980. Meanwhile, the analysis finds no effect of the fertility policies among women of minorities or men of Han, which, first provides some evidence for the exogeneity of the fertility fines used in this paper, and also suggests that the differential effects of OCP could explain 50 percent of the narrowing in gender gap in education.

To shed light on possible mechanisms, we examine the associations between the effects of fertility fines on women's education and the effects of the policies on their future outcomes.

Consistent with the previous literature (Goldin and Katz, 2002; Field and Ambrus, 2008; Miller, 2010), we find that the effects of family planning policies on women's education are significantly associated with those on the outcomes, including late marriage, employment and subjective attitudes toward children and gender equality.

Exploiting the unique policy settings in China, we provide a novel explanation to women's rise in education and economy. The evidence from this particular country sheds some light on the emerging worldwide phenomenon by establishing the relationship of the fertility policies with the accumulation of women's human capital and the narrowing of the gender gap in education. Note that the fertility policies were widely implemented in many countries in last century, especially mid- and low- income ones (Miller and Babiarz, 2014), the findings in this paper may put the research in agenda to investigate further evidence from other countries in the future. In addition, the associations of the effects of the policies on education with those on later *ex post* outcomes provide some *qualitative* and *suggestive* evidence on the the possible mechanisms. These basic findings in this paper also call for the studies to further clearly identify the possible mechanisms.

Chapter 2

One-Child Policy, Marriage Distortion and Welfare Loss[1]

"No union is more profound than marriage, for it embodies the highest ideals of love, fidelity, devotion, sacrifice, and family. In forming a marital union, two people become something greater than once they were."

–Justice Anthony Kennedy

2.1 Introduction

Marriage is an important source of happiness and plays an important role in generating and redistributing welfare among individuals (e.g., Zimmermann and Easterlin, 2006; Dupuy and Galichon, 2014). Since Becker (1973, 1974) built up the original transferable utility model for the marriage market over 40 years ago, an established strand of literature has used and applied this model and its wide-ranging implications (Rao, 1993; Edlund, 2000; Chiappori *et al.*, 2002; Botticini and Siow, 2003; Bitler *et al.*, 2004). Choo and Siow (2006) further developed the model to derive a reduced-form testable formula explicitly linking unobserved marriage gains to the observed marriage outcomes. Then they used this formula

[1]Co-authored Yi Zhou.

to estimate the loss of marriage gains due to the national legalization of abortion in the United States in 1973.

However, there is little empirical evidence for the cornerstone of the transferable utility model, i.e., that individual marriage behavior and market equilibrium are shaped by potential marriage gains. The major difficulty is the rareness of exogenous variation in marriage gains since there is almost no such event or policy that assigns different gains to various types of marriages.[2]

This paper first sheds light on this by using the micro-level data from the country with the largest marriage market and estimating the effects of the One-Child Policy (OCP) on marriage equilibrium outcomes.[3] Different from the welfare programs studied in the previous literature (e.g., Bitler *et al.*, 2004), compulsory fertility restrictions would more exogenously reduce potential marriage gains, because children are natural and important fruits of marriages and can be sources of joy and future supporters.[4] As a result, the fertility policy would distort individual incentives in the marriage market and alter the marriage equilibrium outcomes. Some unique features of the OCP implementation make it a natural setting in which to investigate questions about the possible effects of distortions caused by the disincentives to have children on marriage. Above all, the OCP *directly* and *compulsorily* assigned limited birth quotas to couples. These quotas were strictly implemented by the Population and Family Planning Commissions (PFPC) at every level of government. In addition, the OCP had a large ethnic, spatial, and temporal variation in implementation. First, different birth quotas were assigned to both-Han (H-H), both-minority (M-M) and Han-minority (H-M) couples, according to local policies. In almost all provinces, H-H

---

[2]There are some exceptions. For example, Bitler *et al.* (2004) used vital statistics data on marriages and divorces in the United States during 1989-2006 to examine the role of welfare reform (state waivers and implementation of Temporary Assistant to Needy Families) and other state-level variables on flows into and out of marriage. However, the effects of welfare reform on marriage and divorce are not clearly predicted by the theory and using vital statistic data yields inconsistency in the coefficients on some of key variables in the study.

[3]The OCP was initiated in the late 1970s and has restricted the fertility of hundreds of millions of couples for about 35 years. On Oct 29th 2015, China's government announced it would abandon the one-child policy and all couples would be allowed to have two children. It is almost the largest social experiment in human history.

[4]In China, there are very few women giving birth prior to marriage because of culture and social norm.

couples are strictly constrained to only one (or conditionally two) births, while M-M couples were legally permitted to have more births or were not even subject to the OCP (Baochang *et al.*, 2007). About half of the provinces extended the coverage of this exemption to H-M couples (referred to as preferential-policy regions, hereafter), but the others did not (non-preferential-policy regions). Second, different levels of financial penalties were imposed for an illegal birth across provinces and across years. These OCP penalties ranged from one to five times a local household's yearly income, and were applied to any illegal birth above the quota in the corresponding province.[5]

To investigate how expected marriage gains affected marriage outcomes, we derive three intuitive and testable hypotheses after incorporating the OCP in the model of Choo and Siow (2006). The first hypothesis is that the OCP would increase the unmarried rate due to lower expected gains from marriage, especially for Han people;[6] second, the OCP would increase the H-M marriage rate, particularly in the preferential-policy regions; and third, in the preferential-policy regions, the OCP would increase the utility transfer from a Han spouse to his (or her) minority spouse within H-M couples.

Our empirical results provide sound evidence for the hypotheses above. Using the regional and temporal variations in the fertility penalties combined with census data in China, we find that an increase in the fertility penalty at age 18-25 by one yearly local household income increases the unmarried rate by 1.7 percentage points among Han people (39 percent of the mean).[7] Moreover, in the preferential-policy regions, the same increase in fertility penalty also increases H-M marriage rate by 0.6 percentage points for Han people (20 percent of the mean) and 2.1 percentage points for minorities (15 percent of the mean), respectively.[8] Finally, among H-M couples in the preferential-policy regions, higher fertility

[5]Local governments are responsible for collecting the financial penalties, and a number of administrative penalties such as confiscation of property and excluding children born outside from the *hukou* system are employed to assist the OCP penalty collection.

[6]Note that being unmarried does not mean staying single forever here. Many merely delayed marriages. Hence, increased unmarried rate can be also understood as not married in early ages.

[7]There are 4.4 percent of Han people in the sample having a status as single.

[8]The H-M marriage rates are 3.0 percent for Han and 14 percent for minorities in the preferential-policy

penalty is associated with higher education of spouses for the minorities but not for the Han people.[9]

Investigation using several control groups provides further evidence for the hypotheses and the exogeneity of for the variation that we employ. In contrast to the above significant impacts, fertility penalty has imposed a much smaller and insignificant effect on the unmarried rate of the minorities, as well as on the H-M marriages in the non-preferential-policy regions. Also, among the H-M couples in non-preferential-policy regions, the fertility penalty is not correlated with the education of spouse for either Han people or for minorities.

As the model suggests, the OCP affects marriage outcomes through changing the expectations in the number of births. Although the expectations cannot be observed, we still provide some evidence by showing that the policy-induced H-M marriages tend to result in more births *ex post*. More specifically, we show that, in the presence of the preferential policy, the regions with larger positive effects of the OCP on the H-M marriage rate tend to have weaker negative effects of the OCP on fertility; however, the correlation is much weaker in the non-preferential-policy regions. These findings consistently suggest that some H-M marriages were motivated by the reduction in marriage gains due to the OCP in terms of the restriction on the number of permitted children.

The OCP-induced marriage behaviors are fairly consistent with the predictions originating from the transferable utility model of Choo and Siow (2006) and we conclude that the OCP has caused a significant distortion in marriage market because of the policy-induced expectation in number of children to give birth. Therefore, it is natural to ask how much social welfare loss is caused by the distortion since policy-induced behavior distortion is generally associated with a social welfare loss (Chetty, 2009b; Hendren, 2013). Following the *sufficient statistics* approach in (Chetty, 2008, 2009b,a), we derive a formula for the social

---

regions. In the econometric framework, besides the local minority proportion and sex ratios in the birth cohort of local province, we also controlled for the fixed effects for the ethnicities, type of *hukou*, provinces, cohorts, and calendar years, and the province-specific linear trends in birth cohorts throughout the whole analysis.

[9] Previous literature used education as a pre-marital investment (Chiappori *et al.*, 2009; Lafortune, 2013). We follow this literature here and assume that higher education indicates more transfers to spouse in marriages.

welfare loss caused by the OCP in fertility and marriage market as a whole. This formula only depends on the estimated reduced-form elasticities. More specifically, the welfare deadweight loss (DWL) is composed of two parts: the first originates from policy-induced declined fertility ("mechanical" effects); while the second part pertains to the marriage distortion analyzed above ("distortion" effects). The approach to welfare analysis is different from the traditional approach which structurally estimates a model's primitives and then numerically simulates the effects of a policy. Compared to the traditional approach, this approach is less model-dependent and more empirically credible (Chetty, 2008, 2009b). To the best of our knowledge, this is the first study to estimate the welfare loss caused by the OCP and also the first endeavor to apply the *sufficient statistics* approach to the marriage market.

By applying the reduced-form estimates to the model, we show that the total social welfare loss is about 4.9 percent of total yearly household income, of which 0.85 percent originates from the marriage distortion. Therefore, these estimates suggest that, not accounting for the "distortion" effect would substantially underestimate the total social welfare loss by 17 percent. As marriages are almost prerequisites for children in China, marriage choices are distorted by fertility policies and thus a welfare loss based only on the fertility reduction of married couples is not the whole story. These results highlight that the unintended behavioral responses following from the OCP in terms of marriage distortion is a significant component of the welfare loss. These findings also provide some new insights into public economics, namely that the relationship between commodities need to be considered when estimating the welfare loss of taxation.[10]

This paper is organized as follows: section II introduces the context of this study, especially the background of the OCP. Section III develops a theoretical framework for the empirical predictions and the welfare implications. Section IV presents the empirical

[10] In this study, children are considered to be downstream "goods" of marriages. Taxing children (OCP) has brought significant distortions in marriage behaviors because the expected potential marriages would be eroded for most people. Our results are similar to the findings in Busse *et al.* (2013) who found that gasoline prices have significant impacts on the prices and quantities of sales in the new and used car markets.

strategy and the marriage distortion caused by the OCP. Section V calculates the welfare loss caused by the OCP based on the estimates in the previous sections, and section VI concludes.

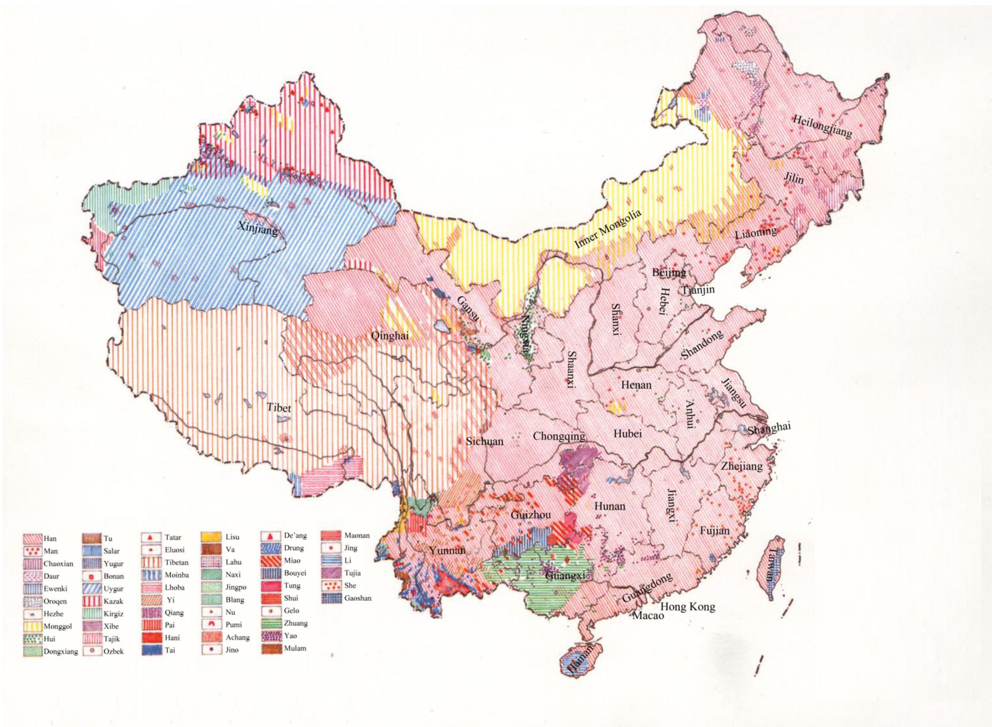
2.2 Context: Ethnicities and One-Child Policy in China

2.2.1 Ethnic Minorities in China

China is a populous country with controversial ethnic issues (Sautman, 1998; Kaup, 2000; Ma, 2007). Ma (2007) listed ten of China's ethnic issues that are worthy of academic attention, and the first one among them is ethnic identification and nationalism. China officially has 56 ethnicities. Soon after the founding of the People's Republic of China in 1949, the central government initiated a monumental project of ethnic identification. In the 1953 population census, more than 400 groups applied for national minority status (Fei, 1979). With guidance from a few Western-educated anthropologists, hundreds of research teams were sent to conduct fieldwork and collect information about the history, language and customs of each group. The main work of ethnic identification was finished in 1957, but follow-up revisions continued until the 1970s. The most recent revision was the recognition of the Jino people in 1979, right before the implementation of the OCP. Based on cultural characteristics and the will of the groups concerned, most of these self-nominated groups were recognized as minority people, and they were officially reclassified into 55 groups. According to Regulations on Household Registration of People's Republic of China, every newborn's ethnicity should be registered in the *hukou* system in the first month after birth. Ethnic identity is mainly determined by parents' ethnicities. The children of intermarried families are permitted to follow either the father's or mother's ethnicity (Jia and Persson, 2015). Ethnic identity is strictly controlled, and thus it is difficult for individuals to make a fake claim.

According to the 2010 census, the Han ethnicity make up 91 percent of the population, while all of the other 55 ethnic groups account for the remainder. The largest minority group

Figure 2.1: *Geographic Ethnicity Distribution in China*



Notes: This map is from the book "A Mosaic of Peoples: Life Among China's Ethnic Minorities" (1992) by China Nationality Art Photograph Publishing House.

currently in China is Zhuang, with a population of 16.9 million in 2010.The smallest minority group, the Keba, has only 3682 members. Figure 2.1 shows the geographic distribution of all the 56 ethnic groups. As shown in the map, most ethnic minority groups live in regions on the western or northeastern boarder. The current geographic pattern of ethnic distribution is mainly caused by the migration history of the Han Chinese (Poston Jr and Shu, 1987).

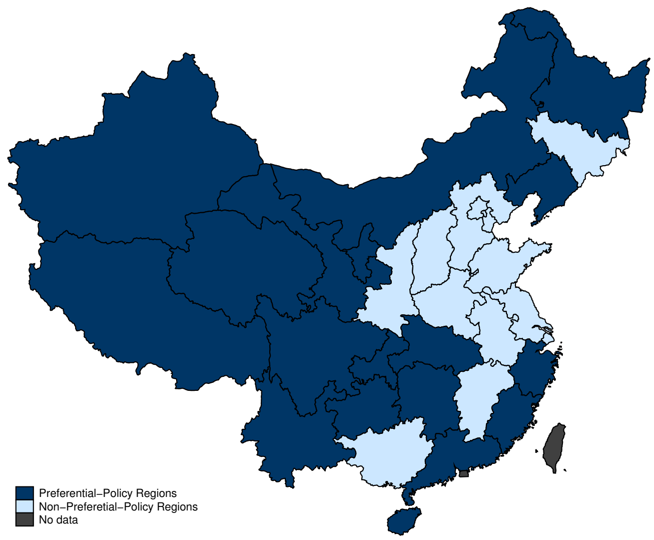
2.2.2 One-Child Policy

China's OCP was first announced in 1978, and it appeared in the amended Constitution in 1982. Legal measures such as monetary penalties and subsidies were employed for the effective enforcement of OCP since 1979 (Banister, 1991). In early 1984, the Communist Party Central Committee issued Central Document 7 as a guideline for local implementation of fertility policies (Greenhalgh, 1986). Because of the "practical difficulties" experienced in earlier years, one important feature of Document 7 was greater flexibility in local practices. As a slogan at that time said, "Open a small hole to close up a big one." The central government believed that some small compromises would make the whole policy more acceptable.

The central government authorized provincial governments to design specific regulations according to their local conditions. Indeed, both the effectiveness of the implementation of the OCP and inter-ethnic harmony were on the list of evaluation criteria for local officials. Therefore, preferential terms were exclusively granted to M-M or H-M couples (Baochang *et al.*, 2007). Han residents living in urban areas were mostly allowed to have only one child, but those living in rural areas could have one or two. M-M couples were legally permitted to have more births or even were not subject to the OCP. About half of the provincial governments extended the coverage of this exemption to H-M couples but the others did not. For example, the Population and Family Planning Statute of Qinghai states, "Families can have one more births, if one or both sides of the couple are from minority groups." We collected data regarding the exemption terms for H-M couples in every province from the website of the National Health and Family Planning Commission of China. According to the historical policies that that we can find, there was no temporal variation in the existence of preferential terms for H-M couples in various provinces.[11] Figure 2.2 plots the provinces with exemption terms geographically.

[11]The data source is the website of the National Health and Family Planning Commission: http://www.nhfpc.gov.cn/zhuzhan/dftl/lists.shtml. However, the exemption terms may have some variations within provinces in rural/urban regions and in different ethnicities. For example, Zhuang ethnicity may not have a preferential policy in certain regions, and some provinces also specify that the preferential policy may only apply for rural areas but not for urban areas. Our analysis also accounts for this variation.

Figure 2.2: *Preferential-Policy Regions v.s. Non-preferential-policy regions*



Notes: Data source for the preferential-policy regions is from the website of National Health and Family Planning Commission of the people's Republic of China. The website is http://www.nhfpc.gov.cn/zhuzhan/dftl/lists.shtml (Chinese Website accessed in November 2015).

We additionally use the average fertility financial penalties for one unauthorized birth on the province-year panel from 1979 to 2000. The OCP penalties or fertility fines, also known as "social maintenance fees" in China, were formulated in multiples of yearly household income, which is consistent with its wide use in previous literature (Ebenstein, 2010; Wei and Zhang, 2011; Huang *et al.*, Forthcoming). The variation has been used in Chapter 1 in the thesis. We use the variation to identify the effects of the OCP on marriage outcomes in following analysis.

2.3 The Model and Implications

2.3.1 Marriage Distortion under the One-Child Policy

We follow the framework of Choo and Siow (2006) to analyze the impact of the OCP on marriage equilibrium outcomes. People are divided into two types: Han (H) or minority

(M). Under the circumstance of the OCP, we suppose there are two periods: in the first period, people decide whether to marry and to whom they marry; in the second period, married people decide how many children to have. However, people are able to anticipate the number of children to have according to the potential types of marriages and local fertility policies, and thus behave correspondingly in the marriage market.

Fertility choice under the OCP We solve the problem backwardly. A certain couple (i, j) choose the number of children to give birth to, n_{ij}, in order to maximize the household utility under the fertility policy depicted by $(\overline{n}_{ij}, f)$, where $\overline{n}_{ij}$ is the birth quota assigned to the couple (i, j), and f is the fertility penalty for an additional illegal birth. For simplicity, the households solve the problem as follow:

$$\max_{n_{ij}} u(n_{ij}) + y_{ij} - n_{ij}C - \delta_{n_{ij} \geq \overline{n}_{ij}}(n_{ij} - \bar{n}_{ij})f \tag{2.1}$$

where $u(.)$ is the utility from the number of children given birth to and is uniformly applied to all couples, with $u' > 0$ and $u'' < 0$; y_{ij} is the exogenously given household income and C is the fixed cost of raising up a child. $\delta_{n_{ij} \geq \overline{n}_{ij}}$ is an indicator function which equals to 1 if $n_{ij} \geq \overline{n}_{ij}$, and 0 for otherwise. We assume that utility function is quasi-linear such that the utility can be interpreted in monetary units directly.[12] For simplicity, we also assume the couple can choose any positive number of children ($n_{ij} \in \mathbb{R}^+$) , and define u_{ij} as the maximized household utility for couple (i, j). Then we have $\dfrac{\partial u_{ij}}{\partial f} = -\delta_{n^*_{ij} \geq \overline{n}_{ij}}(n^*_{ij} - \overline{n}_{ij}) \leq 0$, implying that fertility penalties would reduce the utility when the birth quota is binding. Therefore, the fertility penalty would have larger negative effects on the household utility gained from the number of children when the birth quotas are fewer. According to the OCP implementation, the negative marginal effects of fertility penalty on the household utility would be larger for H-M couples in non-preferential-policies regions and for all H-H couples, and smaller or even equal to zero for H-M couples in preferential-policy regions

[12]The quasi-linear utility function simplifies the welfare implication in previous literature. See Chetty (2009b) for examples.

and for all M-M couples.

Marriage market distortion Following the setting in Choo and Siow (2006), for a type i man to marry a type j woman, he must transfer an amount of income τ_{ij} to her. The marriage market clears when, given equilibrium transfers τ_{ij}, the demand by type i men for type j spouses is equal to the supply of type j women for type i men for all i, j. We assume the numbers of men and women of Han people are both $\overline{H}$ and those of minority people are $\overline{M}$, with $\overline{H} > \overline{M}$. Every individual considers matching with a member of the opposite gender or staying single. Let the utility of a type i man g who marries a type j woman be $V_{ijg} = \tilde{\alpha}_{ij} - \tau_{ij} + \epsilon_{ijg}$, where $\tilde{\alpha}_{ij}$ denotes the gross marriage gains to the man i in potential marriage (i, j). For simplicity, suppose the above utility gained from the number of children are divided between men and women equally. Therefore, $\tilde{\alpha}_{ij} = \frac{1}{2}u_{ij} + \tilde{a}_{ij}$, where $\frac{1}{2}u_{ij}$ denotes the utility that a type i man expects to gain from the quantity of children if he marries a type j woman. $\tilde{a}_{ij}$ represents the systematic gross return to a type i man married to a type$j$ woman other than that from the quantity of children. The payoff to a type i man g from remaining unmarried is denoted by $V_{i0g} = \tilde{\alpha}_{i0} + \epsilon_{i0g} = \tilde{a}_{i0} + \epsilon_{i0g}$.

The women's problem is parallel. We let the utility of a type j woman k who marries a type i man be $W_{ijk} = \tilde{\gamma}_{ij} + \tau_{ij} + \varepsilon_{ijk}$, in which $\tilde{\gamma}_{ij} = \frac{1}{2}u_{ij} + \tilde{b}_{ij}$, where $\frac{1}{2}u_{ij}$ denotes the utility gained from the expected number of children to women j, and $\tilde{b}_{ij}$ represents the systematic gross return to a type j woman married to a type i man other than that from the quantity of children. The payoff of remaining unmarried is given by $W_{0jk} = \tilde{\gamma}_{0j} + \varepsilon_{0jk}$.

Following the assumption in Choo and Siow (2006), we also assume that ϵ_{ijg}, ϵ_{i0g}, ε_{ijk}, and ε_{0jk} are independently and identically distributed random variables with a type I extreme-value distribution. A man g of type i will choose according to $V_{ig} = max_j\{V_{i0g}, V_{iHg}, V_{iMg}\}$. A women k of type j will choose according to $W_{jk} = max_i\{W_{0jk}, W_{Hjk}, W_{Mjk}\}$. Defining $\alpha_{ij} = \tilde{\alpha}_{ij} - \tilde{\alpha}_{i0}$, $\gamma_{ij} = \tilde{\gamma}_{ij} - \tilde{\gamma}_{i0}$, and μ_{ij} as the number of (i, j) marriages, we consider the following symmetric equilibrium for men and women (i.e., $\mu_{ij} = \mu_{ji}$) : for i,$j \in \{H, M\}$,

$$\tau_{ij} = \frac{ln\mu_{i0} - ln\mu_{0j} + \alpha_{ij} - \gamma_{ij}}{2}; \; ln\mu_{ij} = \frac{ln\mu_{i0} + ln\mu_{0j}}{2} + \frac{\alpha_{ij} + \gamma_{ij}}{2}$$

$$\text{with } \mu_{H0} + \mu_{HH} + \mu_{HM} = \overline{H}, \, \mu_{M0} + \mu_{MH} + \mu_{MM} = \overline{M}.$$

For type i individuals ($i \in \{H, M\}$), we denote the married rate as r^i_m, and the H-M marriage rate (conditional on married) as r^i_{HM}. Assuming that the fertility penalty does not affect the utility of being single or the systematic gross return other than that from the number of children, we have the following empirically examinable implications (See Appendix B for detailed mathematic proof).

Prediction 1: The fertility penalties increase the unmarried rate of Han people, especially in non-preferential-policy regions.

Because birth-quota restrictions and fertility penalties have reduced the potential gains from marriage, more Han people would choose to delay their marriages or stay single. In this paper, we do not differentiate these two, and only investigate whether the respondent was married or not at the time of the survey. In addition, because minority people are generally not subject to the restrictions, we expect that such effects would be significant only for Han people. Since the Han people in preferential-policy regions may marry minorities to escape from the fertility restrictions, the effects should be weaker in these regions.

Prediction 2: The fertility penalties increase the H-M marriage rate for both Han and minorities only in preferential-policy regions.

In the presence of preferential policies, because H-M marriage is a way to bypass the fertility restrictions legally for Han people, a higher level of penalty would induce greater incentives for Hans to marry minorities. In contrast, people would not have policy-induced incentives to participate in such marriages when the preferential policy is absent because there is no additional birth quota for the H-M couples.[13]

Prediction 3: The fertility penalties increase the marriage transfer from Han to minorities only in preferential-policy regions.

In the preferential-policy regions, the price of a minority spouse in the marriage market would increase when the fertility penalty becomes heavier, because a minority spouse is

[13]We investigate interethnic marriage not only because interethnic marriage is an important marriage outcome impacted by the OCP but also because interethnic marriage has been widely used as an indicator of social boundaries between two ethnic groups in sociological and economic studies (Kalmijn, 1991; Fryer, 2007).

associated with additional birth quotas in these regions for Han people. Therefore, the transfers from the Han spouse to the spouse in a H-M marriage would increase when fertility penalty increase. However, since the transfers cannot be directly observed in the data, this paper investigates the association of the minority ethnicity with the education level of the spouse in H-M couples to test this hypothesis.[14]

2.3.2 Welfare Implications

The model above yields the probability that a utility-maximizing man of type i marries a woman of type j is $P_{ij} = \frac{exp(\tilde{\alpha}_{ij})}{\sum_j exp(\tilde{\alpha}_{ij})}$ and the expected utility of a man of type i is $S_i(\tau) = ln(\sum_j exp(\tilde{\alpha}_{ij}))$. Since the utility is in monetary units under the quasi-linear utility setting, social surplus is the summation of the expected utilities of men and women, and the fertility penalties collected by the government:

$$\Pi = \sum_i \overline{m}_i ln(\sum_j exp(\tilde{\alpha}_{ij})) + \sum_j \overline{n}_j ln(\sum_i exp(\tilde{\gamma}_{ij})) + \sum_{i,j \neq 0} \mu_{ij} c_{ij} f \tag{2.2}$$

where $\overline{m}_i$ and $\overline{n}_j$ denote the number of men of type i and that of women of type j, respectively. We denote $c_{ij} = \delta_{n_{ij} \geq \overline{n}_{ij}}(n_{ij} - \overline{n}_{ij})$ as the number of illegally-born children for couple the (i, j), divide the above equation by the total population of men (or women) (i.e., $\overline{H} + \overline{M}$, the number of the households), and then take the derivatives with respect to the penalty fine rate (See Appendix B for mathematic proofs). Then, we have

$$\frac{d\pi}{df} = \sum_{i \in \{H,M\}} P_i \Big(\sum_{j \in \{H,M\}} r_m^i r_{ij}^i c_{ij} (e_m^i + e_{ij}^i + e_{ij}^c) \Big) \tag{2.3}$$

where π is the surplus *per* household, P_i is the proportion of type i people in the population, r_m^i is the married rate for type i people, and r_{ij}^i is the proportion of married type i people involved in type $i - j$ marriages with $i, j \in \{H, M\}$. And e_m^i, e_{ij}^i and e_{ij}^c are the elasticities of r_m^i, r_{ij}^i and c_{ij} with respect to the penalties f, respectively. The equation 2.3 indicates that the welfare loss depends *only* on the basic statistics and behavior responses to the penalties,

[14] There is also a strand of economic literature studying marriage transfers in terms of dowries (Botticini and Siow, 2003; Anderson and Bidner, 2015), bride exchange(Jacoby and Mansuri, 2010).

suggesting that the corresponding elasticities (i.e. e^i_m,e^i_{ij} and e^c_{ij}) are *sufficient statistics* to estimate the social welfare loss (Chetty, 2008, 2009a,b; Hendren, 2013). More importantly, these behavioral responses can be derived directly from OLS estimations.[15]

In the equation 2.3, the welfare loss for the ethnicity i is $\sum_{j\in\{H,M\}} r^i_m r^i_{ij} c_{ij}(e^i_m + e^i_{ij} + e^c_{ij})$, and the social welfare loss is the population weighted mean of it. Within the parentheses, the first term captures the part whether individuals choose to marry or not: it is expected to be negative due to the lower expected gains from marriage. The second term captures potential welfare gain or loss from the policy-induced changes in marriage matching for different types of people: it may be positive or negative depending on the assignment of expected marriage gains. Therefore, the first two terms capture the welfare loss caused by the distortion in the marriage market and we term it "distortion" effects.

The final term originates from the fertility restrictions by the penalties and we name it "mechanical" effect. It is expected to be negative. Had we followed the traditional way to consider the tax incidence on the "taxed goods", the estimated total welfare loss of the OCP would only account for the part caused by the reduction in fertility, which is captured by this final term. However, the "distortion" effects in the marriage market would be ignored, and it is thus an empirical question as to how much the welfare loss caused by marriage distortion contributes to the total.

It should be noted that the model only looks into the effects of OCP on marriage market and fertility in a partial equilibrium framework. It does not take into account the potential externalities of number of children or other dimensions, including the impacts of the fertility policies on the status of women and the quality of children, though some of these factors are investigated in previous literature (Miller, 2010; Rosenzweig and Zhang, 2009). The next few sections provides empirical evidence of the effects of OCP, and estimate the social welfare loss as well as the proportion caused by the distortion effects.

[15]We thank Professor Raj Chetty and Professor Nathan Hendren for their encouragement and guidance for this part. Any errors are ours.

2.4 Data

The main data used in this study are the 2000 Population Census and the 2005 One Percent Population Survey (referred as Census 2000 and 2005, thereafter). Both of the data sets contain gender, education level, year and month of birth, region of residence, type of *hukou* (urban/rural), *hukou* province, ethnicity, marital status, and number of children. For each household, the relationship of each member with the household head is also available, which may include spouse, offspring, siblings, parents, etc. We use this information to identify couples in the households. Sampling weights are applied throughout the whole analysis.

We restrict our sample to those aged between 25 and 55. We keep those aged 25 or above because the outcomes in marriage market would be close to equilibrium and the late marriage age in China usually refers to age 25. We also drop those aged 55 or above because seniors may suffer from mortality selection and then be widowed after this age. The cohorts in the sample are those born during 1945-1980. Since the OCP started in 1979, almost all people in the sample were born before the OCP, and there are many cohorts with their marriage market affected and not affected by the OCP. Our results are robust with different sample restrictions in age.

In the questionnaire, marital status is categorized into five groups: 1 for unmarried, 2 for those in a first marriage, 3 for remarried, 4 for divorced, and 5 for widowed. For accuracy and simplicity, we only keep the sample who were single or in their first marriages (96 percent of the original sample). Based on the answers to the marital status and the ethnicity of spouse, we examine two outcomes in the marriage market: unmarried status and H-M marriage.[16] Because we analyze the sample by Hans and minorities, these two outcomes fully capture whether the respondents were married and to whom they were married. To make the empirical results easier to interpret and to derive the needed parameters in equation 2.3 for estimating the welfare loss, we use different samples for these two outcomes.

[16]Note that being unmarried does not mean staying single forever here. Unmarried rate can be also understood as being not married at certain ages.

First, we use the full sample derived above to study the impact of the OCP on whether the person is married or not. Then, we keep the married ones with information on spouse (88 percent of the sample) to study the impact of the OCP on whether people married others of their own ethnicity (Han/minorities), or of different ethnicities. When information on the spouses is missing, it is mainly because the spouses were not currently living in the household or they refused to answer.[17]

Table 4.1 shows the mean values and standard deviations for the main variables used in this study, by Hans and minorities. The first three columns are for the full sample, and the next three are for married people. Panel A presents the results for marriage outcomes. According to the results, 4.6 percent of people (i.e., 4.4 percent of Han and 6.6 percent of minorities) were unmarried at the time of the survey. Among married people, 2.9 percent were involved in H-M marriages. Because the number of Han people and the number of minority people involved in H-M marriages are the same but the population size of the minorities is much smaller (8.4 percent of the population), the H-M marriage rate is 1.6 percent for Hans and 17.4 percent for minorities. Given that the H-M marriage rate would be 6.5 percent if Han and minorities had married randomly, H-M marriages are still relatively rare compared with homogamy. The reasons could be 1) that people prefer homogamy partially because of the shared culture and language, and lower communication costs; and 2) that the interaction across different ethnicities is relatively less than that within the same ethnicity because the minorities tend to inhabit certain geographical regions.

Panel B presents descriptive statistics of the demographic and socioeconomic status variables. On average, minorities are of lower socio-economic status than Han people. The proportion of Hans living in urban regions (43 percent) is much higher than that of minorities (26 percent). Gender composition is almost balanced and the average age is about 39 across all samples.

[17] The use of all the married couples also gives consistent results. In doing so, we first assume that all the married ones with missing information of spouse are homogamy because most of the marriages are within ethnicity. Then, instead of assuming they are homogamy, we assume they are in another group, namely "missing" group, and repeat our analysis. Both of the two yield consistent results.

Table 2.1: *Summary Statistics*

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full sample			Married sample		
	Full	Han	Minority	Full	Han	Minority
Panel A: Marriage outcomes						
Unmarried (%)	4.62	4.44	6.57			
	(21.00)	(20.59)	(24.78)			
H-M marriage (%)				2.94	1.61	17.38
				(16.88)	(12.58)	(37.89)
H-H marriage (%)				90.10	98.39	
				(29.86)	(12.58)	
M-M marriage (%)				6.96		82.62
				(25.45)		(37.89)
Panel B: Demographics and Education levels						
Minority (%)	8.64			8.42		
	(28.10)			(27.78)		
Male (Yes = 1)	0.50	0.50	0.51	0.49	0.49	0.49
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Urban (Yes = 1)	0.41	0.43	0.26	0.41	0.42	0.26
	(0.49)	(0.49)	(0.44)	(0.49)	(0.49)	(0.44)
Age	39.40	39.49	38.42	39.82	39.91	38.89
	(8.21)	(8.21)	(8.21)	(8.03)	(8.02)	(8.04)
Observations	5,677,311	5,223,157	454,154	4,692,977	4,330,059	362,918

Notes: Data source is Census 2000 and 2005. Sampling weights applied. Standard deviations are in parentheses.

2.5 Econometric Framework and Empirical Results

2.5.1 Marriage outcomes responding to the OCP fertility penalties

We start the analysis by applying an "event study" to investigate how marriage outcomes respond to the variations in the fertility fines. For each group based on the type of *hukou* (urban/rural), ethnicity (Han/minorities), and survey year (2000/2005), we first calculate the changes in the fertility penalties at ages 18-25 and the changes in marriage outcomes (i.e., unmarried rate and H-M marriage rate) in two consecutive birth cohorts, within the same *hukou* province. We use the fertility penalties at age 18 to 25 because this is when most individuals prepare for marriage and seek spouses.[18]

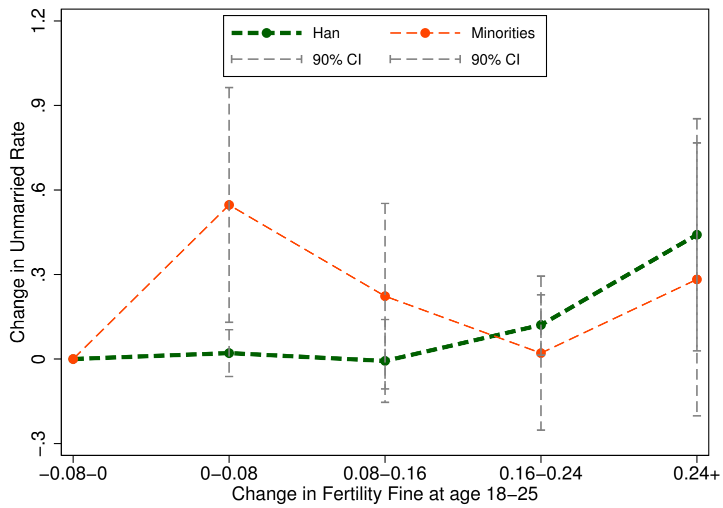
We then plot the changes in the marriage outcomes against those in the fertility penalties, weighted by the corresponding population size. Figures 2.3a and 2.3b show the results. For the outcome of unmarried status, we divide the sample into Hans and minorities because the OCP mainly restricted the fertility of Han people rather than minorities. For the outcome of H-M marriage, we divide the sample into preferential-policy regions and non-preferential-policy regions because the positive correlation is supposed to exist only in preferential-policy regions. The change in the penalty rate is divided into five categories. A higher value means a stricter policy at age 18-25 compared to the prior birth cohort. The positive slopes for the thick blue lines in both figures indicate that stricter fertility policy increases the unmarried rate as well as the H-M marriage rate in corresponding treated groups. In contrast, the increase in the penalty rate appears to be uncorrelated with unmarried rate for the minorities, and H-M marriage rate for the non-preferential-policy regions.

---

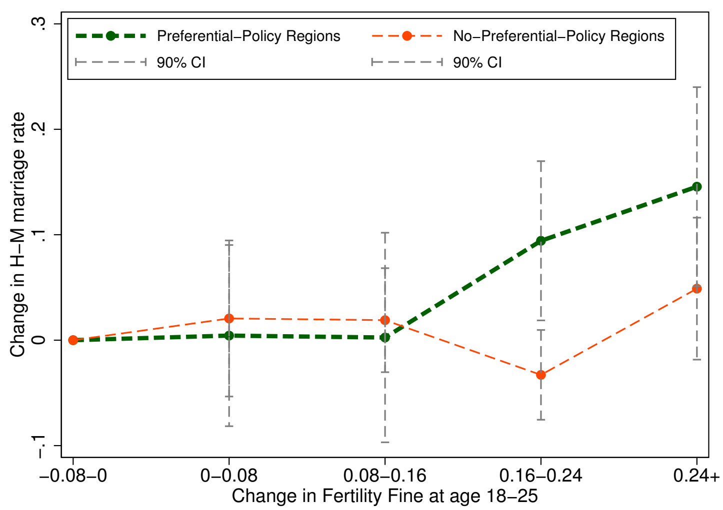
[18]In the sample, about 80 percent of marriages are formed during this age period. We also tried the penalties at other age periods and the results are consistent. We also trim the sample for this event study to those born later than 1950, because those born earlier would not have been subject to variations in the fine rate when they were 18-25.

Figure 2.3: *Marriage Outcomes Changed according to the Changes of the OCP Penalties at age 18-25*

(a) Unmarried status, by Han and Minorities



(b) H-M marriage, by Preferential-Policy or No-Preferential Policy Regions



Notes: The data source is Census 2000 and 2005. X-axis is the categories of changes in the OCP penalties at age 18-25 in two consecutive birth cohorts and the Y-axis is the corresponding changes in unmarried rate (for Figure a) and H-M marriage rate (for Figure b) in each category. Standard errors are clustered at province level and 90% CIs are reported. The estimation is weighed by the population size of each birth cohort. The treated groups for unmarried rate and H-M marriage are Han ethnicity people and the people in preferential-policy regions, respectively.

2.5.2 Econometric framework

To estimate the impact of the OCP on the marriage outcomes, we conduct the following regressions:

$$Y_{ijbt} = \beta_0 + \beta_1 Fine_{jb}^{18-25} + X_{ijbt} + D_{ijbt} + \gamma_j Prov_j \times YoB_b + \epsilon_i$$

where the dependent variable, Y_{ijbt}, is the marriage outcome variable of an individual i of birth cohort b in *hukou* province j and year t. $Fine_{jb}^{18-25}$ denotes the mean value of the fertility penalties in province j for birth cohort b when aged 18-25. The coefficient, $\beta_1(s)$, is of central interest because it captures the effects of the OCP penalties on marriage outcomes. We match the penalty rate according to the individual's *hukou* province rather than their current living province, because most inter-province migrants in China cannot change their *hukou* place. By doing so, we actually assume that the marriage market is independent within each province, and that individuals seek for potential spouses in the same *hukou* province. Therefore, it may be a potential concern if many people once changed their *hukou* provinces or met their spouses in other provinces. We argue this may not be a first-order issue. First, the cohorts we choose in this analysis are born before 1980 and most of the people did not change their *hukou* province. Although the census data do not provide information about previous *hukou*, we calculate this using other micro-level data sets and find the proportion with changed *hukou* provinces is smaller than 5 percent in the same cohorts. Second, the individual activities over the life cycle, such as migration and other social activities, are mostly conducted within the same province. For example, migration in China is mostly intra-province rather than inter-province; the proportion of people whose current living province is the same as their *hukou* province is over 93 percent in our sample. In addition, 97 percent of individuals have the same *hukou* province and birth province, and over 90 percent of people have the same current-living and birth provinces, according to birth province information in the 2005 census. These findings provide validity for matching the fertility penalties according to the *hukou* province. But it is noteworthy that our results are consistent if we use the birth province in census 2005 or current living province to match the information.

The term, X_{ijbt}, includes continuous variables such as the male and female proportions of minorities, and the Han and minority proportions of men, in the local province j of birth cohort b, which are used to control the relative size of Han and minorities as well as the gender composition in the local marriage market. The other term, D_{ijbt}, includes a series of other covariates: dummies for ethnicities to capture the time-invariant differences among the different ethnicities, such as time-invariant cultures or attitudes toward interethnic marriages; dummies for gender, age, and the interaction of the two, to allow for the time-invariant and age profile differences between men and women; dummies for province, type of *hukou* and their interaction, to control for the geographical fixed effects; and dummies for the year and their interaction with birth cohort dummies, to allow for the changes of the age profile over time. Finally, we also control for the provincial specific linear trends in the birth cohort, $Prov_j \times YoB_b$, to capture the potential changes in local subjective attitudes towards staying single or being in an interethnic marriage. This framework is the main identification strategy throughout our analysis, and the standard errors are clustered at the province level to allow autocorrelation within the same province over time.

Considering the differences in marriage markets between Hans and minorities, we allow for this heterogeneity by dividing the sample by Hans and minorities to conduct regressions for the two groups separately. There are some other good reasons for doing so. First, the OCP aims to restrict the population of Han people rather than minorities, and thus we expect differential effects of fertility penalty on unmarried status for the two groups. Second, since the numbers of Han and minority people involved in H-M marriages are the same but the H-M marriage proportion in each group are significantly different because of different population sizes, it may be more straightforward to interpret the estimates if we conduct regressions separately for the two groups. Finally, the coefficients β_1 could be interpreted at the individual level rather than at the couple level for different ethnic groups, and then it would be easier to estimate the corresponding elasticities in equation 2.3 directly to calculate the potential welfare loss.

2.5.3 The OCP increased the proportion with an unmarried status

Table 2.2 reports the OLS estimates for the impacts of the OCP on unmarried status. The first three columns are the results for Han people and the rest are for minorities. The estimates suggest that an increase in OCP penalties by one year of local household income predicts an increase of 1.7 percentage points in the unmarried rate for Hans, while the estimate is insignificant and much smaller for minorities (0.46 percentage point). Since the mean value of the unmarried rate is 4.4 percent for Han people and 6.6 percent for the minorities, the effects on unmarried status for Han are larger than those for minorities on both absolute and relative scales.

By dividing the Han sample into preferential-policy regions and non-preferential-policy regions, we find the effects of the OCP on unmarried status are greater and more significant in non-preferential-policy regions. An increase in OCP penalties by one year of local household income leads to an significant increase of 1.97 percentage points in the unmarried rate for Hans in non-preferential regions but only 0.93 percentage point for Hans in preferential-policy regions. Since there is no significant difference between the mean unmarried rates for the two different types of regions, the gap in effects is larger on both absolute and relative scales. In contrast, for the minorities, the effects of the fertility penalties are insignificant, much smaller, and even opposite in the preferential-policy regions. One possible explanation is that minority people would become relatively more valuable in the marriage market if the penalty has increased, because they have additional birth quotas. These results for minorities also provide some supportive evidence to the exogeneity of the fertility-penalty rate. That is, we should also find some effects of the OCP for minorities if the effects were driven by some omitted variables correlated with both penalty rates and the unmarried rates, such as economic development or changes of attitudes towards marriage.

Figures 2.4a and 2.4b show the gender-specific point estimates for $\beta_1(s)$, as well as the corresponding 90-percent confidence intervals. Figure 2.4a presents the results for the Han people. An increase of one year of local household income in the penalty rate causes an increase of 2 to 3 percentage points in unmarried rates among Han people. Also,

Table 2.2: *Impact of OCP on Marriage Outcomes: Unmarried Status*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Unmarried status (Yes = 100)					
	Han sample			Minority sample		
Sample	Full sample	Preferential-policy regions	No-preferential policy regions	Full sample	Preferential-policy regions	No-preferential policy regions
Mean of Dep. Var.	4.44	4.95	4.21	6.57	6.61	6.49
Fertility penalties at age 18-25	1.746***	0.934	1.971***	0.457	-0.204	0.439
	(0.528)	(0.537)	(0.611)	(0.619)	(0.411)	(0.583)
Observations	5,223,157	1,622,652	3,600,505	454,154	289,864	164,290
R-squared	0.102	0.104	0.102	0.125	0.119	0.145

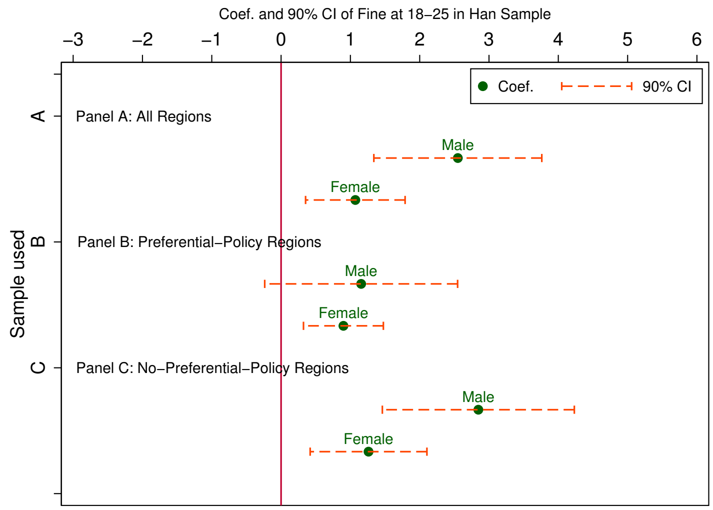
Notes: Data source is Census 2000 and 2005. Dependent variable is multiplied by 100 so the coefficients can be interpret in percent. The covariates include the local minority proportion in the birth cohort (Local Minority Prop.), local male proportion in the birth cohort (Local Male Prop.), dummies for ethnicities (Ethnicity FE), gender, age and their interaction (Gender & Age FE), *hukou* province, type of *hukou* and their interaction (Province & *Hukou* FE), survey year and its interaction with age (Age & Year FE). The province-specific year of birth linear trends (Province-Yob Trends) are also included to control the potential changes in local subjective attitudes towards marriage. Sampling weights are applied and robust standard errors in parentheses are clustered at province level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

the magnitudes of the coefficients are larger for men than for women. For example, in non-preferential-policy regions, the coefficient for men is 2.5 times larger than that for women. But this may not mean that the effects for men are larger. Because the mean values of the unmarried rates for men is also much higher than that for women, the effects of the OCP on unmarried status are similar for men and women on a relative scale.[19] In contrast, Figure 2.4b shows that the impact of the OCP on the unmarried rate is consistently much smaller and more insignificant for the minorities in all the subsamples.

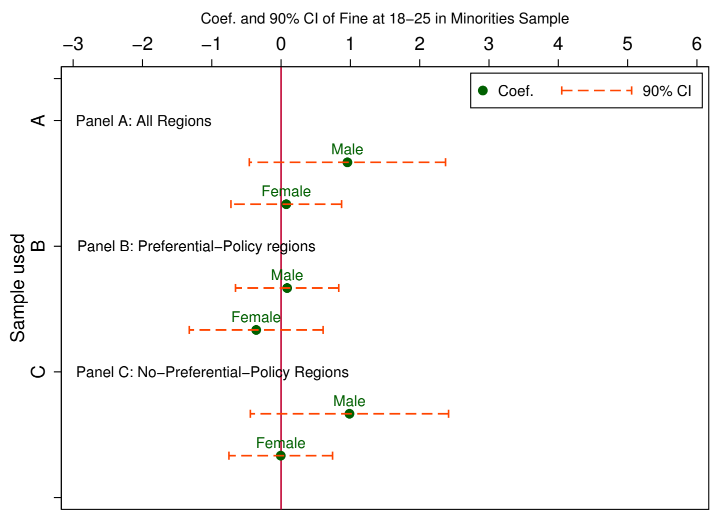
[19]In the sample, 7.2 percent of Han men and 1.6 percent of Han women are unmarried, and 10.3 and 2.6 percent of minority men and women, respectively, are unmarried.

Figure 2.4: *Impact of the Fine Rate of the OCP at age 18-25 on Unmarried Status, by Gender, Region and Ethnicity*

(a) Impact of the OCP on unmarried for Han ethnicity, by Gender and Region



(b) Impact of the OCP on unmarried for the Minorities, by Gender and Region

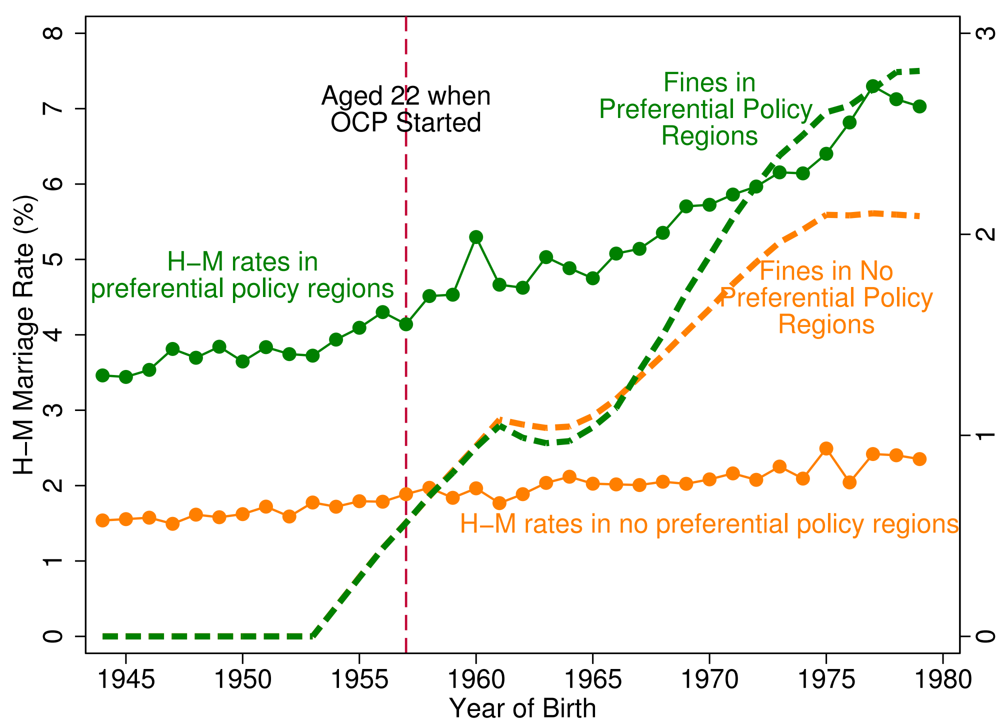


Notes: The data source is Census 2000 and 2005. We estimate the effects of the OCP penalties on the unmarried rate. Figure a and figure b report the OLS coefficients on the fertility penalties at age 18-25 and the corresponding 90% confidential intervals for the Han people and minorities, respectively. Standard errors are clustered at province level and sampling weights are applied.

2.5.4 The OCP increased H-M marriages

We investigate the effects of the OCP on H-M marriages in this section. As mentioned above, some regions consistently allowed H-M couples to have more children, while others did not. The non-preferential-policy regions are used as the control group in this section. Before the regression analysis, we plot the H-M marriage rate of all couples over the birth cohorts in Figure 2.5, based on whether the local region had the preferential policy.

Figure 2.5: *H-M Marriage rate and Fertility Fine at 18-25 over Year of Birth, by Preferential-Policy or No-Preferential Policy Regions*



Notes: The data source is Census 2000 and 2005. The H-M marriage rates and penalties are plot against the birth cohorts, by whether the region has preferential policies or not. Sampling weights are applied.

Figure 2.5 shows fairly parallel trends in the H-M marriage rate across the two types of regions before the early 1950s cohorts. The preferential-policy regions saw an increase from 3.5 to 7 percent and the non-preferential-policy regions saw an increase from 1.5 to 2.3 percent. However, the two lines start to diverge after the 1955 birth cohorts, who were aged 25 at the start of the OCP. The preferential-policy regions increased by 2.8 percentage points from 4.1 to 6.9 percent while those without the preferential policy only increased by 0.3 percentage points from 1.9 to 2.2 percent. However, the birth cohort trends for the average fine rates at age 18-25 for both types of regions, as presented by the two dashed lines, are very similar. This implies that the strictness of the OCP itself may not have created significant differences. Thus, the divergence of the H-M marriage rate of the two types of regions should be mainly caused by the preferential policy for H-M couples.

Since the increase in the H-M marriage rate in the preferential-policy regions, as shown in Figure 2.5, may merely be caused by a higher minority proportion in the local population, we conduct the regression analysis, and report the results in Table 2.3. The estimates in columns 1 and 4 of Table 2.3 show positive impacts of the OCP on the H-M marriages for both Han and for minorities. And the rest of the columns show that the effects are larger and more significant for the preferential-policy regions for both Hans and minorities, suggesting that the local minority proportion may not be the first-order factor that leads to the pattern in Figure 2.5. Specifically, an increase in the penalty rate by one year of local household income is associated with an increase of 0.6 percentage points in the H-M marriage rate for Han people and with an increase of 2.1 percentage points for the minorities. But the effects are much smaller and insignificant for the Han people in the non-preferential-policy regions. We also find that, in the non-preferential-policy regions, the minorities became less likely to marry Han people because doing so would "waste" the birth quota which is valid only if they were to marry other minorities.

Figures 2.6a and 2.6b show consistent results with the gender-specific subsamples. Also note that the impact of the OCP is quite similar between men and women; we do not find a significant gender difference in the marriage-behavior response to the OCP, either in absolute or relative scales.

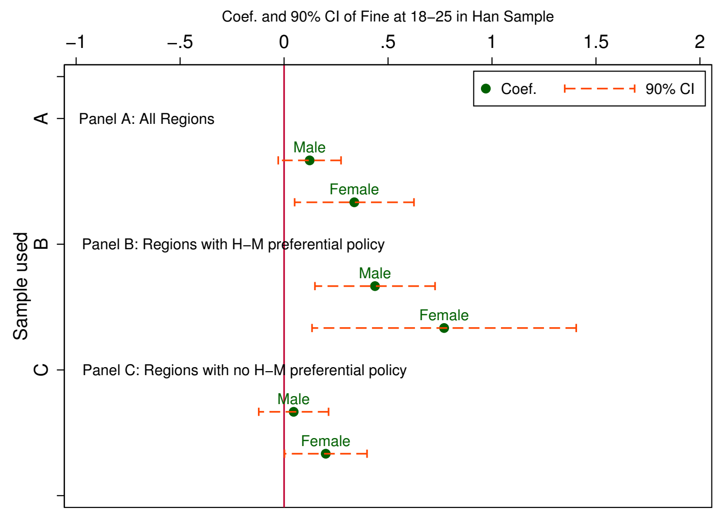
Table 2.3: *Impact of OCP on Marriage Outcomes: Han-Minority Marriage*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Han-Minorities Marriage (Yes = 100)					
	Han sample			Minority sample		
Sample	Full sample	Preferential-policy regions	No-Preferential policy regions	Full sample	Preferential-policy regions	No-Preferential policy regions
Mean of Dep. Var.	1.61	3.00	1.01	17.4	14.3	23.7
Fertility penalties	0.227*	0.607***	0.116	0.863	2.063*	-0.666*
at age 18-25	(0.122)	(0.269)	(0.074)	(0.773)	(1.139)	(0.359)
Observations	4,330,059	1,320,064	3,009,995	362,918	231,661	131,257
R-squared	0.037	0.037	0.028	0.194	0.154	0.256

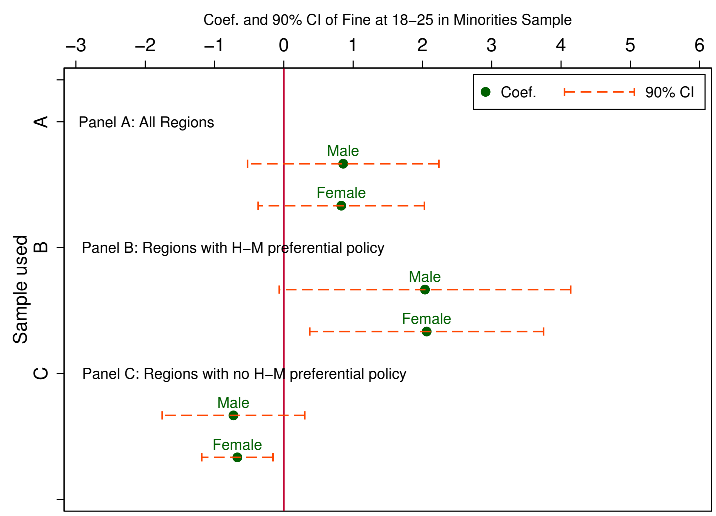
Notes: Data source is Census 2000 and 2005. The covariates are the same as those in Table 2.2. Sampling weights are applied and robust standard errors in parentheses are clustered at province level.
*** p<0.01, ** p<0.05, * p<0.1

Figure 2.6: *Impact of the Fine Rate of the OCP at age 18-25 on H-M Marriages, by Gender, Region and Ethnicity*

(a) Impact of the OCP on H-M Marriages for Han ethnicity, by Gender and Region



(b) Impact of the OCP on H-M Marriages for the Minorities, by Gender and Region



Notes: The data source is Census 2000 and 2005. We estimate the effects of the OCP penalties on the H-M married rate. Figure a and figure b report the OLS coefficients on the fertility penalties at age 18-25 and the corresponding 90% confidential intervals for the Han people and minorities, respectively. Standard errors are clustered at province level.

As mentioned earlier, we use the married-couple sample where information is complete for both spouses, so the effects estimated here must be interpreted as those effects that are conditional on being married. The first concern is that marriage ages are different across groups: if H-M marriages systematically have a higher or lower marriage age and this difference is correlated with the fertility-penalty rate, then the estimates of the impacts of the OCP on H-M marriages could be biased. However, we argue that this may not be a serious issue. First, the difference in the age of first marriage between H-M marriages and other types of marriages is small,[20] and we find no evidence that those involved in H-M marriages tend to marry later because of the OCP. Also, if we trim the sample to those aged over 30, we still find consistent effects. Note that over 95 percent of all marriages are formed before age 30, for any ethnicity and for any type of marriages.[21]

Another concern is that the OCP induced delayed marriages and thus people would get to meet more people before marriage and thus the marriage outcomes tend to be more diversified, especially for the preferential-policy regions because of the higher minority rate. To rule out this possibility, we conduct a similar analysis of interethnic marriages among minorities and report the results in Table 2.4. If the above hypothesis is true, we could expect that higher fertility penalties would also lead to more interethnic marriages among the minorities. The results suggest that the OCP did not motivate minorities to marry other minorities, and indicate that the above concern may not be an important issue.

---

[20]For men, the average age of H-M marriages is 23.8 and that of the other marriages is 24.2; for women, the ages are 22.0 and 22.2, respectively.

[21]The results are available upon request.

Table 2.4: *Impact of OCP fine on Interethnic Marriages among Minorities*

	(1)	(2)	(3)
Dependent variable	Interethnic marriage among minorities (Yes = 100)		
Sample	Full sample	Preferential-Policy Regions	No-Preferential-Policy Regions
Mean of Dep. Var.	3.777	3.950	3.417
Fertility fine rate	0.245	0.0604	0.502
at age 18-25	(0.157)	(0.150)	(0.343)
Observations	362,918	231,661	131,257
R-squared	0.045	0.040	0.072

Notes: Data source is Census 2000 and 2005. The covariates are the same as those in Table 2.2. Sampling weights are applied and robust standard errors in parentheses are clustered at province level.

*** p<0.01, ** p<0.05, * p<0.1

2.5.5 Children: Incentives for H-M marriage

We argue above that a primary motivation for the H-M marriages in the preferential-policy regions is to have more children legally. This section provides evidence to support this argument. The main difficulty in performing such a test is that the expectation about the number of children is unobservable. Based on the *ex post* data, we examine this by checking whether the regions with a more positive impact on H-M marriages are also the regions with less negative impacts on the number of children of H-M couples. The rationale is straightforward: if policy-induced H-M couples are formed to seek additional childbirth quotas, they would be more likely to have more births *ex post*, and thus the negative effect of the penalties on the number of children should be smaller.[22]

The presence of non-preferential-policy regions provides a natural control group. In these regions, we expect that the impact on H-M marriages should not be correlated with the impact on the number of children because individuals have no policy-induced incentives to form H-M couples. Specifically, we divide Han people into 62 subsamples by the *hukou* province and by the type of *hukou*. Then for each subsample, we conduct the following regressions:

$$HM_{ibt} = \theta_1 Fine_b^{18-25} + X_{ibt} + D_{ibt} + \epsilon_{i1}$$

where the dependent variable, HM_{ibt}, denotes whether an individual i is involved in a H-M marriage; $Fine_b^{18-25}$ denotes the average penalty rate at age 18-25 for the birth cohort b in the local province j; X_{ibt} denotes the minority proportion for both males and females in the birth cohort b of the local province; and D_{ibt} denotes a set of control variables, including indicators for education levels, gender, calendar year, and groups of birth cohorts (i.e., for every 10 years).[23] Then we keep the Han people involved in H-M marriages and conduct the following regressions on each subsample:

$$Children_{ibt} = \theta_2 Fine_b^{18-25} + X_{ibt} + D_{ibt} + \epsilon_{i2}$$

Here we keep all the other control variables the same and only switch the dependent variable to the number of children ever born to the mother in the household. For each subsample (s), we can get a θ_1^s and θ_2^s. We plot θ_2^s against θ_1^s and investigate how they are correlated, weighted by the population size in each cell. Figure 2.7a shows the pattern in non-preferential-policy regions and Figure 2.7b shows the pattern in the presence of

[22]We thank Professor Lawrence Katz for providing great suggestions for this part. Any errors are ours.

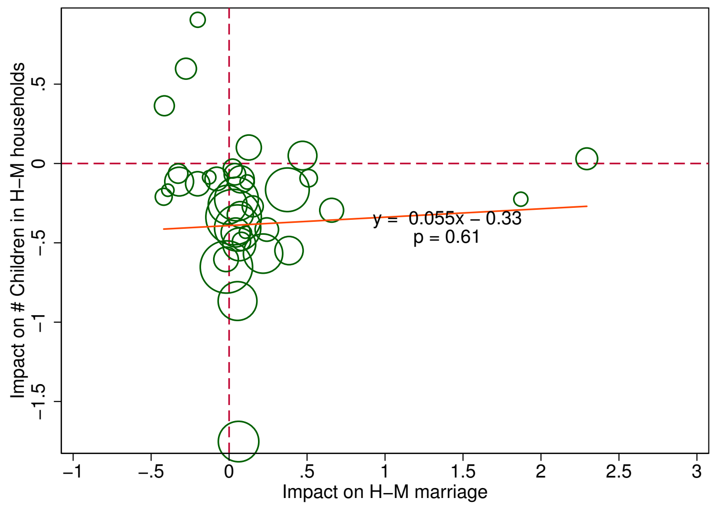
[23]We cannot control for the specific year of birth dummies here because the $Fine_b$ is in the level of the year of birth. The results are robust to the different years of birth categories.

preferential policies.[24] We find a very weak correlation between the impact on fertility and the impact on H-M marriages in Figure 2.7a, but a significantly positive correlation in Figure 2.7b, which implies that the effect of the OCP on fertility would be partially offset by the policy-induced H-M marriages. Therefore, Figure 2.7 provide some evidence that the expected number of children is an important factor that individuals consider in their marriage decision.

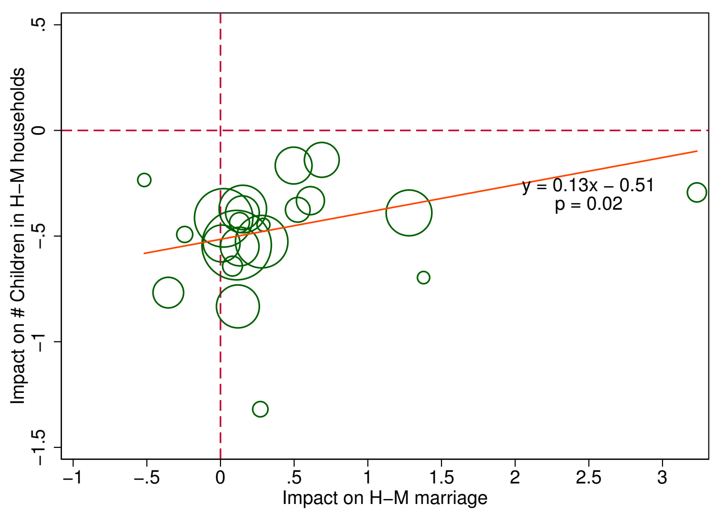
[24]Consistent with the finding that the policy-induced H-M marriages mostly happened in the preferential-policy regions, the weighted mean value of the impacts on H-M marriage is 0.1 in Figure 2.7a but 0.3 in Figure 2.7b.

Figure 2.7: *Associations between Impacts of the OCP on H-M marriages and those on Fertility of these couples, by Preferential-Policy or No-Preferential Policy Regions*

(a) Regions with no-preferential policy to Han-Minority couples



(b) Regions with preferential-policy to Han-Minority couples



Notes: The data source is Census 2000 and 2005. The full sample is divided by the province and for each subsample. The X-axis is the effects of the OCP penalties on H-M marriage rate and the Y-axis is the effects on number of children of those couples. Then we divide the sample by whether the region has the preferential policy or not, and report them in figure a and figure b, respectively. The size of the circle reflect the population size.

2.5.6 More "Transfers" to Minority Spouses in H-M couples

The third hypothesis of the model states that more "transfers" from Han spouses to minority spouses in H-M couples will happen if the implementation of the OCP becomes tougher and a preferential policy is in place. This is because the value of a minority partner as reflected by the additional birth quotas can be brought into marriage. However, the "utility transfers" cannot be directly observable. Thus we examine, in the preferential-policy regions, whether the minorities in H-M marriages marry more highly educated people in presence of higher fertility penalties.[25] We expect that, in preferential-policy regions, the educational attainments of the spouses of minorities should be higher in H-M couples since the minorities are more "valuable" in the marriage market as the penalty rates increase. In contrast, this should not hold true for either the spouses of the Han people in the same regions, or for the minorities in the non-preferential policy regions. Therefore, we trim the sample to those H-M couples, and divide the sample into regions with preferential policies and those without, and then conduct the following regression separately by Hans and minorities:

$$Education_{ijbt}^{spouse} = \alpha_0 + \alpha_1 Fine_{jb}^{18-25} + X_{ijbt} + D_{ijbt} + \gamma_j Prov_j \times YoB_b + \epsilon_i$$

where the dependent variable is education level of the spouse, on a scale of 1 to 5; the larger the value, the higher the education level. All the other variables are kept the same. Panel A and Panel B of Table 2.5 report the ordered logit estimates for Hans and for minorities, respectively. Consistent with our expectation, the estimates show that higher penalty is significantly associated with a higher education level of the spouses of the minorities in H-M couples, but this positive association *only* exists for the minorities in the preferential-policy regions. The coefficient is as high as 0.96. By comparison, the coefficient for the Hans in the same regions is 0.019, and that for minorities in non-preferential-policy regions is 0.03. Both of the latter two are insignificant.

[25] In previous literature (Chiappori *et al.*, 2009; Lafortune, 2013), education is viewed as pre-marital investment and predicts higher household income, and we consider the education of spouse as received utility transfer in marriages.

Table 2.5: *Ordered Logit Estimation: Impact of the OCP Penalties on Education of Spouse among H-M marriages*

	(1)	(2)	(3)
Dependent variable	Education Level of Spouse (1-4, larger for higher education)		
Sample	Full sample	Preferential-policy regions	No-Preferential policy regions
Panel A: The sample is the Minorities in the H-M marriages			
Fertility penalties	0.0578**	0.0955*	0.0306
at age 18-25	(0.0273)	(0.0507)	(0.0339)
Observations	63,005	34,566	28,439
Panel B: The sample is the Han people in the H-M marriages			
Fertility penalties	0.0380	0.0187	0.0579
at age 18-25	(0.0259)	(0.0428)	(0.0450)
Observations	63,005	34,566	28,439

Notes: Data source is Census 2000 and 2005. Only H-M couples are included. The covariates are the same as those in Table 2.2. Ordered logit estimation is applied. Sampling weights are applied and robust standard errors in parentheses are clustered at province level.

*** p<0.01, ** p<0.05, * p<0.1

2.6 Welfare Analysis

Recalling that reduced-form elasticities are sufficient statistics for the deadweight loss of social welfare, this section applies the individual behavioral response to the OCP penalties to the equation 2.3 to calculate the welfare loss caused by the distortion. The most important parts of equation 2.3 are the three terms in parentheses. The first two terms reflect the distortion in the marriage market and the third term captures the reduction in fertility.

Based on the data of the number of children observed in the each household, we can directly calculate the number of illegal birth children c_{ij}. Then we use the same identification strategy above to estimate the effects of penalties for different types of marriages to make the whole analysis consistent. Table 2.6 reports the results. Consistent with our expectations, the effects are mainly from H-H couples. The insignificant but sizable coefficient for Han-Han couples reflects a large heterogeneity within the population and are consistent with the ongoing debate about the magnitude of the policy-induced fertility decline (Schultz and Zeng, 1995; McElroy and Yang, 2000). Columns 3 and 4 shows that the effects of illegal birth are one scale smaller for H-M couples, and around zero for M-M couples.

In Table 2.2, 2.3, and 2.6, we have estimated the needed marriage and fertility responses to the OCP penalties to calculate the welfare loss. Table 2.7 reports the results. We calculate the loss by Han and minorities, respectively. Panel A reports the basic statistics in equation 2.3 (i.e., P_i, r^i_m, r^i_{ij} and c_{ij}). Panel B reports the elasticities of unmarried, intra- or inter-ethnicity marriage, and number of illegal children born with respect to the fertility penalties, by the ethnicity combinations of i and j. Panel C reports the welfare gain/loss induced by one unit increase in the penalty (the unit is yearly local household income) for each ethnicity i. Along with the notation in the equation 2.3, we specifically calculate the marriage distortion and the fertility reduction in the parentheses for each ethnicity combination, and report them in the first two rows. The unit for welfare loss is the percentage of yearly household income. So for the Han ethnicity, the welfare loss originates from both fertility reduction (-3.32) and marriage market distortion (-0.71), indicating that the distortion of the marriage market actually captures 18 percent of the total welfare loss for the Han people. For the

Table 2.6: *Effects of the OCP Penalties on Number of Illegal Births*

	(1)	(2)	(3)	(4)
Dependent variable	Number of illegal births			
Sample	All the couples	Han-Han Couples	Han-Minority Couples	Minority-Minority Couples
Fertility penalties at age 18-25	-0.0191	-0.0231	-0.00620	0.000130
	(0.0285)	(0.0326)	(0.00928)	(0.000142)
Observations	4,692,977	4,263,273	133,375	296,329
R-squared	0.263	0.248	0.255	0.016

Notes: Data source is Census 2000 and 2005. The covariates are the same as those in Table 2.2. Sampling weights are applied and robust standard errors in parentheses are clustered at province level.
*** p<0.01, ** p<0.05, * p<0.1

minorities, some of them actually were better off from the OCP in the marriage market and the welfare loss in the fertility reduction is also smaller in magnitude than that for Han people. The final column reports the social welfare loss by calculating the mean values weighted by the population proportion P_i. These estimates suggest that the one unit increase in penalty will induce a welfare loss, which is 3.75 percent of local yearly household income. Because the average penalty at age 18-25 is 1.3 (times of household income) for those birth cohorts born later than 1955, by assuming that the elasticities are constant across the birth cohorts afterwards, we conclude that the total welfare loss caused by the OCP is 4.9 percent of yearly household income, to which marriage distortion contributes 0.85 percent of yearly household income. It indicates that the traditional way to calculate the policy-induced welfare loss, which does not consider the distortion in marriage market (i.e., the distortion effects), would significantly underestimate the total welfare loss.

Therefore, these findings highlight the importance of considering the "distortion effects" when calculating relevant welfare loss. This raises the question as to under what circumstances do we need to consider the "distortion effects" and why most of the previous studies did not take them into account in their welfare analyses. Children ("the taxed good") are different from most normal goods in the market, because most children are born in wedlock and thus children are the natural fruits of marriages. A higher tax will prevent more people from marrying because their expected marriage gains become lower than the "married or not" threshold. The "mechanical effects" only consider the welfare loss among those who are married, and cannot take into account those whose expected marriage gains would fall below the threshold of "married or not" because they are censored when conducting the traditional analysis.[26]

[26]However, this study is not the first one to reveal the relationship between different "goods" and its consequences. For example, Busse *et al.* (2013) found that gasoline prices have significant impacts on prices and quantities of sales in the new and used car market.

Table 2.7: *Welfare loss caused by the OCP, by types of marriages*

	(1)	(2)	(3)	(4)	(5)
Ethnicity i	Han		Minority		Welfare gain/loss
Ethnicity j	Han	Minority	Han	Minority	(Weighted by P_i)
Panel A: Basic statistics in the data					
(P_i) Prop. of i	0.93		0.07		-
(r^i_m) Married rate of i	0.96	0.96	0.94	0.94	-
(r^i_{ij}) Prop. of i married to j	0.98	0.02	0.17	0.83	-
(c_{ij}) Illegal births	0.25	0.10	0.10	0.00	-
Panel B: Estimated Elasticities with respect to penalties					
(e^i_m) Married rate of i	-0.03	-0.03	-0.01	-0.01	-
(e^i_{ij}) Marriage $i-j$	0.00	0.22	0.08	-0.02	-
(e^c_{ij}) Illegal births	-0.14	-0.09	-0.09	0.00	-
Panel C: Welfare gain/loss of unit change in fine rate (% of yearly household income)					
Marriage Distortion	-0.71		0.11		-0.65
Fertility Reduction	-3.32		-0.15		-3.10
Total	-4.02		-0.04		-3.75

Notes: Data sources are Census 2000 and 2005. Statistics in Panel A are calculated from the corresponding samples. The estimates in Panel B are calculated from the results in Tables 2.2, 2.3, and 2.6 . The estimates in Panel C are calculated from the results in Panel A and Panel B by plugging them into the equation 2.3. Welfare loss estimated in column 5 is the population weighted mean of those for the Han people and minorities.

2.7 Conclusions and Discussion

This study provides new evidence on the implications and extensions of the transferable utility model by exploiting the plausibly exogenous deductions in marriage gains that are caused by the strict fertility policies in China, and estimates the welfare loss caused by the OCP in both fertility and marriage.

Using plausibly exogenous variations in the ethnicity-specific assigned birth quotas and different fertility penalties across provinces over time, we provide new evidence for the transferable utility model by showing that 1) The higher the OCP penalty at age 18-25 is, the higher the unmarried rate is, especially for the Han ethnicity; 2) an increase in the penalty rate induces more H-M marriages, but only in the preferential-policy regions; and 3) the minorities in interethnic marriages are more likely to marry highly-educated Han spouses when the penalty rate is higher in the presence of preferential policies.

Based on the theoretical framework, we further estimate the welfare loss induced by the OCP. The welfare loss is composed of two parts: one is the reduction in individual fertility (the "mechanical" effect), and the other is the distortion in the marriage market ("distortion" effects). More importantly, the welfare loss depends only on the fertility and marriage outcome elasticities, with respect to the fertility penalties. Applying the estimated reduced-form elasticities to the model shows that the distortion of the marriage market actually brings about a welfare loss approximately equal to 0.85 percent of the yearly household income, which captures about 17 percent of the total loss caused by the OCP. The estimates suggest that the OCP has led to a large distortion in marriage equilibrium outcomes. The large impact on H-M marriage outcomes implies that the unintended but rational behavioral responses to the policy potentially create large and persistent impacts on the culture, development, and societies of minorities. This calls for future studies on the behavioral and social impacts of other similar ethnic-specific policies.

Our findings also suggest a significant welfare loss caused by the OCP in both fertility and marriage. This paper enhances the current literature by studying the largest fertility policy in the world and by extending the *sufficient statistic* approach to the marriage market.

The estimates suggest that the relationship between different goods needs to be considered when studying the potential consequences of policies or taxations. Children (the "goods" that are taxed by the OCP) are different from other normal goods because they are the natural fruits of marriage. Our findings suggest the heavy tax on children has distorted the marriage market, which has contributed a significant proportion of welfare loss.

This study also suffers from some limitations. First, the government implemented other strict OCP regulations at the same time. For example, workers in the public sector risked losing their jobs if they did not comply with the OCP, and this is not covered by the monetary penalty we consider here. Although the evidence in this paper shows that the fertility penalty may be a good measure as suggested in previous literature (Edlund, 2000; Wei and Zhang, 2011; Huang *et al.*, Forthcoming), we need to bear in mind when interpreting the estimates that they only reflect the impacts of the monetary penalty rather than the overall effects of the OCP. In addition, some social conflicts have happened in the process of collecting the OCP penalties, especially in remote and poor regions. There are also some illegally born children who were not registered and were not eligible to receive formal education. These facts suggest that the deadweight loss induced by the OCP may be beyond the numbers in our study. Finally, our model and empirical analysis look into the effects on marriage and fertility only, but do not take into account other dimensions, including the impacts of the fertility policies on the status of women and the quality of children, as well as some possible spillover effects on human capital and social burden, though some of these factors are investigated in previous literature. We are looking forward to future studies, which may shed light on these questions.

Chapter 3

One-Child Policy and the Rise of Man-Made Twins[1]

"One is too few," said a woman (in China) waiting at the hospital. "People want to have a second child." —Elizabeth Grether, ABC News, August 3, 2011

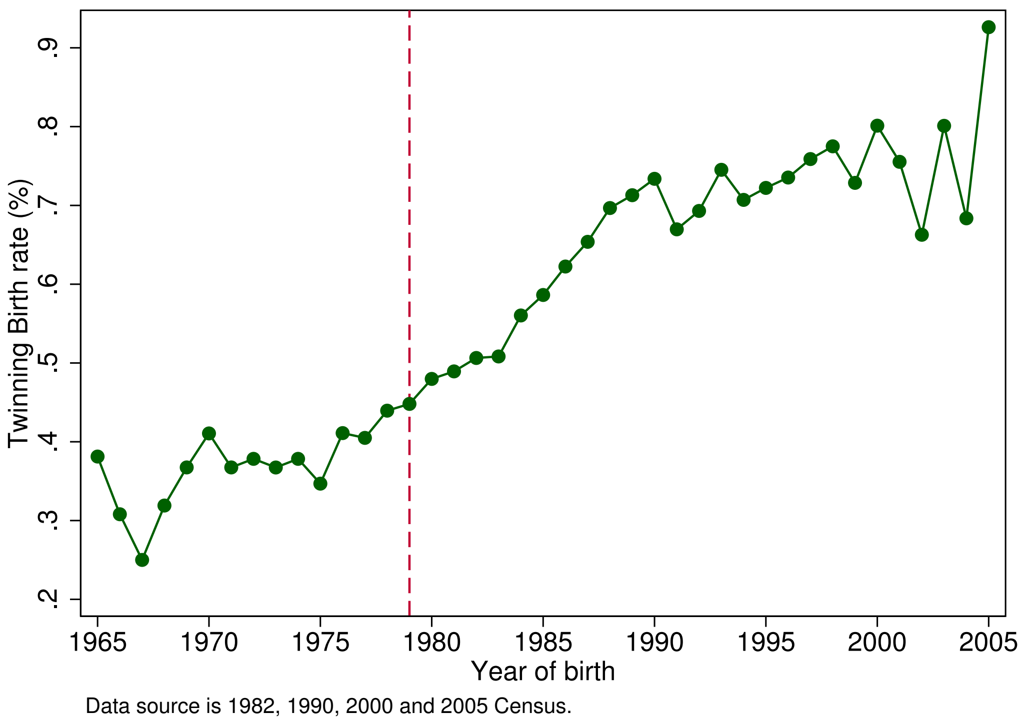
3.1 Introduction

Since Rosenzweig and Wolpin (1980a,b), an established strand of the economics literature has used samples of twins to deal with biases from unobserved factors, particularly in estimating educational returns and the effects of family size (e.g., Ashenfelter and Krueger, 1994). This methodology has been further developed and followed by studies in various contexts, including developing countries such as China (e.g., Li *et al.*, 2008; Rosenzweig and Zhang, 2009).

Previous studies generally assume that twins are born randomly across the population conditional on observed biological factors, such as the mother's age at childbirth and ethnicity. But the studies do not consider the possibility that the birth of twins, as a fertility behavior, could be manipulated in response to the distorted incentives imposed by relevant

[1]Co-authored with Xiaoyan Lei and Yaohui Zhao.

Figure 3.1: *Twining Birth Rate against Year of Birth, 1965–2005*



Notes: Data are from Census 1982, 1990, 2000, and 2005. Twinning rates in each birth cohort are plotted against year of birth. The vertical dashed line marks 1979, when the One-Child Policy was formally introduced.

policies. For the first time, our paper investigates the impact of fertility policy on births of twins to examine the behavioral responses to the distorted incentives of the One-Child Policy (OCP) in China.

A first glance, the timing of OCP and the trend in the rate of twin births suggest strong correlation between them (Figure 3.1): the rate of twin births reported in population censuses more than doubled between the late 1960s and the early 2000s, from 3.5 to 7.5 per thousand births. The rate of increase in twin births was greater after 1979, when OCP was fully implemented.[2]

[2]The pattern of increasing twins in China is not unique and the rate of twin births is heterogeneous across different populations. For example, the rate of twin births increased from 18.9 to 33.3 per thousand births from 1980 to 2009 in the United States (Martin et al. 2012). Therefore, no conclusion can be drawn before more serious analyses.

There are reasons to believe that the correlation may not be accidental. When a couple is allowed one birth, the only legitimate way to have two children is to give birth to twins, which is supposedly out of the control of the couple. One option for achieving this goal is to take fertility drugs. Although fertility drugs are meant to be used for infertility by inducing ovulation, there is some anecdotal evidence that women sometimes intentionally take fertility drugs to obtain twins.[3] For those who fail to have twins, an alternative option is to report fake twins, i.e., to register two consecutive siblings as twins. For example, officials in Yunnan province identified 700 pairs of fake twins in 342 villages in 2000. More surprisingly, among 23 pairs of twins reportedly born in 1999, 18 were identified as fake.[4]

Using the timing and geographical variation of policy violation fines, we find that the policy accounts for at least one-third of the increase in twin births since the 1970s. Such policy-associated twins are more likely to be found in rural areas and in observed second births.[5] The results are robust to a set of alternative regression specifications.

After getting evidence consistent with the hypothesis of man-made twins, we further investigate the mechanism through which this happens, i.e., whether they get twins by measures like reporting fake twins. We find that the birth gap between the first birth and the observed second twin birth is 0.08 year longer relative to an observed second single delivery after the policy was introduced. We also find that height difference within twins is larger where the OCP fine rate is higher. These findings are consistent with the hypothesis that OCP incentivizes parents to report non-twin children as twins.

The structure of the paper is as follows. Section 2 provides background on China's OCP. Section 3 describes the data and the empirical results. Section 4 explores the potential measures that people take to have twins. Section 5 concludes with a discussion of the

[3]Because of the lack of stringent regulation, these drugs are easily accessible, for example through online pharmacies and private hospitals. News Source: http://abcnews.go.com/Health/chinese-women-fertility-drugs-bypass-child-policy/story?id=14219173.

[4]News Source (Chinese): http://www.people.com.cn/GB/channel1/13/20000728/163617.html.

[5]Note that observed second births may not be real second births if the parents report fake twins. As discussed later, the parents possibly report the second birth child and the third birth child together as twins, and these twins will be observed as the second birth.

findings, policy implications, and suggestions for further research.

3.2 Background

OCP was introduced in 1979 to alleviate social, economic, and environmental problems in China (Greenhalgh, 1986; Wang *et al.*, 2012). Legal measures, such as monetary penalties and subsidies, have ensured the effective enforcement of OCP since 1979. Because of the heterogeneous regional development across China, Central Party Committee "Document 7" devolved responsibility from the central government to the local and provincial governments. The devolution allowed for regional variation in family planning policies, such as the amounts for monetary penalties or subsidies (Greenhalgh, 1986). However, OCP mainly focused on the Han ethnicity, the largest ethnic group in China, with more than 90 percent of the population.

In addition to the timing of its implementation, the additional measure of OCP in this study is the average monetary penalty rate for one unauthorized birth in the province-year panel from 1979 to 2000.[6] The OCP regulatory fine (policy fine) is formulated in multiples of annual income (Ebenstein, 2010; Wei and Zhang, 2011).

Since a period of approximately nine months is needed from the beginning of pregnancy until birth, parents' decision to have a child should be made close to a year in advance. For each birth, we construct a variable, policy fine rate, which is the weighted mean value of the fine rate in the 12 months just before the pregnancy in a given province.[7] The effective fine rate was zero for children born before 1979, when OCP started. We drop the children born after 2001, because the fine rate is not available after 2000.

[6] Details on the construction of this variable can be found in Ebenstein (2010).

[7] Because the 1982 Census does not have birth month information, we assume the children surveyed in 1982 were born in June and conduct the same procedure. The estimates do not rely on the OCP measure we constructed here: results are consistent if we simply use the fine rate one year before the child was born. (The results are available upon request.) When matching the policy fine rate to the current local province, we assume that the province of birth is the province of current residency, which may not be true due to migration. Using census data after 1990 with information on place of birth, we found over 95 percent of the children live in the same province where they were born, indicating that interprovincial migration should not be a big issue of concern in the analysis.

3.3 Data and Empirical Results

3.3.1 Census Data

The data used in this study are from the 1982, 1990, and 2000 Population Censuses and the 2005 One-Percent Population Survey. All the data sets contain years of birth, region of residence, type of residence (urban/rural), gender, ethnicity, education, and relation to the household head. The data sets after 1982 also include month of birth. For women older than 15, the data also provide information about their fertility history, including number of children ever born and number of living children.

For the analysis of twin births and family background, we first keep only those households with at least one child and with information available for the mother. We restrict the sample to those whose household heads and spouses are their first marriage and further restrict the sample to those households with equal numbers of reported living children, children ever born, and children observed in the survey. Doing so ensures that all the children in each household have the same mother observed in the household and that the sample covers the information needed for the children. In case we miss children who have moved from the household, we further drop households with children over age 17 in the survey. We finally drop households where the mother's age at childbirth is either younger than 15 or older than 50, as these subsamples may be too special or may contain recording errors.

Table 3.1 presents the summary statistics: the rate of twin births is 0.58 percent for Han (the majority) and 0.44 percent for minorities. Twins are defined as children in the same household with the same birth year and birth month.[8] Observations are at the birth level, so twins are treated as a single observation, because they are in the same observed birth.

[8]Because the 1982 Census data do not have information on birth month, we define twins in that year as those children born in the same household with the same birth year only. The results are almost the same when we drop the 1982 Census or define twins only using the year of birth in all the other data sets.

Table 3.1: *Summary statistics*

Variables	(1) Full sample	(2) Parents are Han	(3) Either parent is minority
Twinning Rate (%)	0.58	0.59	0.43
	(7.58)	(7.65)	(6.56)
Rural area	0.73	0.72	0.81
(Yes = 1)	(0.45)	(0.45)	(0.40)
Both parents Han ethnicity	0.93		
(Yes = 1)	(0.26)		
Age in years	8.04	8.08	7.62
	(4.67)	(4.67)	(4.62)
Mother's age at childbirth	23.25	23.28	22.82
	(2.97)	(2.95)	(3.18)
Birth order			
First (Yes = 1)	0.57	0.57	0.51
	(0.50)	(0.50)	(0.50)
Second (Yes = 1)	0.30	0.29	0.32
	(0.46)	(0.46)	(0.46)
Third or above (Yes = 1)	0.14	0.14	0.18
	(0.35)	(0.34)	(0.38)
Observations	6,071,870	5,654,203	417,667

Notes: The data are from the Census 1982, 1990, 2000, and 2005. The sample is restricted to births before 2001. Standard deviations are in parentheses.

3.4 Impact of OCP on Twinning Rates in China

To evaluate the effects of OCP on the twinning rate, we estimate the following equation:

$$Twins_{ijky} = \beta_0 + \beta_1 Fine_{jym} + \delta_k + \delta_y + \delta_{ky} + \delta_j + X_{ij} + \epsilon_{ij} \tag{3.1}$$

where the dependent variable, $Twins_{ijky}$, denotes whether birth i in year y and province$j$ is a twin birth in survey year k. $Fine_{jym}$ is the OCP fine rate defined above in province j for children born in year y and month m. The main coefficient of interest, β_1, gives the association of the OCP fine with the reported twinning rate and is interpreted as the impact of OCP. δ_k,δ_y and δ_{ky} are indicators for year of birth y, survey year k, and their combinations, respectively. δ_j denote the province dummies. X_{ij} is a set of covariates, including dummies for residence type (urban/rural), parents' ethnicity (both Han or either a minority), birth order, birth month, mother's education level, and mother's age at childbirth as well as the provincial specific linear trends in birth cohorts.

The first three columns in Table 3.2 report the OLS estimates for β_1 in Equation 3.1 with standard errors clustered at the provincial level. The results indicate that an increase equivalent to one year's income in the policy fine is associated with a 0.066 percentage point increase in the twin birthrate among the whole sample. The estimates in column 2 suggest that 36 percent of the increase in twins in the Han ethnicity sample can be attributed to OCP.[9] As expected, the association and significance survive in the sample for Han ethnicity but diminish in the sample for minority groups.

We further interact the policy fine with Han ethnicity and report the results in column 4. Consistent with the above, the main effect of the fine diminishes and the interaction is positive and significant, with similar magnitude as in column 2.

Li *et al.* (2011) argue that spatial and temporal variation of OCP may be endogenous. They find that the policy fine increases with community wealth and the local government's

[9]The twinning rates before and after OCP are 0.39 percent and 0.67 percent, respectively. The mean value of the fine rate increases from 0 to 1.4 years of local household income. The part of the increase in the twinning rate that can be explained is 0.072*1.4/0.28 = 0.36.

Table 3.2: *Impact of One-Child Policy Fines on the Reported Birthrate of Twins in China, 1965-2001*

	(1)	(2)	(3)	(4)	(5)
	Full sample	Parents Han	Either Parent Minority	Full sample	
Variable	Reported Twinning Birth (Yes = 100)				
Policy fine rate (years of local household income)	0.066*	0.072*	0.011	0.022	
	(0.038)	(0.041)	(0.029)	(0.035)	
Policy fine rate * parents Han				0.052***	
				(0.016)	
Born after 1979 * parents Han					0.154***
					(0.030)
Observations	6,071,870	5,654,203	417,667	6,071,870	6,071,870
R-squared	0.001	0.001	0.001	0.001	0.001

Notes: Data source is Census 1982. 1990, 2000 and 2005. Robust standard errors in parentheses are clustered in provincial level. The One-Child policy fine is measured in years of local household income. Sampling weights are applied. Coefficients should be interpreted as percentage because all dependent variables have been multiplied by 100. Covariates include residency type, province, birth order, mother's education level, mother giving birth age, year of birth, survey year, and interactions between year of birth and survey year. Parents' ethnicity dummy is controlled for in columns 1, 4 and 5.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

birth-control incentives and decreases with the local government's revenue incentives. In particular, spatial and temporal variation in OCP may be affected by the local fertility rate, which, in turn, may correlate with the incidence of twins. To test this possibility, we regress future policy fines on prior twin birthrates and see if the latter has predictive power on the former. If the association is significant, then the endogeneity problem is worthy of concern. More specifically, in each regression, we use whether the observation is a twin birth as the key independent variable and the amount of policy fine required in the next year, three years later, or five years later as the dependent variable, respectively. Because the local government started to have the local policies in 1984, the year of "Document 7", I kept the post-1980 birth cohorts to check this. As shown in Table 3.3, the rate of twin births does not seem to have any predictive power on the amount of the policy fine over the next one, three, or five years, suggesting that the reverse causality problem may not be serious in this study. Also note that the coefficients are really small, which suggest that the predicted fine rate would change smaller than 0.01 even if the current twining rate had increased from 0 to 10 percent.

Then, we conduct another set of regressions without using the policy fine rate, exploring only the timing of OCP:

$$Twins_{ijky} = \beta_0 + \beta_2(Policy_{1980} \times Han_i) + \delta_k + \delta_y + \delta_{ky} + \delta_j + X_j + \epsilon_i \tag{3.2}$$

where $Policy_{1980}$ denotes an indicator of whether birth i was in 1980 or after, and Han_i is an indicator for Han ethnicity of both parents. This difference-in-differences (DID) estimate in the final column of Table 1 indicates that OCP explains over 54 percent of the increase in twins, which is larger than that from estimates based on the policy fine rate.[10] The result is reasonable, since the policy fine captures only one means of punishment and the fine rate is averaged at the provincial level, which may miss some contributing variation within a province.

An important assumption of the DID estimation is that the trend for Han ethnicity

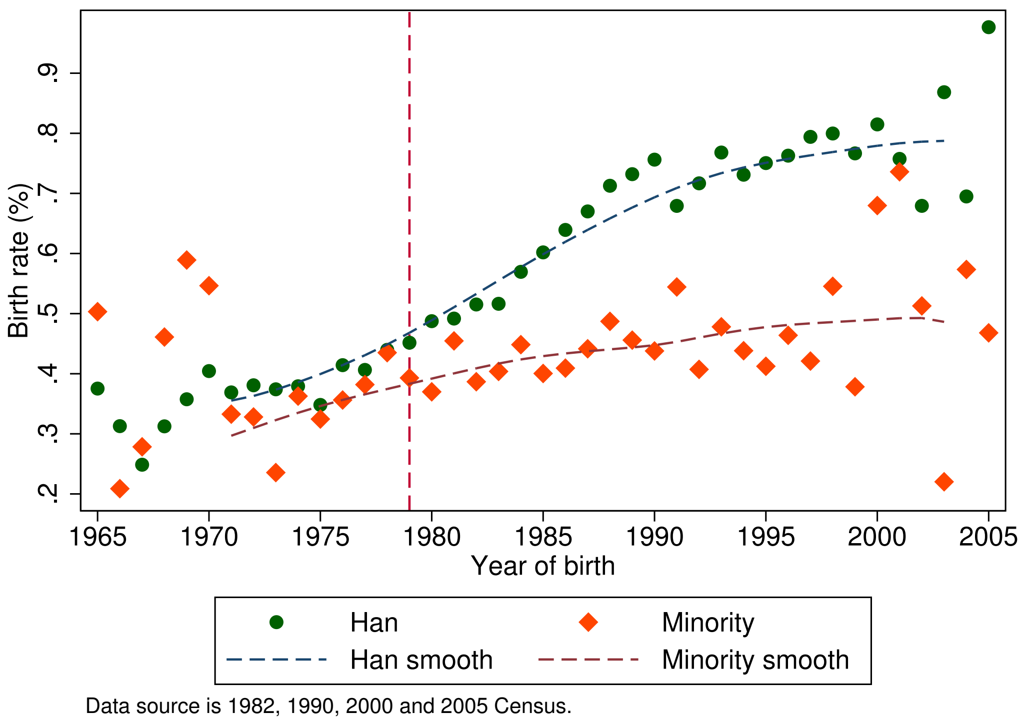
[10] The proportion that can be explained by OCP is 0.15/0.28 = 0.54.

Table 3.3: *One-Child Policy Fine Predicted by the Prior Rate of Twin Births, Post-1980*

	(1)	(2)	(3)
	One-Child Policy Fine Rate (Years of local household income)		
Variable	1 year later	3 years later	5 years later
Twinning birth	0.013	0.005	-0.002
(Yes = 1)	(0.008)	(0.005)	(0.006)
Observations	3,846,783	3,617,403	3,412,645
R-squared	0.754	0.778	0.800

Notes: The dependent variable the One-Child Policy fine rate in years of local household income. The data are from the Census 1982, 1990, 2000, and 2005. Post-1980 births are used because the One-Child Policy fine started in 1979 and provincial governments started to have local policies in 1984. Micro-level data are used to test whether contemporaneous twin births in local provinces have predictive power on the fine rates in the future, by regressing the fine rates in the next one, three, or five years on the indicator of twinning in the birth-level data, respectively, in columns 1, 2, and 3. The control variables in all columns are the same as those in column 1 of Table 1. They include continuous variables, such as the provincial time trend, and indicator variables, such as residency type, parents' ethnicity, province, birth order, mother's age at first birth, year of birth, survey year, and interactions between year of birth and survey year. Sampling weights are applied. Robust standard errors in parentheses are clustered at the provincial level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 3.2: *Twining birth rates against year of birth, by parents' ethnicity*



Notes: The figure plots the twin birthrate by the ethnicity group of the parents against the year of birth. The dashed lines plot Lowess-smoothed trends with bandwidth 0.8. The figure shows that the trends in the two groups are almost identical prior to the introduction of the One-Child Policy. However, the difference between Han and minority ethnicity increases after the introduction of the One-Child Policy in 1979 (the vertical dashed line). The increased difference is partly because the One-Child Policy mainly restricts the fertility of families of Han ethnicity rather than minorities, which motivates the Han to "make" twins.

without the policy would be similar to that for the minority group. Figure 3.2 examines this assumption by plotting the twin birthrate by the ethnicity group of the parents against the year of birth. As the figure shows, the trends in the two groups are almost identical prior to the introduction of OCP, indicating that the two trends without the policy are likely to be the same.

3.4.1 Heterogeneous Impact of OCP on Twinning Births, by Type of Residence and Birth Order

Enforcement of OCP differs in urban and rural areas. For example, urban areas strictly enforce the policy, while many rural areas allow a couple to have a second child if the first is a girl. This varying enforcement, together with other potential differences between the two areas, may result in heterogeneous effects.

The regression results by residence are reported in Table 3.4. The policy fine is positively correlated with the incidence of twins in urban and rural areas. The association in rural areas is larger and more significant, indicating the incentive to have twins in rural areas may dominate that in urban areas.

The incentive to have a twin birth may be stronger at different points in the two areas: urban parents may have to manipulate the birth date for the first child, while rural parents can wait until the second. We examine this heterogeneity by interacting policy fines with birth order dummies in the regressions. As shown in panel B of Table 3.4, twins in the second birth are the most policy relevant and the association is mainly reflected in rural areas, which is consistent with OCP enforcement.[11]

3.5 Mechanisms: Reporting Fake Twins and/or Taking Fertility Drugs

As mentioned above, people may either report single children as twins ex post or take fertility drugs ex ante to raise the probability of multiple children in a single birth. In this section, we try to identify the two channels by examining the birth gap between the first two observed births and the height difference within twins, respectively.

[11] Note again that the second birth here may not be the real second birth, because fake twins may be reported. As discussed in section 4, parents might report the second and third births together as twins, who are misleadingly seen as second births.

Table 3.4: *Heterogeneous Impacts of the Policy Fine on Reported Birth of Twins, by Type of Residence and Birth Order*

	(1)	(2)	(3)
	Parents Han	Subsamples by type of residence	
		Urban	Rural
Variable	Reported Twinning Birth (Yes = 100)		
Panel A: Using the policy fine only as the key independent variable			
Policy fine rate (years of	0.072*	0.060	0.084*
local household income)	(0.041)	(0.041)	(0.042)
Panel B: Interacting the policy fine with dummies for birth order			
First birth	0.039	0.056	0.033
* policy fine rate	(0.041)	(0.044)	(0.041)
Second birth	0.151***	0.093	0.168***
* policy fine rate	(0.046)	(0.057)	(0.047)
Third or above birth	0.072	0.022	0.088*
* policy fine rate	(0.045)	(0.060)	(0.043)
Observations	5,654,203	1,422,624	4,231,579

Notes: The dependent variable is whether the birth is twins or not (yes = 100). The data are from the Census 1982, 1990, 2000, and 2005. Column 1 restricts the sample to those births with parents of Han ethnicity; columns 2 and 3 further divide the sample into urban and rural subsamples. Coefficients should be interpreted in percentage because the dependent variables in all columns have been multiplied by 100. Covariates include dummies for residency type, parents' ethnicity, province, birth order, mother's age at birth, year of birth, survey year, and interactions between year of birth and survey year. Sampling weights are applied. Robust standard errors in parentheses are clustered at the provincial level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3.5.1 Impact of OCP on Birth Gap between the First Two Observed Births

If an elder child is registered with a younger one as twins, the birthdate of the reported twins tends to be registered as that of the younger child, since the parents have to wait until both children are born. If fake twins are reported to be born as observed second births, there should be a longer birth gap between the first two observed births because it is actually the gap between the first birth and the actual third birth. If parents plan to have twins by taking fertility drugs, however, there is no reason why they would report a delivery date that is later than normal. Therefore, we restrict the sample to the observed second births and conduct the following regressions:[12]

$$Birthgap_i = \beta_0 + \beta_3(Policy_{1980} \times Twins_i) + \beta_3^{'} Twins_i + \delta_k + \delta_y + \delta_{ky} + \delta_j + X_j^{'} + \epsilon_i \qquad (3.3)$$

In Equation 3.3, the variable $Birthgap_i$ denotes the observed birth gap between the current (second) and previous (first) delivery for birth i. $Twins_i$ denotes whether the current birth of the same parents for birth i is a twin birth or not, which captures the potential difference, if any, in the birth gap between single births and twin births driven by factors other than OCP. The coefficient, β_3, on the interaction term is of central interest because it reflects how much additional time is needed to give birth to twins than to a single child after OCP was implemented. If there are reported fake twins in the observed second births after OCP, we should expect the estimated coefficients for β_3 to be positive, because the birth gap between the first and second observed births would be lengthened by a report of fake twins.

The covariates δ_k, δ_y, δ_{ky} and δ_j have the same definitions as before. $X_i^{'}$ includes dummies for residence type parents' ethnicity, mother's education level, and mother's age at first childbirth as well as the provincial specific linear trends in birth cohorts.

Table 3.5 reports the OLS estimates for β_3 and $\beta_3^{'}$ for different samples. For the full

[12] It is possible that parents have twins in the first observed birth. We do not consider this case here, because it is not possible to calculate the gap between the first observed birth and the previous one, and the results in the previous section suggest that policy-related twinning is mostly concentrated in observed second births.

Table 3.5: *Difference-in-Differences Estimation for the Impact of thE One-Child Policy on the Age Gap between First and Second Births*

	(1)	(2)	(3)
	Full sample	Parents Han	Either parent minority
Variable	Age gap between first birth and second birth (in years)		
Twinning in second birth	0.078***	0.077**	0.100
* born after 1980	(0.027)	(0.028)	(0.108)
Twinning in second birth	0.188***	0.182***	0.288***
(Yes = 1)	(0.016)	(0.016)	(0.096)
Observations	1,822,396	1,690,608	131,788
R-squared	0.467	0.471	0.460

Notes: Data source is Census 1982. 1990, 2000 and 2005. Robust standard errors in parentheses are clustered at the provincial level. Sample is restricted to the second births. Sampling weights are applied. Covariates include residency type, province, birth order, mother's education level, mother's age when giving first birth, year of birth, survey year, and interactions between year of birth and survey year.
*** p<0.01, ** p<0.05, * p<0.1.

sample (column 1), we obtain a positive and significant estimate for β_3, showing that a twin birth needs an additional 0.08 year to achieve than a singleton birth does after the OCP was introduced. The next two columns provide results for Han ethnicity and minorities, respectively, and the significant estimates only appear in Han ethnicity.

3.5.2 Impact of OCP on Height Difference within Twins

In this section, we use height differences of the two children within a twin birth to test whether they are "man-made". We do the test for same-gender twins and mixed-gender twins, as the results of each have different implications. The detailed derivations are relegated to Appendix C due to space constraints. Here we summarize some basic results from the derivation: (a) if OCP is not relevant for twin births, the height difference within twins should be uncorrelated with the policy; (b) if OCP leads to more fertility drug use rather than reporting fake twins, a larger height difference is predicated to exist only within same-gender twins under OCP; and (c) if there are any reported fake twins, a larger height difference is expected both within same-gender twins and within different-gender twins.

For this analysis, we turn to the data from the China Health and Nutrition Survey (CHNS), which provides information on height. We try to use the same definition of twins as that with the Census data in the above analyses.[13] We match the twin sample with the fine data and drop those born after 2001, as we did for the Census sample. After these restrictions, we have 72 pairs in total, among whom 53 are same-gender twins and the rest are different-gender twins. Table 3.6 reports the summary statistics for the height gap and the gap/mean ratio. The mean height gap is 2.2 centimeters and the gap/mean ratio is 1.75 percent. As expected, different-gender twins have a much larger difference in height than same-gender twins (4.65 versus 1.30 cm).

To estimate the relationship between height difference and OCP, we conduct the following estimation:

$$HD_{ijyk} = \alpha_0 + \beta_4 Fine_{jym} + \delta_y + \delta_j + \delta_j + Z_i + \epsilon_i \tag{3.4}$$

where the dependent variable, HD_{ijyk} denotes the height difference within twin pair i in province j born in year y of wave k. The coefficient β_4 gives the association between the policy fine and the height difference within twins. Because there are only 72 pairs of twins observed in total and 33 different pairs in the regressions, we just combine the birth years into four groups (every five years as one group), δ_y, to capture the birth year effects. We also combine the neighboring provinces into five groups and control for the region fixed effects δ_j. The other covariates, Z_i, include indicators for whether the twins are same-gender twins, urban residence, whether the boy within different-gender twins is taller than the girl,

[13]The China Health and Nutrition Survey (CHNS) includes 26,000 individuals in nine provinces that contain approximately 56 percent of the population of Mainland China. The nine provinces vary substantially in geography, economic development, public resources, and health indicators. Data collection began in 1989 and has been implemented every two to four years since then (Jones-Smith and Popkin 2010). We first keep the children younger than 18 years in CHNS and select out the twins. Twins are defined as the children with exactly the same birth year and birth month within the same household. We define the height difference of each pair as the difference between the taller and the shorter child. Considering that the height gap may change as the children grow up, we also introduce another measure, the ratio of the height gap to the mean height of the pair (gap/mean).

Table 3.6: *Summary statistics in CHNS*

Variable	(1) All twins	(2) By type of twins: Same gender	(3) By type of twins: Different gender
Height gap (cm)	2.18	1.30	4.65
	(2.74)	(1.59)	(3.68)
Mean height (cm)	121.15	119.58	125.55
	(27.59)	(28.47)	(25.17)
Gap/mean ratio in percent	1.75	1.02	3.78
	(2.14)	(1.14)	(2.90)
Observations	72	53	19

Notes: Data source is CHNS. Standard deviations are reported in parentheses. Data source is China Health and Nutrition Survey. Twins are defined as children (aged below 18) born in the same household within the same month. For each pair of twins, height gap is defined as the height of the taller member minus that of the shorter one. Gap/Mean ratio is defined as twins' height gap divided by the mean height of the pair, and the values reported are multiplied by 100.

and continuous variables like average height of pair i, mother's age at childbirth, and age and age squared of the twins.

Panel A in Table 3.7 reports the estimates of Equation 3.4 with standard errors clustered at the provincial level. The first column shows that increasing the fine by one year's income is associated with an increase in the height gap of twins by 1.8 centimeters, suggesting that couples may have employed methods to make twins. Interacting the fine variable with the indicator variable of same- and different-gender twins in column 2 shows that the policy fine is positively associated with height differences for both types of twins. Columns 3 and 4 use the gap/mean ratio as the dependent variable and the results are similar.

Since CHNS is a panel data set and the same twins may be surveyed in different waves, we alternatively keep observations from only the latest wave for each pair and run the same regressions as above. Panel B in Table 3.7 reports the estimates for the key variables and the results represent the same pattern, with coefficients of larger magnitude than those in panel A. Altogether the results in this panel provide supportive evidence of parents reporting

Table 3.7: *Impact of the One-Child Policy Fine Rate on Height Difference Between Reported Twins*

Variable	(1)	(2)	(3)	(4)
	Height Difference (cm)		Height Difference/Mean Height	
Panel A: Full sample				
Policy fine rate (years of	1.85***		1.71***	
local household income)	(0.31)		(0.23)	
Policy fine rate *		2.10***		1.79***
same-gender		(0.38)		(0.26)
Policy fine rate *		0.69**		1.33***
different-gender		(0.25)		(0.27)
Observations	72	72	72	72
R-squared	0.63	0.64	0.63	0.63
Panel B: Only the latest wave is kept				
Policy fine rate (years of	3.29**		2.71***	
local household income)	(1.08)		(0.59)	
Policy fine rate *		3.37*		2.59**
same-gender		(1.56)		(0.83)
Policy fine rate *		2.92		3.21**
different-gender		(1.94)		(1.15)
Observations	33	33	33	33
R-squared	0.74	0.74	0.77	0.77

Notes: Data source is CHNS. Standard errors in parentheses are clustered in provincial level. Data source is CHNS. The twins sample used in this table are those born between 1979 and 2001. Height gap is defined as the height of the taller twin minus that of the shorter one, and gap/height ratio is defined as twins' height gap divided by the mean height of the pair. Each observation is derived from one pair of twins. Covariates include age and age squared of the twins, mother's age at childbirth, and indicator variables such as whether they are same-gender twins, urban residence, whether the boy is taller if the pair is different genders, provinces, year of birth categories, survey year, and provincial linear time trend.
*** p<0.01, ** p<0.05, * p<0.1.

consecutive children as twins to avoid the policy violation punishment.

3.6 Conclusions and Discussion

In 1979, the Chinese government launched the One-Child Policy, which led to hundreds of millions of couples involved in this strict family-planning program lasting for more than 30 years. In this paper, we find that an increase in the policy fine of one year's income is associated with an increase in twin births by approximately 0.07 per thousand births, indicating that at least one-third of the increase in twins since the 1970s can be explained by OCP. We then examine the heterogeneous effects by residence and find that the impact of the policy is larger in rural areas, where the policy raises the twin birthrate mainly for second births.

Furthermore, we find that since OCP was put into effect, the birth gap between the first two observed births is 0.08 year longer when the observed second birth is a twin than when it is not. In addition, the height difference within twins is positively associated with the policy fine and the association exists in both same- and different-gender twins. These findings support the hypothesis that OCP has incentivized people to have twins by reporting non-twin children as twins.

Since behavior response is related to deadweight loss of social welfare (Hendren, 2013), economics literature usually examines individuals' behavioral response to or against government policies. This study builds up the literature by examining people's behavioral response to OCP. Our estimates indicate a sizable behavioral response from Chinese couples. This finding should motivate future studies to examine individuals' behavioral responses to other undesirable public policies, and calculate the associated welfare gain or loss, especially in China, the largest developing country. This study also helps to explain the context and background of studies that use data on twins in China. Since couples can intentionally have twin births to bypass OCP, the distribution of reported twins in China may not be random. It is worth noting that the results do not rule out the possibility that women may take fertility drugs to have twins, and that the results allude to the importance of carefully screening twins with observable characteristics when analyzing Chinese data sets of twins.

There are some possible limitations in the current study and we leave these issues for future study. First, the evidence for fake twins provided in this paper is suggestive rather than determinant, because we cannot observe fake twins directly in our data. Identification of twins would require careful and detailed field surveys. It would be interesting to see how many twins are fake and how these fake twins are distributed across household socioeconomic status. Second, although CHNS is the largest data set we can find for the analysis on height difference, the sample size is too small; therefore, the results are sensitive to model specification and any attempt to generalize the results should be treated with caution. We hope the issue can be better addressed in the future with a larger data set. Finally, the fine rate measure we use in this study captures only one dimension of OCP. We hope that future studies can explore more comprehensive measures of the policy to test our story and investigate which dimensions of the policy lead to more fake twins.

Chapter 4

Understanding the Effects of Education on Health: Evidence from China

4.1 Introduction

The causal effects of education on health are of central interest to the economists. These effects are crucial parameters in the classical theoretical models of demand for health capital (Grossman, 1972) and the influences of childhood development on adult outcomes (Heckman, 2007, 2010; Conti *et al.*, 2010). Moreover, quantifying the extent to which education causally affects on health is essential to the formation and evaluation of education and health policies.

However, the empirical findings on causality are mixed. For example, Lleras-Muney (2005) used state-level changes in compulsory schooling laws (CSLs) in the United States as instruments for education and identified large effects of education on mortality.[1] In contrast, Clark and Royer (2013) used two education policy reforms in the United Kingdom and found no impact on mortality. The effects of education on mortality have also been found

[1]Identification of this effect is achieved by exploiting variation in the timing of the changes in the law across states over time such that different birth cohorts within each state have different compulsory schooling requirements.

in the Netherlands (van Kippersluis *et al.*, 2011) and Germany (Kemptner *et al.*, 2011) but not in France (Albouy and Lequien, 2009) or Sweden (Lager and Torssander, 2012).[2] The inconsistent findings in the literature reflect scarce evidence on the mechanisms, which is largely due to data limitation. Since most education reforms in industrial countries usually happened early and the changes were small in general, the affected cohorts were really old when surveys took place and the policies only induced small increase in education. For example, the education reforms in Lleras-Muney (2005) happened between 1914 and 1939 and in most of the states the changes in minimum school-leaving age were less than two years.[3] And the two reforms in Clark and Royer (2013) happened in 1947 and 1972, both increasing the minimum school-leaving age by only one year.

To shed light on the causal effects of education and the mixed findings in the literature, this study explores the compulsory schooling laws (CSLs) in China to investigate the causal effects of education on health and explores the possible mechanisms. The unprecedented nationwide education reform initiated in 1986 made nine-year schooling (i.e., up to the junior high school) compulsory and 16 years the minimum school-leaving age for all the regions in the largest developing country.[4] This education reform resulted in great achievements: the enrollment rate for junior high school increased by 26 percentage points, from 69.5 percent in 1986 to 95.5 percent in 2000, and the number of students enrolled in junior high school increased by 8.9 million.

Following the previous literature (Lleras-Muney, 2002, 2005), I first exploit the variation in the different timing of policy adoption across the provinces. Because the central government allowed the provincial governments to implement the policy separately, I construct a *CSLs-eligibility* indicator for the birth cohorts in the corresponding provinces. Since the timing variation across provinces is small (the gap between the earliest and latest provinces

[2]Some mixed findings are even found within the same country; Fletcher (2015) revisited the case for the United States and did not find evidence for causality on mortality.

[3]See the Appendix of Lleras-Muney (2005). This could be a reason why the results are not robust when state-specific time trends are added, since they may absorb most of the variations.

[4]The surveys span from 1995 to 2012 and the CSLs started in 1986, so I keep the 1955-1993 birth cohorts and aged between 18 and 50 at the survey to conduct this study.

is only five years in the sample), I further explore the cross-sectional variation in the potential increase in education across the regions. Because *all* the provincial governments were required to enforce the "nine-year" compulsory schooling laws, the years of education in the provinces with more people with less than nine years of schooling before the enforcement of the law should potentially increase more after the law was enforced.[5] The estimates provide sound evidence for this. The CSLs significantly increased the schooling by 1.1 years on average (i.e., 12 percent of the mean value); the effect is 1.6 years in the regions with lower education before (lower than median) but is only 0.6 years for the rest (i.e., 19 percent of the mean value in the lower education regions and 7.1 percent of the mean in the higher education regions). Compared to the exogenous shocks in the previous literature, the effects of the CSLs in China are much larger in magnitude, both in absolute and relative scales.[6]

Since the identification is based on the different timing of the enforcement of the laws and the heterogenous effects across regions, there are some concerns about the identification. First, the potential cohort trends across the provinces caused by other factors, such as heterogeneous economic growth, may drive the estimates. I further control for province-specific birth cohort linear trends, and this yields fairly consistent results. Second, the constructed variables may pick up the effects of other reforms, since China implemented a couple of policies during that period. However, exactly consistent with the "nine-year" compulsory schooling, the results show that the effects of CSLs on education only exist *if and only if* the number of years of schooling is less than or equal to nine. Third, the associations of CSLs with education may reflect the "regression to the mean" rather than the actual effects, because regions with lower education may increase more probably because of lower marginal cost. I conduct a placebo test for the CSL-ineligible cohorts and find no evidence for this. Finally, greater increase in education in the regions probably reflects the larger

[5] In practice, I calculate the proportion of individuals with fewer than 9 years schooling among the CSLs non-eligible cohorts in the local province (the mean value is 0.37 and the value ranges from 0.05 to 0.79 in the sample), and interact it with the CSL-eligibility in the regressions.

[6] For example, Clark and Royer (2013) found that the both education reforms in the UK increased years of schooling by 0.3-0.5 with mean values of years of education around 15-16. Thus, both reforms increased education by 1.9-3.3 percent.

improvement in nutrition, because these regions probably had poorer nutrition status in the beginning. But I find the policy has no effects on height, which is a widely used measure for nutrition status of younger adulthood (Thomas *et al.*, 1991; Deaton, 2003; Huang *et al.*, 2013).

The estimates from the reduced forms and the two-stage least squares (2SLS) both find pronounced effects of education on health outcomes. Specifically, the 2SLS estimates show that one additional year of schooling leads to 2-percentage points decrease in reporting fair/poor health (10 percent of the mean), 1.1-percentage points decrease in the rate of underweight (14 percent of the mean), and 1.3-percentage points decrease in the rate of smoking (5 percent of the mean).

Apart from the remarkable increase in education, another virtue of using the variations in the CSLs in China is that they happened much later (i.e., 1986-1991 in the sample) than the reforms examined in the literature. Thanks to the series of surveys conducted since the 1990s in China, I can use detailed individual information collected in the micro-level data sets to provide some *quantitative* evidence on several candidate mechanisms. For example, income is usually used as an explanation for the impact of education on health because richer people can afford healthier foods since higher education predicts higher income.[7] Another one is that education increases people's cognition, so that they are able to obtain more health knowledge and know how to take care of themselves better. The final one could be the externalities or spillover effects of education. For example, increased education of the population over all by the CSLs would improve the health behaviors in general and generates better sanitary conditions, and thus lead to different health outcomes.

Therefore, I examine the above three mechanisms. The estimates show that income and cognition only explain a small proportion of the effects of CSLs on self-reported health; income explains 7 percent and cognition explains 15 percent. However, the empirical results suggest a more important role of the externalities of education, especially among those with lower education. Among those received no formal education, the empirical estimates also

[7] Higher incomes increase the demand for better health, but they affect health in other ways as well. For example, richer people can also afford more cigarettes; higher wage also means the higher opportunity cost of time: because many health inputs require time (such as exercise or doctor visits or cooking).

suggest a better health among those CSLs-eligible cohorts than that among the CSLs non-eligible cohorts. A conservative calculation suggest the externalities explain over 25 percent of the effects of the CSLs.[8] In addition, the roles of income, cognition, and externalities are different for different health measures. When underweight is the outcome, empirical results suggests a much more important role of income (i.e., income explains 20-30 percent of the effects of CSLs on underweight), but a less important role of spillover effect (i.e., the empirical estimates provide no evidence for this). For the smoking behaviors, however, spillover effect is a more important mechanism, while income and cognition together explain less than 10 percent.

The findings in this paper contribute to several strands of literature. First, the findings provide evidence of the effectiveness of education policies in improving education and health status, and build up the literature by studying causality between education and health for the working-age population in a developing country. Second, the findings about BMI and cognition are consistent with the results in Cutler and Lleras-Muney (2012),[9] Aaronson and Mazumder (2011) and Carlsson *et al.* (2012).[10] Finally, this study fills a gap in the literature by examining the potential mechanisms through which education affects health, which helps to explain the large heterogeneity in the impact of education on health across different nations and in different periods.

---

[8] This is a little bit different from the "peer effects" documented in the literature (e.g., Jensen and Lleras-Muney, 2012). The externalities or spillover effects here emphasize that the people around have higher education caused by the CSLs would improve individual own health even though there is no increased in own education.

[9] First, the findings highlight the effects of education in a developing country: education increases BMI in China because it reduces the underweight rate but has no effects on obesity, while the previous literature (e.g., Brunello *et al.*, 2013) found negative effects of education on BMI because it mostly reduces the obesity rate. The reason may be that the underweight is a more serious health problem in the developing countries like China while obesity matters more for the countries in those developed ones like Europe and US.

[10] The former found that the construction of Rosenwald schools had significant effects on the schooling attainment and cognitive test scores of rural Southern blacks and the latter found that 180 days extra schooling increased cognition test scores by approximately 0.2 standard deviations among the 18-years-olds adolescents in high schools in Sweden. The findings in this paper provide consistent evidence to this.

4.2 Background and Data

4.2.1 Compulsory Schooling Laws in China

China's Compulsory Education Laws were passed on April 12, 1986, and officially went into effect on July 1, 1986. This was the first time that China used a formal law to specify educational policies for the entire country. This law had several important features : 1) nine years of schooling became compulsory; 2) children were generally supposed to start their compulsory education at six years of age in principle, 3) compulsory education was free of charge in principle; 4) it became unlawful to employ children who are in their compulsory schooling years; and 5) local governments were allowed to collect education taxes to finance compulsory education (Fang *et al.*, 2012). Different from the United States and European countries which increased the compulsory schooling by one or two years , the laws in China actually use the uniform "nine years" for the length of years of compulsory schooling no matter where it is.

Local provinces were also allowed to have different effective dates for implementing the law, because the central authorities recognized that not all provinces would be ready to enforce the law immediately. But the variation in the timing is not large, and the gap between the earliest and latest provinces is only 5 years in our sample.[11] Therefore, I further explore the cross-sectional variations in the enforcement of the laws. The central government planned to have different levels of implementation across different regions because of large inequality in education levels across regions, and thus it decided to mainly support the less-developed regions. A government document, "Decisions about the Education System Reform," in 1985 said "the nation will try best to support the less-developed regions to reduce the illiterate rate." One direct consequence is that the CSLs have compressed educational inequality across the nation. For example, the illiterate rate for those over age 15 years in rural areas declined by 25 percentage points, from 37.7 percent in 1982 to 11.6 percent in 2000, while that in urban areas only declined by 12 percentage points,

[11]Note that our sample covered 26 provinces in China. The latest two provinces are Hainan and Tibet, whose CSLs starting year are 1992 and 1994. But these two are not covered in our sample.

from 17.6 percent to 5.2 percent in the same period (Yearbooks Population Survey, 1982 and 2000). Therefore, this study explores both the temporal and geographical variations in the enforcement of the law to identify the effects of education. Sections 3 and 4 provide empirical evidence.

The CSLs in China produced great achievements: the enrollment rate for junior high school increased by 26 percentage points, from 69.5 percent in 1986 to 95.5 percent in 2000, and the number of students enrolled in junior high school increased by 8.9 million. The CSLs made China the first and only country attaining the "nine-year compulsory schooling" goal among the nine largest developing countries.[12]

It was the first time for the largest developing country to enforce such compulsory schooling laws. It would be unrealistic to require those over age 10 years with no formal education but to complete the full nine-year compulsory schooling because they are legal to work at age 16. Those aged 12, for example, are required to go to school to receive education until they are reach age 16 years. They can stop their education legally and go to work because they are no longer age-eligible. Thus, the laws actually defined the age-eligible children as those between ages 6 and 15 years, and required the minimum school-leaving age to be 16 rather than truly "9-year" formal education, at least for the first few cohorts.

4.2.2 Data and Variables

The main sample used in this study is from the Chinese Family Panel Studies (CFPS), Chinese Household Income Project Series (CHIPs), and China Health and Nutrition Survey (CHNS), three ongoing and largest surveys in China. The Appendix D provides a detailed description for each of them. I keep the variables consistently measured across the data sets, if possible: 1) demographic variables: gender, year of birth, *hukou* province (i.e., the province where the household was registered), and type of *hukou* (rural/urban); 2) socioeconomic

[12]The nine countries are China, India, Indonesia, Pakistan, Bangladesh, Mexico, Brazil, Egypt, and Nigeria.

variables: years of schooling and marital status; 3) health and health behavior variables.[13]

Because the CSLs were announced and implemented in 1986, I keep those birth cohorts born after 1955 and earlier than 1993 and surveyed between 1995 and 2011, so that there are almost as many affected as unaffected cohorts in the sample. Furthermore, I restrict the sample to individuals over age 18 years because most of the respondents have completed their education by then. For simplicity, I also drop those over age 50 years because all of them are ineligible to the CSLs and the mortality rate start to increase. I pooled the samples from three data sets together, and the total number of observations is more than 100,000, making it one of the largest micro-level samples to analyze the impact of education on health so far.[14] Table 4.1 reports the mean and standard deviation of the key variables used in the study.

Self-reported health and reported fair/poor health Previous literature suggests that self-reported health is highly predictive of mortality and other objective measures of health (Idler and Benyamini, 1997), and thus this study uses this measure as a major individual health outcome.[15] The measure of self-reported health is based on the answer to the question "How is your health in general?" in the three surveys, with the response ranging from 1 to 5: 1 for excellent, 2 very good, 3 good, 4 fair and, 5 poor. Indicator for reported fair or poor health is equal to one if the answer is 4 or 5, and zero for otherwise. Table 4.1 shows that 19 percent of respondents reported fair or poor health in the sample.

[13]CHNS was collected in nine provinces and almost every two years since 1989: 1989, 1991, 1993, 1995, 1997, 2000, 2004, 2006, 2009, and 2011. The CHIPs and CFPS data are sampled nationwide. But the CHIPs data used here include those collected in 1995, 2002, 2007, and 2008; the CFPS data here are those surveyed in 2010 and 2012. More details can be found in the Appendix D.

[14]Since the three different datasets were collected in different years and different provinces, I allow the systematic differences across the different datasets by including dummies for the province, survey year, data sources and all the possible interactions between the three.

[15]Although individual mortality is a more accurate and objective measure for health and has been widely used in previous literature, the sample here is much younger than those examined in previous literature, and the mortality rate for this age group is too low.

Table 4.1: *Summary Statistics*

Variables	(1) Obs.	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
Panel A: Health and Health Behaviors					
Health Fair or Poor	88,971	0.19	0.39	0	1
Health Excellent	88,971	0.28	0.45	0	1
BMI	85,275	22.5	3.18	12.1	50
Underweight	85,275	0.08	0.27	0	1
Obese	85,275	0.02	0.15	0	1
Smoke	105,634	0.26	0.44	0	1
Panel B: Education and Demographics					
Years of schooling	114,647	8.86	3.91	0	23
Male	114,647	0.50	0.50	0	1
Age	114,647	32.5	9.16	18	50
Urban	114,647	0.39	0.49	0	1
Married	114,647	0.54	0.55	0	1

Notes: Data source is CFPS, CHIPs and CHNS. The variables are measured consistently across the data sets. The sample is composed of the 1955-1993 birth cohorts, aged between 18 and 50, and surveyed between 1995 and 2011.

BMI, underweight and obesity BMI is also a widely used variable in the literature to depict the individuals' nutritional situation and has shown to be correlated with mortality and economic growth (Fogel, 1994; Cutler *et al.*, 2003). All three surveys provide the information needed for calculating BMI,[16] and I define underweight status as BMI being less than 18.5 and obesity as BMI greater than 30. Table 4.1 reports that the underweight rate is 8 percent and the obesity rate is only 2 percent,[17] indicating that the obesity problem seems not to be a big issue compared to the popular obesity in the developed areas like the United States and Europe.

Smoking Because of the high smoking rate in China and the close relationship between smoking and mortality (Wasserman *et al.*, 1991; Cutler and Lleras-Muney, 2010), this study also examines the effects of education on smoking. In most of the surveys, respondents were asked "Do you smoke now?" or "Did you smoke last week?" I then code the respondents as current smokers, which equals one if the answer to these questions is "yes," and zero if otherwise. The smoking rate is 26 percent for the full population and most of the smokers are men, whose smoking rate is higher than 50 percent, almost three times of that in the United States.

Cognitive abilities Cognition refers to mental processes that involve several dimensions, including the thinking part of cognition, which includes memory, abstract reasoning, and executive function, and the knowing part, which is the accumulation of influence from education and experience (Hanushek and Woessmann, 2008; Hanushek, 2013). The CFPS measured cognitive abilities by two sets of tests. For the words recall test, interviewers read a list of 10 nouns, and respondents were asked immediately to recall as many of the nouns

[16]Height and weight are reported by respondents themselves in CHIPS and CFPS but are measured by professional nurses in CHNS. This study simply takes the BMI derived from the reported variables and that from measured variables equally. In our regressions, we controlled for the indicators for calendar year, data source and hukou provinces and all of their interactions to capture any possible systematic bias. I also drop those BMI with values being smaller than 10 or larger than 50 (less than 1 percent of the sample) because these outliers are mostly due to falsely reporting

[17]In the sample, 12 percent of the women are underweight, although this is not reported in this table.

as they could in any order. The test would stop if the respondents continuously mentioned three nouns that were not in the list. The other test is about mathematical calculation ability: the respondents were asked to answer 8 or 10 math calculation questions and the test would also terminate if the respondents answered three questions in a row incorrectly. Because of different number of questions are used in the different survey years, I calculate the proportion of correct answers for each test and use the Z-score in each year as the cognition measures.

Demographics and education The basic demographic variables, such as education, gender, type of *hukou* (urban/rural), and year of birth (or age) are consistently collected in the surveys. For all the surveys, information on years of schooling is provided. Panel B of Table 4.1 reports the basic statistics for these variables; the people in the sample are age 30 years on average, and 33 percent of them lived in urban areas.

4.3 Graphical Analysis

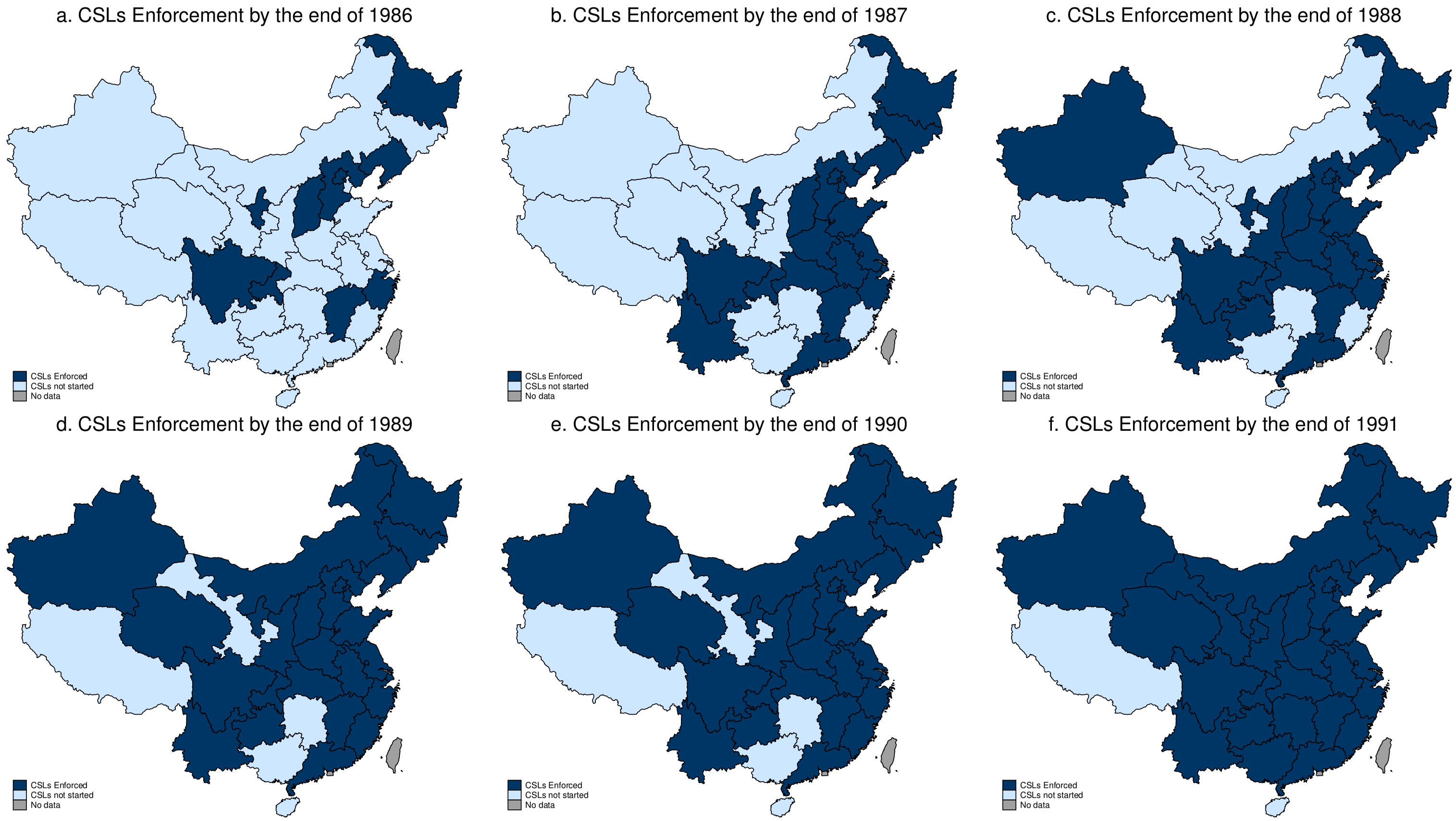
Because the central government allowed the provincial governments to implement the policy separately, I collected the formal official documents in each province and report the initial year in which the CSLs were effective in each province in column 1 of Table 4.2, and report the first cohort affected in column 2.[18] Figures 4.1 a-f graphically show the CSLs enforcement across different provinces over time. Almost all the provinces enforced CSLs within the 1986-1991 period.[19]

An important feature of CSLs in China is the *uniform* nine-years compulsory schooling. I thus hypothesize that the increase in years of education in provinces with lower education

[18]The timing of the CSLs, as shown in Table 4.2, is weakly correlated with the education level of each province (correlation coefficient = 0.2). Regressing the year when the law became effective on the education level prior to the CSLs yields an insignificant (p-value = 0.27) though positive coefficient. In further analysis, this study also allows the provinces to determine endogenously when to start the CSLs, finding the results are also consistent. The results are available upon request.

[19]There are only two provinces in mainland China which did not start the CSLs in 1991, Hainan and Tibet. These two provinces are not surveyed in the three data sets.

Figure 4.1: *CSLs Enforcement in Different Provinces over Time*



Notes: Data source is the education year books for each province. Every figure shows the CSLs enforcement across China at the end of each corresponding year. Two regions not starting CSLs in 1991 are Hainan and Tibet, which are not included in the sample. The data on Taiwan are missing.

prior to the CSLs be greater after the CSLs enforcement. So I first calculate the proportion of those with fewer than nine years education in the birth cohorts prior to the CSLs (within 15 years) in each province, as reported in column 3. It ranges from 0.05 for Beijing to 0.79 for Fujian and has a large variation, suggesting a large regional inequality in education in China before the enforcement of the CSLs. Figure 4.2a plots the values geographically.

I divide the provinces by the median value of column 3 into high-education provinces and low-education ones. Then I regress the schooling years on the dummies of different birth cohorts relative to the CSLs eligibility for each group, controlling for gender, *hukou* province, survey year, sample source (CHNS/CFPS/CHIPS) and all of their interactions. The the reference group is the *just-eligible* cohort (i.e., the birth cohorts aged 15 the CSLs became effective in the local province). Figure 4.2b reports the point estimates and the corresponding confidence intervals for each birth cohort (i.e., from those born 4 years earlier than the reference cohort to those born 14 years later than the reference cohort). These birth cohorts cover those totally *non-eligible* ones (i.e., age sixteen years or older when CSLs enforcement), those *partially-eligible* ones (i.e., age between seven and fifteen years when CSLs enforcement), and those *fully-eligible* ones (i.e., age six years or younger when CSLs enforcement). Initially, there is more years of schooling among those non-eligible cohorts in higher-education regions. However, the difference is much narrowed among the partially-eligible cohorts, and is even reversed among the fully-eligible cohorts. The years of schooling in the low-education provinces increased about 1.6 on average, while that in in the high-education provinces only increased about 0.7.

Figure 4.2c reports the results of parallel analysis when the dependent variable is self-reported health (i.e., the value ranges from 1 to 5, and the higher value indicates unhealthier status). The figure shows that the relative levels and cohort trends in self-reported health (compared to the reference group in each sample) among non-eligible cohorts are similar in the two groups; however, self-reported health improved more from the non-eligible cohorts to the fully-eligible cohorts in the regions with lower education prior to the CSLs enforcement. Therefore, Figure 4.2b and 4.2c together provide some evidence for the causal

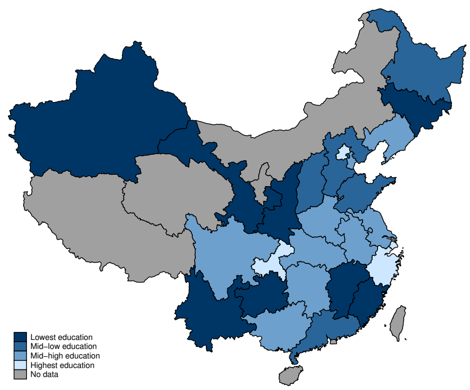
Table 4.2: *Compulsory Schooling Laws by Province*

Province	Law effective year	First affected birth cohort	Prop of earlier cohorts with less 9-years education
Beijing	1986	1971	0.053
Tianjin	1987	1972	0.285
Hebei	1986	1971	0.401
Shanxi	1986	1971	0.394
Liaoning	1986	1971	0.352
Jilin	1987	1972	0.487
Heilongjiang	1986	1971	0.385
Shanghai	1987	1972	0.220
Jiangsu	1987	1972	0.306
Zhejiang	1986	1971	0.249
Anhui	1987	1972	0.302
Fujian	1989	1974	0.790
Jiangxi	1986	1971	0.672
Shandong	1987	1972	0.392
Henan	1987	1972	0.358
Hubei	1987	1972	0.288
Hunan	1991	1976	0.357
Guangdong	1987	1972	0.382
Guangxi	1991	1976	0.381
Chongqing	1986	1971	0.226
Sichuan	1986	1971	0.318
Guizhou	1988	1973	0.475
Yunnan	1987	1972	0.499
Shaanxi	1988	1973	0.409
Gansu	1991	1976	0.577
Xinjiang	1988	1973	0.581

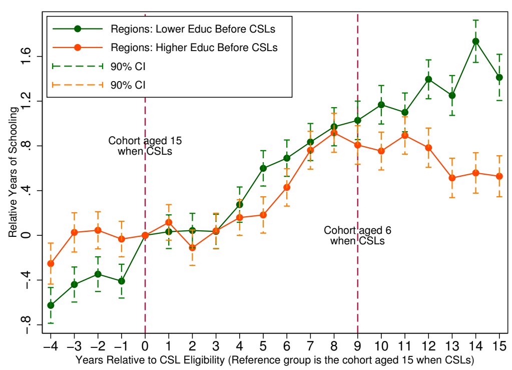
Notes: Data are from the education yearbooks for each province.

Figure 4.2: *Lower Prior Education, More Improvement in Education and Health after CSLs*

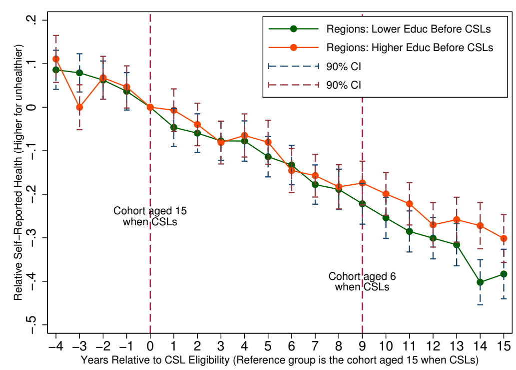
(a) *Geographical Distribution of Education Levels before the Laws*



(b) *Increased Education over Birth Cohorts, by Local Education Level among Earlier Cohorts*



(c) *Improvement in Health over Birth Cohorts, by Local Education Level among Earlier Cohorts*



Note: Data source is CFPS, CHIPs and CHNS. Figure 4.2a categorizes the values in column 3 of Table 4.2 into four groups and plotted them geographically. For Figures 4.2b and 4.2c, I divide the sample by the median value of the proportion of people with less than 9-year education prior to the CSLs, then conduct regressions to estimate how the years of schooling or self-reported health change over birth cohorts relative to CSLs eligibility for each subsample, controlling for gender and dummies for *hukou* province, survey year, sample and all of their interactions. The reference group is the *just-eligible* cohort for the CSLs for each subsample.

effects of education on self-reported health. The following sections further provide further evidence by conducting regression analysis.

4.4 First Stage: Impact of CSLs on Education

4.4.1 Econometric Methodology

I estimate the following equation to to test the hypothesis formally:

$$Edu_{ijbt} = \alpha_0 + \alpha_1 Eligible_{bj} + \alpha_2 prop_j^{prior<9} \times Eligible_{bj} + \alpha X_{ijbt} + \delta_{sjt} + \epsilon_{it} \tag{4.1}$$

The subscripts i, j, b, and t denote the individual i, province j, birth cohort b, and survey year t, respectively. The dependent variable Edu_{ijbt} denotes years of schooling of individual i, and $Eligible_{bj}$ denotes the CSL-eligibility for birth cohort b in province j, which equals one if the individual is fully-eligible for the CSLs and equals zero if the individual is non-eligible. Then I assume the eligibility follows a linear function in between ages six and sixteen years. The results do not rely on the linear-function assumption. I also used a step function (i.e., every three years or five years) and find consistent results.

One potential issue here is that the *hukou* province may be not the province where they received education. But this may not be a first-order issue driving the results: the proportion of individuals whose *hukou* province is the same with their birth province is more than 93 percent for the same cohorts, according to the author's calculation based on the 2005 census.

X_{ijbt} denotes a set of control variables, including dummies for gender, type of *hukou* (urban/rural), married status (married or not), age, and year of birth. δ_{sjt} denotes a set of dummies, including data sample s (CHNS/CFPS/CHIPs), province j, and survey year t and all of three interactions. Adding δ_{sjt} into the equation controls for not only the potential systematic difference existing across data sets but also the different contemporaneous conditions in each province.

$prop_j^{prior<9}$ denotes the proportion of people with fewer than nine years schooling in the population born prior to the CSLs in province j (i.e., the value in column 3 in Table

4.2). Since the proportion varies at the province level, the main effect would be absorbed by the province dummies. The coefficients of eligibility (α_1) and the interaction (α_2) are of main interest because they capture the main effect of the CSLs, and the differential increase in education after the CSLs between the provinces with lower and higher prior education. In practice, I interact the CSL-eligibility with the *demeaned value* of $prop_j^{prior<9}$. Thus the coefficient on eligibility (α_1) can be interpreted as the impact of CSLs on education at the mean level of prior education, which is expected to be positive. I also expect $\alpha_2 > 0$, which suggests those with lower education prior to the CSLs will have a greater increase in years of education after the enforcement of CSLs.

4.4.2 Empirical Results

Table 4.3 reports the OLS estimation for α_1 and α_2, with the standard errors clustered at the province-year of birth level. Column 1 presents the results without the interaction term, showing that CSLs increase the years of schooling by 1.1 years on average. The estimates in column 2 show that $\alpha_1 > 0$ and $\alpha_2 > 0$, and both of them are significant. The magnitude of the coefficient suggests that the policy-induced increase in years of education in regions with lower education before the CSLs (e.g. Fujian, Jiangxi and Gansu) would be 1.5 years more than the regions with higher education before the CSLs (e.g., Beijing, Tianjin, and Shanghai).

One potential issue is that time trends across the different regions, caused by other factors like economic growth, may drive the estimation. This issue is also relevant to Stephens and Yang (2014), who found the results in previous literature become insignificant and wrong-signed when region-specific linear trends are included. I thus control for province-specific birth cohort linear trends in column 3. The estimates show that the impact of the CSLs is robust to including these, suggesting that the other birth cohort linear trends across different regions should not be the first-order factors.

Table 4.3: *OLS Estimation for Impact of Compulsory Schooling Laws on Years of Schooling*

	(1)	(2)	(3)
Variables	Dependent variable is Years of Schooling		
CSLs Eligibility	1.116***	1.136***	1.242***
	(0.381)	(0.360)	(0.382)
Pr(less than 9-year education)		4.065***	6.124***
* CSLs Eligibility		(0.646)	(1.445)
Observations	114,647	114,647	114,647
R-squared	0.243	0.245	0.249
F-statistic for all the variables	8.572	23.25	16.19
P-value for the F-test	0.003	0.000	0.000
Province-YoB Linear Trends	No	No	Yes

Notes: Data source is CFPS, CHIPs and CHNS. Robust standard errors in parentheses are clustered at the province-year of birth level. Covariates include indicators of type of *hukou* (urban/rural), year of birth, age (three-year categories), *hukou* province, survey year, and all interactions of province, year, and sample. The Pr(less than 9-year education) variables are de-meaned value so that the coefficient on CSLs Eligibility can be interpreted as the impact where the Pr(less than 9-year education) has the mean value.

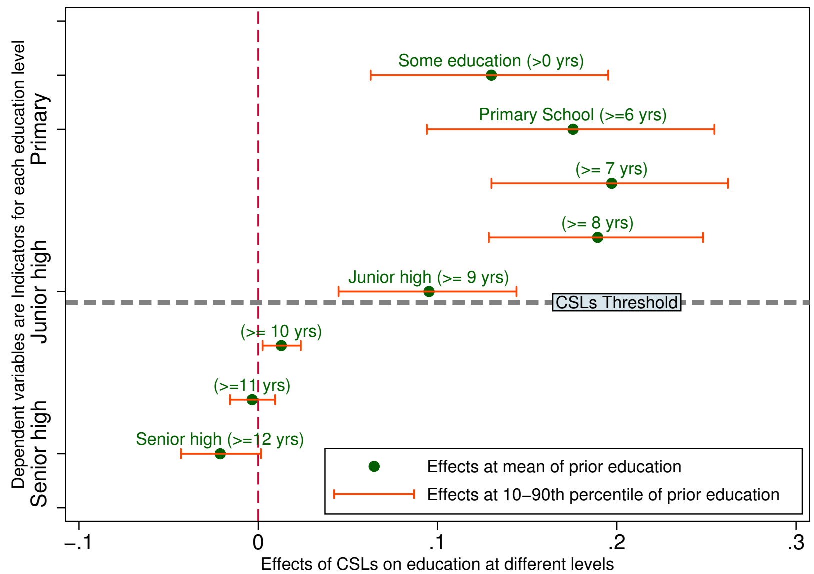
4.4.3 Evidence of Exogeneity of the CSLs

Other Confounding Factors or Other Policies?

Comparison between before and after CSLs across the provinces captures the differential increase in years of education across the regions. However, the timing of the CSLs and the interaction may pick up variations in other policies, because China experienced a series of different reforms in the 1980s. But it seems to be unrealistic to list all contemporaneous policies in different regions during that period and test their correlation with the timing and enforcement of the CSLs. Instead, I directly test to what extent that CSLs increased the years of schooling . The education reform requires nine years of compulsory schooling for all the provinces. Therefore, the constructed variables based on the CSLs may increase the years of education up to and only up to nine years. However, there is no evidence that other confounding factors, such as local opinions toward education or other policies, would increase the years of schooling only up to nine years.

To test this, I construct a set of indicators for different years of schooling, use these indicators as dependent variables, and conduct the regressions as in equation 4.1. Because the effects of CSLs are depicted by the coefficients α_1 and α_2 together, I use the estimated coefficient in each regression to calculate the impact of CSLs on education at the mean level of prior education, and those at 10th and 90th percentile level of prior education. The points in Figure 4.3 reports the impact of CSLs on education when the prior education equals to the mean value of all the provinces. For each dependent variable, left end of the interval is the effect of CSLs when prior education is at the 10th percentile; while the right end indicates that when prior education is at the 90th percentile. The wider the intervals are, the larger heterogenous effects of CSLs have across the regions. When the years of schooling do not exceed the threshold of nine, the points are obviously positive and the range is wide. Once the years of schooling are greater than, however, the impact of the policy diminished dramatically both for the main effects (the points are much closer to zero and are not significant) and the heterogeneous effects across regions (the intervals are much narrower). These findings suggest that the positive association between education and the

Figure 4.3: *Impact of CSLs on Years of Schooling at Different Education Levels*



Notes: Data source is CFPS, CHIPs and CHNS. Each row reports a specific OLS estimation when the dependent variable is the indicator for completing the corresponding years of schooling (as marked). The independent variables are described in equation 4.1. The points in the figure report the coefficients on CSLs-eligibility and the intervals show the impact from the 10th to 90th percentile of the prior education level based on the OLS estimates.

constructed variables in Table 4.3 should originate from the CSLs rather than from other unobserved factors like the implementation of other policies or reforms.

"Regression to the Mean" and Nutrition Status?

I also conduct two sets of placebo tests to provide further evidence on excludability of the constructed CSL variables. The first set aims to test whether the impact or associations in Table 4.3 are only "regression to the mean." First, I restrict the sample to those cohorts earlier than the first affected cohort (i.e. the cohorts 2-15 years earlier than the first affected cohort). And then I suppose the year of implementation of the CSLs was five years earlier, estimate the same regressions as equation 4.1, and report the results in the first two columns in Table 4.4. The results provide no evidence that pre-trends or regressions to the mean matter much in this analysis.

Table 4.4: *Placebo Tests for Impacts of Compulsory Schooling Laws*

	(1)	(2)	(3)	(4)
Settings	CSLs ineligible (2-15 years earlier) and suppose CSLs 5 years before		Use Height as Dep. Var.	
Dependent variable	Years of Schooling		Height (cm)	
CSLs Eligibility	0.266	0.257	0.466	0.463
	(0.622)	(0.617)	(0.447)	(0.448)
Pr(less than 9-year education)		1.415		-0.353
* CSLs Eligibility		(0.940)		(0.570)
Observations	39,511	39,510	87,137	87,137
R-squared	0.305	0.305	0.546	0.546
F-statistic for all the variables	0.183	1.185	1.086	0.728
P-value for the F-tests	0.669	0.306	0.298	0.483

Notes: Data source is CFPS, CHIPs and CHNS. Robust standard errors in parentheses are clustered at the province-year of birth level. Covariates and variable definitions are the same as those in Table 4.3.

The second set of placebo tests are conducted to test whether the impacts of the CSLs reflect better nutrition in individuals in childhood or young adulthood. I use height as an independent variable since height is proved to be a good measure for health and nutritional status in childhood and young adulthood (Thomas *et al.*, 1991; Deaton, 2003; Currie and Vogl, 2013). If the impact of the CSLs reflects the improvement in nutrition , the effects should be captured in height. The estimates in the last two columns of Table 4.4 provide no evidence of this.

4.5 Effects of Education on Health

4.5.1 Baseline Results

I begin the analysis by first conducting the OLS estimation for following equation as a benchmark:

$$Health_i = \theta_0 + \theta_1 Edu_i + \theta X_i + \delta_{sjt} + \epsilon_i \tag{4.2}$$

the dependent variable, $Health_i$, denotes the health outcome variables, which may be self-reported health, underweight, smoking, or cognition, and all the other variables are the same as those in equation 4.1. Panel A in Table 4.5 reports the OLS estimates of θ_1, showing that higher education is correlated with better health in general. The sample size varies across columns because some surveys may not collect the corresponding health information. For example, the cognition tests (i.e., words recall and math calculation) are only collected by CFPS.

Following previous literature, I conduct 2SLS estimation:

$$Health_i = \beta_0 + \beta_1 \widehat{Edu}_i + \beta X_i + \delta_{sjt} + \varepsilon_i \tag{4.3}$$

$\widehat{Edu}_i$ is the predicted education value of equation 4.1 and all the other variables are the same as those in equation 4.1. Panel B presents the results. Because of the different samples, the F-tests in the first stage (i.e., weak instrumental variable tests) and Hansen tests

Table 4.5: *Effects of Education on Health*

Dependent variables	(1) Health Fair or Poor (Yes = 1)	(2) Underweight (Yes = 1)	(3) Smoker (Yes = 1)	(4) Words recall Z-score	(5) Math Ability Z-Score
Mean of Dependent Var.	0.190	0.077	0.264	0.000	0.000
Panel A. OLS Estimation					
Years of Schooling	-0.00761***	0.000155	-0.00389***	0.107***	0.0834***
	(0.000448)	(0.000325)	(0.000465)	(0.00142)	(0.000843)
Observations	88,971	85,275	105,634	34,999	34,985
R-squared	0.095	0.053	0.356	0.382	0.809
Panel B. 2SLS Estimation					
Years of Schooling	-0.0205***	-0.0115*	-0.0134*	0.158***	0.0694***
	(0.00642)	(0.00636)	(0.00723)	(0.0265)	(0.0114)
Observations	88,971	85,275	105,634	34,999	34,985
First Stage F-statistics	26.87	27.67	25.78	12.15	12.20
Over-identification P-values	0.125	0.263	0.004	0.06	0.435
Panel C. Reduced Form Estimation					
CSLs Eligibility	-0.0628***	-0.00282	-0.0713***	0.320***	0.150***
	(0.0217)	(0.0174)	(0.0208)	(0.0808)	(0.0496)
Pr(less than 9-year education) * Eligibility	-0.0759**	-0.0693**	-0.0123	0.335***	0.103
	(0.0328)	(0.0311)	(0.0358)	(0.111)	(0.0839)
Observations	88,971	85,275	105,634	34,999	34,985
R-squared	0.090	0.053	0.355	0.185	0.684

Notes: Data source is CFPS, CHIPs and CHNS. Robust standard errors in parentheses are clustered at the province-year of birth level. Covariates and variable definitions are the same as those in Table 4.3.

(over-identification Tests) for the instruments are reported at the bottom of each column. The large F-statistics reject the null hypothesis and provide evidence of a significant first stage for all the columns. This study did not report the detailed first stage for different outcomes, but the results are available upon request. In general, the instruments also passed the over-identification tests, except for smoking.

The 2SLS estimates are about three times larger in general. It is possible that the effects among the compliers (i.e., those with increased education under the CSLs and not without the laws) are larger because the effects identified from the 2SLS are local average treatment effects (LATE).[20] In addition, the OLS estimates may be biased to zero because of the classic measurement error in education, because the values were reported by the respondents themselves, and these reported values may be inaccurate.

The first column in Table 4.5 provides estimates for self-reported fair or poor health, indicating that an additional one year of schooling decreases the probability of reporting fair/poor health by 2 percentage points.[21] Since there were 19 percent of individuals in the sample reporting fair/poor health, the 2SLS estimates suggest one additional year of schooling reduce the reporting fair/poor health by 10 percent. Column 2 in Panel C shows that an additional year of schooling leads to a drop of about 1.2 percentage points in the underweight rate (14 percent of the mean), suggesting that education improves nutritional status.[22] Column 3 shows the effects of education on smoking. Consistent with the findings in Jensen and Lleras-Muney (2012), the 2SLS estimates suggest that an additional year of

[20]The associations in the lower education group (less than nine years) tend to reflect the impact of education among the "complier" group, since previous analysis shows the CSLs are mainly effective in the lower education group. Hence, I divide the whole sample by whether the individuals completed nine years of education and conduct OLS estimation to investigate the associations of education with the health outcomes for each group. In general, the results are consistent with the hypothesis.

[21]Considering the CHNS used a four-point scale and the other two used a five-point scale, I drop the CHNS sample and re-estimate the effects of schooling, which only yields very consistent results.

[22]However, the results are different from the findings in developed regions like the United States and Europe. Both Kemptner *et al.* (2011) and Brunello *et al.* (2013) found that education has a negative effect of education on BMI. Although not reported, I further find that education increases BMI but the effects only exist in the sample with lower BMI, and do not provide evidence that education increases the rate of obesity in China. These findings suggest that schooling increases BMI in developing countries through decreasing the underweight proportion but decreases BMI in developed countries via reducing the obesity rate. This finding is consistent with Cutler and Lleras-Muney (2012).

schooling reduces the likelihood of smoking by 1.3 percentage points (5 percent of the mean). The last two columns examine cognition. The estimates in the last two columns in Table 3 suggest that an additional year of schooling increases cognition by 0.09 standard deviation for word recall and 0.16 for math calculation.[23]

Panel C shows the reduced form results, whereas education is replaced by the constructed CSLs variables (i.e. $Eligible_{bj}$ and $prop_j^{prior<9} \times Eligible_{bj}$) directly:

$$Health_i = \lambda_0 + \lambda_1 Eligible_{bj} + \lambda_2 prop_j^{prior<9} \times Eligible_{bj} + \lambda X_i + \delta_{sjt} + \epsilon_i \tag{4.4}$$

Since both $Eligible_{bj}$ and $prop_j^{prior<9} \times Eligible_{bj}$ predict higher education, the signs of the coefficients in the reduced form estimations should be negative for poor health and positive for better health. The estimates in Panel B provide consistent evidence of this.

The difference between the reduced form and 2SLS estimates is noteworthy. The 2SLS estimates are based on the exogeneity of the CSLs and estimates the effects of education on health among compliers. However, the 2SLS estimates do not consider the spillover effects or externalities of education. The reduced form estimates, however, estimate the effects of CSLs implementation on health outcomes *directly*, and thus the effects of individual education and effects of the *average* education of the population are mixed together.

4.5.2 Robustness Checks

Considering that health and behaviors may be different in men and women because of biological and cultural reasons, I conduct gender-specific 2SLS estimation, and then report the results in Figure 4.4, respectively.[24] In general, the results provide evidence for the effects of CSLs or education on self-reported health and cognition for both genders. But the effects on underweight are significant only for women and those on smoking are significant

[23]These findings are consistent with Carlsson *et al.* (2012), who found that 180 days extra schooling increased crystallized test scores by approximately 0.2 standard deviation among 18-year-olds adolescents in high schools in Sweden. The findings are also consistent with Aaronson and Mazumder (2011), who found that the construction of Rosenwald schools had significant effects on the schooling attainment and cognitive test scores of rural Southern blacks in the United States.

[24]Reduced form yields very consistent results.

only for men. It makes sense in China because women has a much higher underweight rate (the underweight is 12 percent for women but is less than 3 percent for men) while men has a much higher smoking rate (the smoking rate for men is over 50 percent but for women is less than 3 percent).

Since the CHNS was collected from nine provinces and combining the three samples together might put disproportionate weights on these provinces, I weight the regressions by the population of the province divided by the number of observations, and it yields very consistent estimates.[25] It should be noted that the results are also robust when including province specific linear trends.

4.6 Understanding the Effects of Education on Health

As suggested in Cutler and Lleras-Muney (2012), studies of the effect of education on health will need to understand the pathways that link the two because this would improve our understanding of the education-health link substantially. On one hand, the evidence on mechanisms is somewhat weaker than the evidence on causality, since researchers often have to make assumptions about what constitutes a mechanism, which partly due to the data limitation. On the other hand, the mixed findings in the literature call for studies to investigate the mechanisms through which education affects health. This section aims to shed some light on this issue.

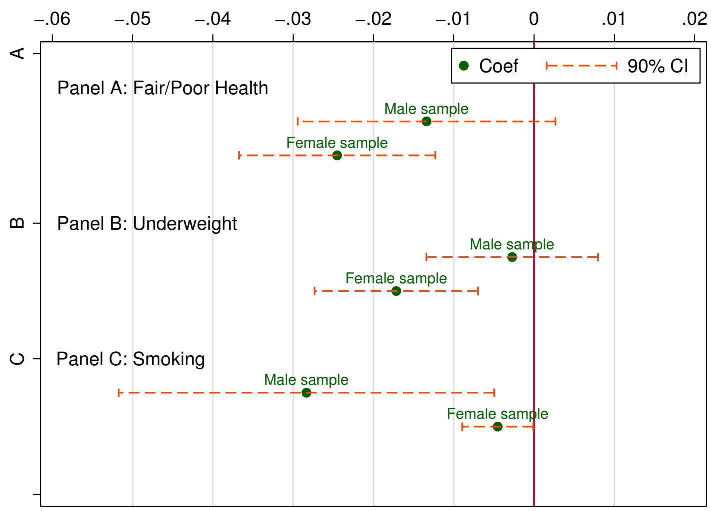
Theoretical foundations for a causal effect of education on health were first provided by the seminal work of Grossman (1972). Current studies such as Cutler and Lleras-Muney (2012) provide some potential mechanism candidates.[26] Due to data limitation, this study examines three possible pathways, including income, cognition and spillover effects. The

[25] Although the results are not reported here.

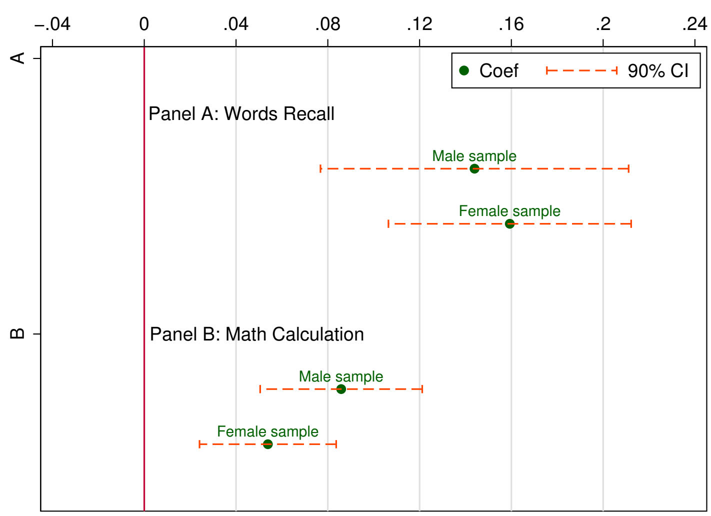
[26] Cutler and Lleras-Muney (2012) classified the pathways of the effect of education on health into four categories. First one is labor market outcomes since higher education yields higher income and safer occupation etc. Second one is the "technology" parameter, such as better use of information. Third one is that education could change the 'taste' for a longer, healthier life, (i.e., the utility function could be changed). Final one is peer effects, which means that people with higher education would be more connected to those with higher education and thus are more likely to develop better health behaviors and have better health.

Figure 4.4: *Effects of Education on Health, by Gender*

(a) Effect of Education on Fair/Poor Health, Underweight, and Smoking



(b) Effect of Education on Cognition



Note: Data source is CFPS, CHIPs and CHNS. Gender-specific 2SLS estimation (equation 4.2) is conducted for each outcome. The points show the coefficients on years of schooling in the 2SLS estimation and the intervals are the 90 percent confidence intervals based on standard errors clustered at the province-year of birth level.

first two are intermediate variables at individual level. Since higher education predicts higher income, this allows people with higher education can have a higher quality, such as living in a house in a safer region and with better environment or having less financial pressure, etc. Higher cognition induced by higher education, as shown above, helps people to get useful information more efficiently and make wiser and more rational choices like choosing proper food, taking drugs in the right way if necessary, evaluating the potential risks in life, and avoiding the potential danger, etc. I also investigate the spillover effects or externalities of education (Borjas, 1995; Ludwig *et al.*, 2012; Wantchekon *et al.*, 2015).[27] For example, increase in education could decrease the smoking rate overall, which would in turn increase the indoor air quality and improve sanitary conditions. In addition, it is also possible that those without any formal education may follow the others with higher education, and they are likely to get more useful suggestions when asking other people around.[28]

4.6.1 Income and Cognition as Mechanisms

To quantify the possible mechanisms, I follow Cutler and Lleras-Muney (2010) and estimate the following two equations:

$$Health_i = \gamma_0 + \gamma_1 Eligible_{bj} + \gamma X_i + \delta_{sjt} + \epsilon_i$$

$$Health_i = \gamma_0' + \gamma_1' Eligible_{bj} + \gamma' X_i + Z_i + \delta_{sjt} + \epsilon_i$$

[27]However, the literature does not reach a consensus about the peer effect or the externalities of human capital, which partly depends on what the outcome is. For example, Borjas (1995) found the average skills of the ethnic group in the parent's generation had some effects on the individual skills; Ludwig *et al.* (2012) found moving to a better neighborhood leads to long-term (10- to 15-year) improvements in adult physical and mental health and subjective well-being. However, Ciccone and Peri (2006) and Acemoglu and Angrist (2001) do not find evidence for externalities for human capital on individual return.

[28]It should be noted that the spillover or externalities here are similar to the "peer effects" documented in the literature such as Jensen and Lleras-Muney (2012) because both of them refer to the effects from people around. But they are different: the peer effects of education usually mean that people with higher education would be more connected to those with higher education and thus are more likely to develop better health behaviors and have better health. But the externalities or spillover effects here emphasize that the people around have higher education caused by the CSLs would improve individual own health even though there is no increased in own education.

the dependent variable $Health_i$ is the main health outcome, which can be reported fair/poor health, underweight and smoking. All the other variables have the same definition as those in equation 4.2. I only use $Eligible_{bj}$ directly here because it captures the average effects of CSLs on the health and thus include both the direct effect of increased own education and the indirect effect of increased education of others in the local region. The estimated effects of CSLs have taken into account of the potential spillover effects. Z_i denotes the potential intermediate variables (i.e., income, cognition or both). Following the methodology in Cutler and Lleras-Muney (2010), I interrupt the change in the magnitude of coefficient on $Eligible_{bj}$ as the part that can be explained by the intermediate variable Z_i (i.e., the explained proportion equals $1 - |\frac{\gamma_1'}{\gamma_1}|$).

Panel A in Table 4.6 reports the results for the proportions explained by the possible intermediate variables when the dependent variable is self-reported fair/poor health. I conduct the analysis by gender with consideration that the effects may differ in between; since only CFPS data measure cognition, I also conduct a parallel analysis for the full and CFPS samples separately. Column 1 reports the original effects of the CSLs. Column 2 reports the conditional effects when income is controlled for and column 3 reports the corresponding proportion that can be explained by income.[29] The part that can be explained by income is 9.9 percent for men and 3.6 percent for women in the full sample, and 7.1 percent for men and 1.2 percent for women in the CFPS sample. One possible reason why the estimates with the CFPS data are smaller is the survey years of the CFPS data are 2010 and 2012, the latest two years in the full sample, when the households and individuals had higher income in general. In addition, the part can be explained by income is consistently larger for men for both samples.

Consistent results of two samples in the first few columns suggest the feasibility of using CFPS data to calculate the part explained by cognition. Column 6 reports the *conditional* effects when only cognition measured by word recall and math calculation is controlled for, and column 7 reports the reduction of magnitude in percent. The proportion that can be

[29]Income here includes both individual income and household income.

Table 4.6: *The Role of Income and Cognition in Effects of Education on Health outcomes*

Sample	(1) Original	(2) Income controlled	(3)	(4) Cognition controlled	(5)	(6) Both controlled for	(7)
	ave. effect	Ave. effect	Explained (%)	Ave. effect	Explained (%)	Ave. effect	Explained (%)
Panel A: Reported Fair/Poor Health							
Both genders in full sample	-0.061	-0.057	6.06				
Men in full sample	-0.048	-0.043	9.87				
Women in full sample	-0.074	-0.072	3.59				
Both genders in CFPS	-0.057	-0.055	3.93	-0.048	16.4	-0.048	16.4
Men in CFPS	-0.068	-0.063	7.10	-0.059	12.5	-0.057	15.8
Women in CFPS	-0.049	-0.049	1.22	-0.038	23.0	-0.039	20.1
Panel B: Underweight							
Women in full sample	-0.011	-0.007	30.7				
Women in CFPS	-0.018	-0.014	20.0	-0.017	7.49	-0.013	25.2
Panel C: Smoking							
Men in full sample	-0.070	-0.066	5.13				
Men in CFPS	-0.199	-0.198	0.86	-0.186	6.67	-0.185	6.96

Notes: Data source is CFPS, CHIPs and CHNS. The corresponding explained proportion is $1 - |\frac{\gamma_1'}{\gamma_1}|$. Because the effects of CSLs on underweight and smoking are only identified among women and men, respectively, this table only examines the corresponding subsample.

explained by cognition is 12.5 percent for men to 23.0 percent for women, implying that cognition is a more important channel among women. In addition, the part that can be explained by cognition is larger than that by income, suggesting that cognition is the most important intermediate variable examined here. These findings are also consistent with the literature that highlights the importance of cognition (e.g., Hanushek and Woessmann, 2008; Aaronson and Mazumder, 2011; Carlsson *et al.*, 2012).

Panel B and Panel C reports the results for underweight among women and smoking among men, respectively. I only keep men or women for these specific outcomes because of no significant effect of CSLs on underweight among men and on smoking among women as shown in Figure A1. The results show that income is an important mechanism to explain the effects of education on underweight since it explains 20-30 percent. But cognition is not since it only explains 7 percent. For the smoking behaviors among men, both income and cognition only explain a small proportion.

4.6.2 Spillover Effects or Externality of Education on Health

The above analysis suggests a small proportion of the effects of education on self-reported health and smoking that can be explained by the individual intermediates such as income and cognition. For self-reported health, around 80 percent of the effects cannot be explained. The natural question is what is the most important factor that may explain the effects of education. As mentioned above, the potential spillover effects may be an important candidate. To provide some evidence of the externalities, I first use the sample composed of those with all education levels, and conduct a reduced form estimation to quantify the effect of CSLs-eligibility on self-reported health in Panel A. The estimates in all the columns show that CSL-eligibility improves health.

Then I restrict to the sample to those without any formal education to conduct the same regression in Panel B.[30] Because the education is unchanged for those receiving no

[30] Age-eligible children may not go to school due to several reasons. First, primary schools in the local regions may not have been built up yet because it takes time to catch up. Second, in some remote villages, the children may not go to school and the punishment of the laws cannot be enforced because the administrative

Table 4.7: *Spillover effects of CSLs on Self-reported Health, Underweight and Smoking*

	(1)	(2)	(3)	(4)	(5)
Dependent variable	Self-reported fair/poor health (Yes = 1)			Underweight (Yes = 1)	Smoking (Yes = 1)
Samples	Full	Male	Female	Female	Male
Panel A: Full sample					
CSLs Eligibility	-0.0607***	-0.0483*	-0.0743**	-0.0106	-0.0698*
	(0.0219)	(0.0263)	(0.0324)	(0.0298)	(0.0391)
Observations	88,971	43,929	45,042	43,516	56,832
R-squared	0.092	0.074	0.104	0.062	0.133
Panel B: People without any formal education					
CSLs Eligibility	-0.167**	-0.275**	-0.110	0.0377	-0.221*
	(0.0705)	(0.124)	(0.0812)	(0.0655)	(0.127)
Observations	8,563	2,901	5,662	5,374	2,962
R-squared	0.120	0.147	0.124	0.080	0.262
Panel C: People with more than nine-year schooling					
CSLs Eligibility	0.0106	0.0551	-0.0431	-0.00442	0.00552
	(0.0360)	(0.0411)	(0.0595)	(0.0612)	(0.0656)
Observations	31,038	16,375	14,663	13,933	21,746
R-squared	0.075	0.069	0.089	0.078	0.165

Notes: Data source is CFPS, CHIPs and CHNS. Robust standard errors in parentheses are clustered at the province-year of birth level. Covariates and variable definitions are the same as those in Table 4.3. Because the effects of CSLs on underweight and smoking are only identified among women and men, respectively, this table thus examines the potential spillover effects only for men when the dependent variable is smoking and only for women when dependent variable is underweight.

formal education, if the individuals without formal education before and after CSLs are comparable, the estimated effects would be only caused by the externalities or education of others. But the condition may not hold because those who had no formal education after CSLs may be more adversely selected. In this case, however, the spillover effects are expected to be underestimated. If CSLs-eligibility is associated with better health in this specific group, it would provide some evidence for spillover effect; if not, it does not mean that there is no spillover effect at all. The estimates here present some evidence for spillover effects for self-reported health and smoking, but not for underweight. Specifically, among those without formal education, the CSLs fully-eligible cohorts have better self-reported health and lower smoking rate, and the magnitude is even two to three times larger than the average effects reported in Panel A.

In Panel C, I conduct the parallel analysis for the sample of those with more than nine-years schooling because Figure 4.3 implies that CSLs did not affect the received education among them. The results show that CSLs do not have any significant effects on health among these people, suggesting little spillover or external effects of CSLs for them. Therefore, these findings suggest that these results provide some evidence of externalities of education, but the externalities mainly exist for those with lower education.

The proportion which can be explained by the externalities would be quantitatively important. However, it is really difficult to accurately estimate this proportion without introducing any additional assumptions. But the above estimates enables a back-of-the-envelope calculation which only takes into account of the spill-over effects among those without any formal education. For example, take the self-reported health as an example. Suppose the estimated coefficients are estimated spill-over effects, and only consider those without any formal education, then my calculation suggests that the proportion could be over 27 percent in full sample, and 36 percent and 22 percent for men and women, respectively. The suggestive evidence shows that the large increase in education caused

---

department may not even have the case because most of the administrative departments were located in urban regions. Third, the CSLs cut the tuition but not abandon the fee. Many primary schools still collect different kinds of fees and there are some poor people still not going to school due to the cost.

by the CSLs may have large spillover effect on self-reported health among the population, especially those with lower education. Based on the conservative estimates, the explained proportion is fairly high compared to that explained by the individual intermediates.

4.7 Conclusions and Discussion

It is important to know whether and why education has a causal impact on health. However, the controversial discussion in the literatures has not come to a consensus that education improves individual health, but reveals the heterogeneity in the effects of education across different countries. This paper uses the exogenous temporal and geographical variation in the establishment of CSLs in China around 1986 to identify the effects of schooling on a series of health outcomes and shed some light on the possible mechanisms.

First stage results suggest that the CSLs significantly increased the education by 1.1 years in China on average. Because of the uniformly "nine-year" compulsory schooling years across all the regions, the results also suggest the policy-included increase in education is significantly larger in the regions with lower education prior to the CSLs were enforced. These variations caused by the CSLs provide valid estimates for the effects of CSLs on health outcomes. In the next, both the reduced form and 2SLS estimates provide sound evidence for the improved health status by the CSLs and the induced higher education. Specifically, the 2SLS estimates show that one additional year of schooling leads to 2-percentage points decrease in reporting fair/poor health (10 percent of the mean), 1.1-percentage points decrease in the rate of underweight (14 percent of the mean), and 1.3-percentage points decrease in the rate of smoking (5 percent of the mean).

The next part of this study aims to unravel the potential mechanisms. I use the framework in Cutler and Lleras-Muney (2010) and examine the potential roles of income, cognition and externalities in effects of education on health. The estimates suggest that income and cognition explain the impact of education on self-reported health by 7 percent and 15 percent, separately. These results suggest helping people to obtain knowledge about health is even more important for health than income. However, the empirical results suggest a more important role of the externalities of education in the effects of education on self-reported health, especially among those with lower education; a conservative calculation suggests the externalities explain over 25 percent. However, the results are different for various dependent variables. For example, income explains the effects on underweight by over 20-30 percent but only explains 5 percent of the effects on smoking. The results also suggest externalities may be important to explain the effects on smoking while hardly explain the effects on underweight.

Although this study provides some suggestive evidence on a couple of mechanisms, it is far from satisfactory. For one thing, it is still a question how much the spillover can explain the effects of education exactly. Further, it is also possible that the heterogeneity in mechanisms exists in different countries and in different periods. Due to data limitations, I leave these questions to future studies that will help us to gain a better understanding of the effects of education on health.

Chapter 5

Support the Elderly: Power of Social Pension[1]

5.1 Introduction

As most of the countries are aging fast, governments nowadays are considering to start or reform social pension schemes to better support lives of the elderly but they usually face tight fiscal budget at the same time. A natural question to ask is how much impact of social pension has on the individual behaviors and health status. The answer is important because the effects on income, expenditure, health and mortality are key parameters to evaluate and design *efficient* pension programs and retirement policies.

Although many economists recognized the importance of above question and there is a long literature on this, most of the empirical evidence is correlation rather than causality due to the severe data limitations and little exogenous variation of pension wealth (Bernheim *et al.*, 2002; Attanasio and Rohwedder, 2003). Although there are some exceptions,[2] the

[1] Co-authored with Chuanchuan Zhang.

[2] For example, South Africa expanded the social pension to black elderly in early 1990s, and some researchers thus used this exogenous shock to identify the effects on income, levels of living, expenditures, private transfers, intrahousehold allocation, and health (Case and Deaton, 1998; Case and Wilson, 2000; Case, 2004; Duflo, 2000, 2003; Jensen, 2004; Edmonds, 2006; Ardington *et al.*, 2009). For another, Snyder and Evans (2006) and Jensen and Richter (2004) used the exogenous variation in pension caused by "Notch" cohorts in the United States and the

answers are still far from satisfactory and even mixed. For example, the universal pension expansion in South Africa has little institutional variation, and the data are usually *ex post* and have very small sample size;[3] the pension benefits in industrial countries are usually based on the previous earnings or wages, and thus are correlated with underlying personal tastes or characteristics (Coile and Gruber, 2000; Chan and Stevens, 2004). For another, Snyder and Evans (2006) found that higher pension income leads to higher mortality due to social isolation, while Case (2004) and Jensen and Richter (2004) found that higher pension income makes people healthier and have lower mortality.

This study builds up the growing literature by providing new evidence on the effects of pension provision on income, labor supply, expenditure, health and mortality. To do so, we first provide cross-country evidence for the effects of social pension on mortality. Matching human mortality database (HMD) to the official starting year of social pension in 10 countries,[4] we find that mortality rate of the age-eligible people (i.e., people whose age is above the pension age threshold) significantly reduced 1.7-2.2 percent *just* after the introduction of pension programs, while that among the age-ineligible group (i.e., people whose age is below the pension age threshold) changed little and insignificantly. However, it is difficult to find micro data evidence for these pension schemes since most of them started in the late 18th or the early 19th century when few individual surveys were conducted.

The New Rural Pension Scheme (NRPS) in China provides a natural setting to fill in the gap and answer the above question. It covered 320 counties (about 10 percent of the total counties) in 2009, remarkably expanded in the next couple of years, and reached full coverage by the end of 2012. Once the county was covered, all the rural *hukou* people who are aged 16 or above can *voluntarily* participate in the scheme.[5] All the enrollees with ages

---

collapse of pension system in Russia, respectively, to identify the effects on mortality.

[3]For example, the sample size of Case (2004) is smaller than 500 and that in Jensen (2004) is smaller than 1,000.

[4]The 10 countries are Belgium, Canada, Denmark, Finland, France, Italy, Norway, Sweden, Switzerland and United States.

[5]A *hukou* is a record in the system of household registration required by law in mainland China. It is hard to change the *hukou* status in China, especially for the elderly. .

60 years or above are eligible to get a fixed amount of money, 55 yuan per month (9 US dollars), regardless of previous earnings or income.[6] But the condition is their offsprings, no matter age-eligible or age-ineligible, need to participate in the program and pay for the premiums.[7] Therefore, the pension scheme can be viewed as a conditional cash transfer program for the rural enrollees with ages 60 or above. By the end of 2014, 65 percent of rural *hukou* people participated in the pension scheme, and 140 million pensioners started to receive pension.

Base on the institutional variation in pension coverage, we use a national representative sample with ages 45 and above from China Health and Retirement Longitudinal Studies (CHARLS) and China Family Panel Studies (CFPS) to identify the effects of pension scheme provision on income, expenditure, transfers, and health. The time period in the sample spans the years from 2011 to 2014, exactly covering the period NRPS expansion.[8] This sample is composed of over 70,000 observations from more than 300 counties in China. Using the regional and temporal variation in NRPS coverage across different counties, we employ the Difference-in-Differences (DID) methodology to identify the effects of NRPS.

We first examine the mechanical effects, which suggests that the rural people aged 60 and over are 25 percentage points more likely to receive the pension just after the NRPS was in present. Consistent with the pension policy, we do not find any significant effect of the NRPS on pension receipt among the rural people aged below 60 or the urban people. Thus we divide the sample into four subsamples based on both age- and *hukou*- eligibility for the pension receipt: both age- and *hukou*- eligible, age-eligible but *hukou*-ineligible, age-ineligible but *hukou*-eligible, neither age nor hukou eligible (for simplicity, we name the first group as pension eligible group and all the rest three pension ineligible groups in the rest of this paper). And we separately examine the effects of NRPS in each subsample. Among the

[6]According to the yearly statistical data released the by the National Bureau of Statistics (NBS) in China, 55 yuan is about 15 percent of the per capita income of the rural households in the median group.

[7]The age-ineligible enrollees need to choose one of the following levels of annual contribution: 100, 200, 300, 400, 500 RMB. They have to pay for the premium yearly until they reach age 60.

[8]CHARLS and CFPS are two on-going surveys and this study used the first two waves of both - CHARLS 2011, 2013 and CFPS 2010, 2012. All these are the latest two waves that have been released for both data.

pension-eligible group, we find that 1) the NRPS significantly increased household income and food expenditure by 17.6 and 9.6 percent, respectively; 2) the scheme significantly reduced the labor supply by 3.0 percentage points (6.2 percent of the mean), and most of this came from the decline in farm work while little change in non-farm work; and 3) the scheme did not significantly affect the chance of receiving private transfers or the total household expenditures. For the pension ineligible groups, however, we do not find any significant effects on income, labor supply, expenditure or private transfers. The only exception is that the NRPS shifted the younger rural people (i.e. *hukou*-eligible but age-ineligible) from farm work to non-farm work - the scheme reduced the farm work by 5.8 percentage points but increased the non-farm work by 3.3 percentage points.

We follow the same methodology to further investigate the NRPS-induced health consequences. Among the pension-eligible people, we find that 1) the rates of reported disability and underweight reduced by 3.2 percentage points (11.4 percent) and 1.8 percentage points (11.3 percent) after NRPS implementation, respectively; 2) the implementation of NRPS reduced the mortality by 2.2 percentage points (14.4 percent) among the rural people with ages over 65 and above in individual-year panel from Chinese Longitudinal Healthy Longevity Survey (CLHLS);[9] and 3) the NRPS crowded out the health insurance participation by 4.2 percentage points (5.7 percent),[10] but it did not significantly affect health behaviors like smoking, or medical care usage such as inpatient and outpatient cares. As a comparison, we do not find that the NRPS had any significant effects on all the above outcomes among those pension ineligible groups with the exception that NRPS reduced health insurance participation by 3.5 percentage points (3.9 percent) for the rural people with ages below 60.[11]

[9]Our calculation suggests that the income-mortality elasticity ranges from 0.18 to 0.60; estimates in Jensen and Richter (2004) suggest the elasticity is 0.21 since the mortality increased by 5 percent when the income reduced by 24 percent.

[10]The results here are consistent with those in Bitler *et al.* (2005). A possible explanation for the crowd-out effect of NRPS on health insurance for all ages is the reduced uncertainty in (future) income because of introduction of social pension. Therefore, the effects of NRPS on health insurance can be viewed as the net effects of the reduced demand and increased income.

[11]The exceptions (i.e., the effects on labor supply and health insurance participation) reflect some effects of

The validity of DID estimation cannot be taken for granted since the counties are not chosen randomly.[12] One obvious concern is the heterogeneous trends across different counties since the effects from DID may just reflect the different trends across the counties. The empirical results above may help to alleviate this concern since they show that the NRPS-induced effects on income and health are much smaller and insignificant among the age-ineligible group and *hukou*-ineligible group in the same counties. Yet we provide more direct evidence by showing the pre-trends for different counties. In practice, for the counties with different stating years of NRPS, we separately plot a series of local macro economy indexes over years before 2009, including GDP per capita, salary of workers, government expenditure, government revenue, number of doctors, and number of beds in hospitals. We do not find any significant unparalleled trends across the counties for these indexes. Similarly, we did not find unparalleled pre-trends in mortality using CLHLS data either.

These findings contribute to several ongoing literatures. First, as mentioned above, the findings build up the literature investigating the effects of welfare or pension programs on individual behaviors such as expenditure, labor supply, retirement and insurance participation (Case and Deaton, 1998; Madrian and Shea, 2001; Attanasio and Rohwedder, 2003; Attanasio and Brugiavini, 2003; Bitler *et al.*, 2005; French, 2005; Ardington *et al.*, 2009; Aizer *et al.*, Forthcoming). In addition, the findings also provide new evidence on effects of income on health in the literature (Case and Wilson, 2000; Case, 2004; Frijters *et al.*, 2005; Jensen and Richter, 2004; Snyder and Evans, 2006; Evans and Moore, 2011, 2012; Aizer *et al.*, Forthcoming). Finally, our results on health insurance participation and labor supply for those *hukou* eligible but age-ineligible people are relevant to the literature on the effects of

cash transfer programs on those *hukou*-eligible people whose age had not reached the pension threshold. One possible explanation can be the NRPS-induced expectation (Atalay and Barrett, 2015; Staubli and Zweimüller, 2013; Gustman and Steinmeier, 2015). Since the program enrollees with ages below 60 would be able to receive pension when they reach the pension age, the higher and more stable (expected) income would change their current behaviors such as labor supply and insurance participation.

[12]To become eligible for NRPS, the counties need to first apply to the provincial government, then to the central government. It is the central government who made the decision to approve the application or not. There are no details about the qualification for pilot site in official documents. But in some news report, the officials of the Ministry of Human Resources and Social Security (who in charge of the NRPS) said the program tends to start earlier in the middle and western regions.

policy-induced expectation and the indirect effects of public policies (Atalay and Barrett, 2015; Staubli and Zweimüller, 2013; Gustman and Steinmeier, 2015).

The paper is organized as follows. Section II provides cross-country evidence for the effects of social pension. We introduce human mortality database, show the RD methodology and relevant results. Section III provides evidence from micro level data in China, including the background of NRPS, data, methodology and empirical results. Section IV concludes.

5.2 Cross-Country Evidence from Cohort Data Analysis

This section investigates whether the introduction of social pension reduced mortality in human history.

5.2.1 Data: Human Mortality Database and Birth Year of Social Pension

Mortality data are taken from the Human Mortality Database (HMD). The HMD contains detailed cohort life tables by year of birth and gender. A typical observation in the HMD is the mortality rate, per 100,000, for men (women) in a particular year in a particular country at certain age ranging from 0 to 110. The HMD data provide the mortality tables with different various years across 38 countries or regions.[13] The country specific timing of introduction of social pension is from Cutler and Johnson (2004) and pension-watch website.

We match the HMD data to the countries with available information of first social pension scheme introduction, and restrict to the countries with both mortality information before and after the introduction of social pension. We have 10 countries in total in the end. These countries are Belgium, Canada, Denmark, Finland, France, Italy, Norway, Sweden, Switzerland and United States. Among these countries, the earliest one introducing social pension is Denmark (1891) and the latest is Italy (1969). Table 5.1 shows the countries and the birth year of social pension in each of them.[14]

---

[13] The country list and available years can be found here: http://www.mortality.org/.

[14] Table 2 of Cutler and Johnson (2004) provides the detailed year of introduction, case of introduction, type of system, and later changes for the social pension in 20 different countries, and the pension watch website

Table 5.1: *Social Pension Scheme in 10 countries*

Country	Year introduced	Age of eligibility
Belgium	1924	65
Canada	1927	65
Denmark	1891	65
Finland	1937	65
France	1956	65
Italy	1969	65 and 3 months
Norway	1936	67
Sweden	1913	65
Switzerland	1948	65 (men) 60 (women)
United States	1937	65

Notes: Data are from Cutler and Johnson (2004) and Pension-Watch website (http://www.pension-watch.net/about-social-pensions).

5.2.2 Methodology and Empirical Results

Because both the level and trends in mortality differ largely when we use a long time panel, we use regression discontinuity (RD) to identify the effects of social pension scheme provision on mortality. We restrict the sample to those aged above age 45 because the elderly people are targeted population for social pension schemes. We also drop those aged above 90 because of possible misreporting issues and large measurement errors. For convenience, we define relative year t as the years difference of current year with the introduction year of social pension scheme. For example, it equals to -1 if the current year is the year before scheme introduction, and equals to 1 if the current year is just one year after that.

To control for the invariant factors such as country, gender, and age may also influence mortality, we keep the sample with 10 years bandwidth, divide the sample into 900 different groups (s) based on country (10), gender (2), and age (45). Within each group s, we detrend the logarithm of mortality rate over the relative year by regressing the logarithm of mortality on relative year and its square and keep the residuals. .We pool the residuals of all the groups, plot the linearly fit lines and confidential intervals (CI) over the relative year in

provide the policy-designed eligible ages for the pension schemes across the countries. The pension watch website is http://www.pension-watch.net/about-social-pensions.

Figure 5.1.[15] Figures 5.1a and 5.1b present the patterns for age-eligible and age-ineligible people, respectively. Figure 5.1a shows the social pension scheme introduction significantly reduced the mortality by 1.6 percent. However, we do not find any significant reduction in mortality in age-ineligible group.

We estimate the following equation to further test robustness of the results:

$$lnMR_{gact} = \alpha Post_{ct} + \delta_{gac} + t_{gac} + t_{gac}^2 + \varepsilon_{cagt} \tag{5.1}$$

The dependent variable, $lnMR_{cagt}$, is the logarithm of mortality rate of people of age a, gender g in country c in relative year t. $Post_{ct}$ is an indicator variable which equals to one if the country c had social pension in year t, and equals to zero if not. The coefficient, α, captures the effects of introduction of social pension on mortality in the interested sample. To control for the potential unobserved confounding factors, we include the fixed effects of gender, age, country and all the three combinations (δ_{gac}) in the regression. And, for each combination of gender (g), age (a) and country (c), we also control for the linear and square trends in relative year, t_{gac} and t_{gac}^2.[16]

Following the graphic analysis, we also divide the sample into age-eligible group and age-ineligible group to conduct the analysis, and report the OLS estimates in Table 5.2. Panel A and Panel B show the results for age-eligible and age-ineligible groups, respectively. Different columns show the RD regression results for different bandwidths - 5, 6 and 7 years.[17] The estimates in panel A consistently show that introduction of social pension significantly reduce the mortality of those age-eligible people by 1.6-2.2 percent. In contrast, we do not find any significant evidence for the effects among age-ineligible group, and the difference between age-eligible group and age-ineligible group are significant for the

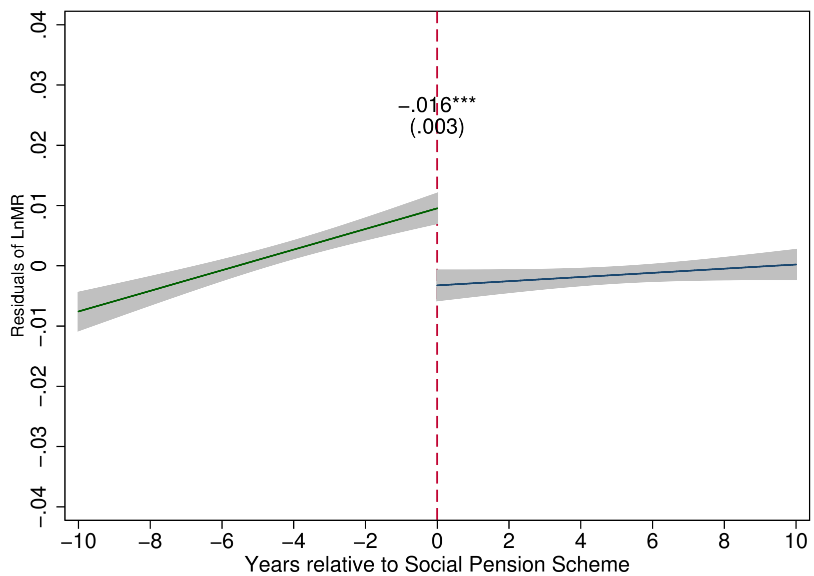
[15] We follow Ruhm (2000) and weight the residuals by the squared root of represented population size.

[16] For example, for those men aged 70 in Belgium, we have both linear and square trends; and we have another two trends for the women of the same age in Belgium. That is to say, if we estimate the equation in the whole sample used, we will have 900 dummies (δ_{gac}) and 1800 trend terms.

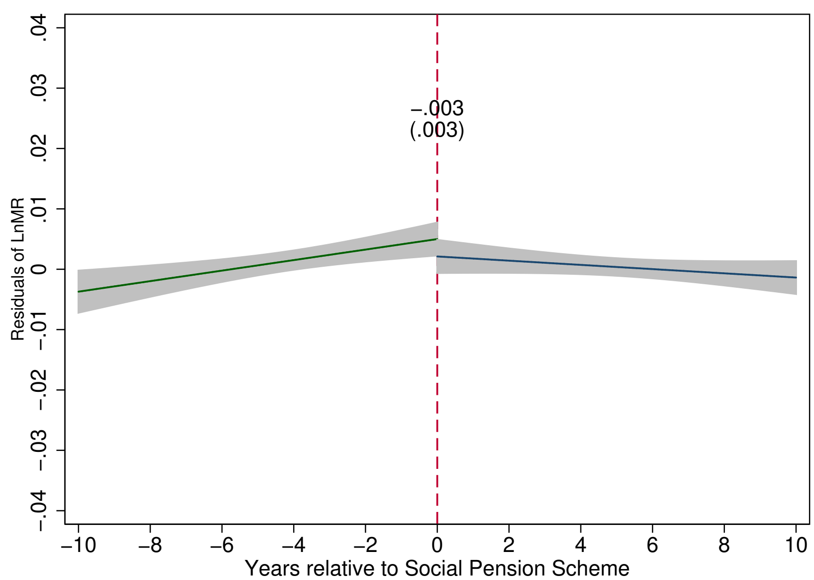
[17] We choose these bandwidths because the "optimal" bandwidth according to Calonico *et al.* (2014) is 6 years.

Figure 5.1: *Regression Discontinuity Estimation for the Effects of Social Pension Introduction on Mortality*

(a) *Age-eligible group*



(b) *Age-ineligible group*



Notes: The mortality data are from human mortality database (HMD) and the data about timing of pension are from Cutler and Johnson (2004) and Pension-Watch website (http://www.pension-watch.net/about-social-pensions). For each country-gender-age cell, we regress the logarithm of mortality on relative year and its square, then keep and pool the residuals of all the groups, and plot the linearly fit lines and confidential intervals (CI) over the relative year.

Table 5.2: *Regression Discontinuity Results for Effects of Introduction of Social Pension*

	(1)	(2)	(3)	(4)
Variables	Logarithm of Mortality Rate			
Bandwidth	5 years	6 years	7 years	6 years
Trends terms	Relative year and its square			Relative year linear trends before and after pension
Panel A: Age eligible group (pension age threshold and above)				
$Post_{ct}$	-0.022***	-0.017***	-0.016***	-0.022***
	(0.003)	(0.003)	(0.003)	(0.003)
Observations	5,605	6,539	7,421	6,539
R-squared	0.996	0.995	0.995	0.996
Panel B: Age ineligible group (45 - pension age threshold)				
$Post_{ct}$	-0.002	0.003	0.005	-0.001
	(0.003)	(0.003)	(0.003)	(0.004)
Observations	4,331	5,053	5,735	5,053
R-squared	0.994	0.994	0.994	0.994
F-statistics	17.68	19.10	20.48	19.45
P-value	0.00	0.00	0.00	0.00

Notes: Data are from Human Mortality Database, Table 2 of Cutler and Johnson (2004) and Pension-Watch website. All the regressions are weighted by the square root of population size. All the standard errors are clustered at country-gender-age level. The F-statistics in the bottom of the table test whether the coefficients difference between Panel A and Panel B. *** p<0.01, ** p<0.05, * p<0.1

three columns. Last column reports the results when we follow (Card *et al.*, 2008, 2009) to control for specific linear trends in relative year before and after the social pension provision. Consistent with the results above, the mortality rate among the age-eligible people drop 2.2 percent just after the introduction of social pension and no significant effect in age-ineligible group.

5.3 Evidence from New Rural Pension Scheme (NRPS) in China

Using the introduction year of pension scheme in different countries in human being history, the above analysis finds that provision of new pension scheme significantly reduced mortality by 1.7-2.2 percent among the age eligible people. However, the aggregated level data cannot tell more. Due to lack of the detailed official documentation, it is difficult to know how much money was spent and how the pension was distributed. In addition, without reliable micro level data, it is even more difficult to know individual associated behavior response to the scheme establishment. These limitations call for micro evidence on the pension-induced individual behavior responses and the health consequences, and the NRPS in China provides a natural setting to fill in the gap.

5.3.1 Background

Some rural regions in China are really poor. According to 2005 inter-census population sample survey, the median earning among the rural adults is about 200 yuan (about 30 US dollars) per month. The poverty is even more serious for the elderly; 67.5 percent of the rural people aged over 60 have no labor earnings in 2005, and 91 percent of them were living with and relied on their offsprings. According to a recent survey on the website, the pension reform is the number one issue among the rural people - 35.4 percent of individuals considered the pension reform as the most important problem in rural China.[18] These motivated the Chinese government to initiate the social pension program in rural regions.

The new social pension program for rural people started in September 2009,[19] and it reached a universal coverage by the end of 2012 after four rounds of expansions - the first round in the end of 2009, the second in middle 2010, and the rest two in middle 2011 and in late 2012. The scheme was implemented at county level, and it is the central government

[18]Source: http://toutiao.com/i6243882674679726593/.

[19]This is a first time of the rural China starting such a large and generous welfare program. It was the "new" rural pension scheme to distinguish it from the old rural pension scheme initiated in 1992. The old rural pension scheme is somewhat like an organized saving account, with premiums accumulated in an individual account and accrued at a low interest rate (Leisering *et al.*, 2002). At the height of the old rural pension scheme, 75.4 million people invested in the accounts, but the amount of pension it afforded was extremely insignificant.The development of the old pension scheme stagnated after 1998, partly because of the widespread mismanagement of the funds and the insignificance of the program (Shi, 2006; Wang, 2006). In 2005, the enrollment rate for the old pension scheme has dropped to less than 3 percent.

who made the decision to approve the counties to initiate the NRPS each year.[20] According to the official government documents we collect, the government aimed to make sure the approved counties distributed "balanced" across different regions in the first wave, but the program tended to start earlier in the middle and western (where is poorer) regions in the next two waves. We requested the data of timing of NRPS coverage in the counties from State Council Leading Group Office of Poverty Alleviation and Development, which officially replied with a formal documentation in two weeks. Figures 5.2a-5.2d show the geographical coverage year by year in mainland China. About 12 percent (about 320) of all the counties were covered in the first wave (2009), and 16 percent (450 counties) were covered in the next year (2010); 38 percent (about 1075 counties) started the program in the third wave (2011) and all the rest 33 percent were covered in the last wave (2012). According to the figures, the NRPS-covered counties are geographically distributed balanced in the first wave. And consistent with the official document, more mid- or west- counties were covered. By the end of 2012, as shown in Figure 5.2d, all the counties were covered. In this study, we exploit the variation in the timing of NRPS implementation, and conduct Difference-in-Differences (DID) regressions to identify the effects of the new pension scheme provision.

After the county was covered, all the rural *hukou* people who are aged 16 or above (not including students) can voluntarily participate in the scheme. All the enrollees with age above 60 years at start of pension scheme are eligible to get 55 yuan (9 US dollars) per month, regardless of previous historical earnings or income.[21] But the pre-requisite is the participation of their offsprings, no matter age-eligible or age-ineligible. The age-ineligible enrollees need to choose one of the following levels of annual contribution: 100, 200, 300, 400, 500 RMB.[22] The distribution method of NRPS pension is determined by local government.

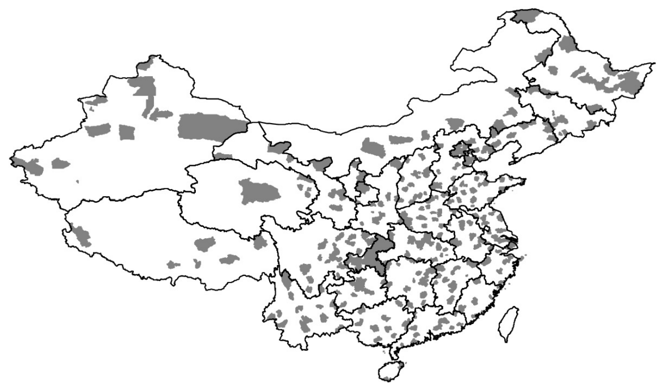
[20]The governments of county level applied to the provincial government, then to the central government. Then the central government select which counties to start the NRPS. There is no official files to document what the candidate counties are and how the central government approve or deny the candidate counties.

[21]In 2014, the benefits increased to 75 yuan per month.

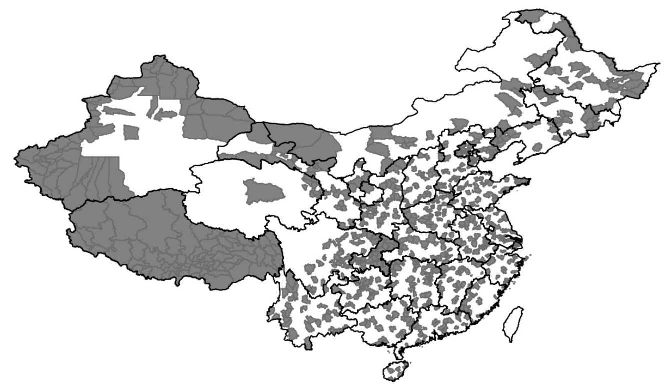
[22]Starting from pension eligible age, 60 years old, the pension benefits for a beneficiary is the sum of

Figure 5.2: *New Rural Pension Scheme (NRPS) Coverage over Time*

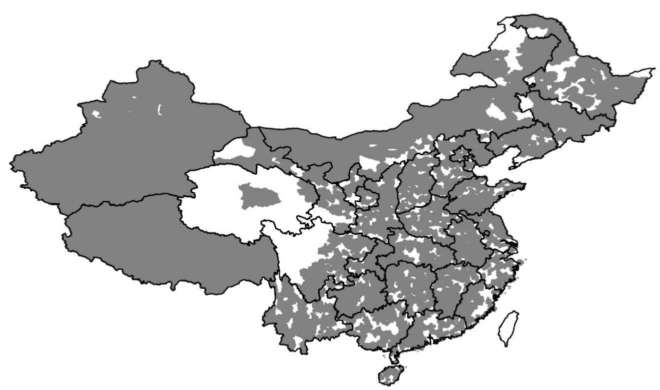
(a) *First round, November 2009*



(b) *Second round, July 2010 - October 2010)*



(c) *Third round, July 2011 - September 2011*



(d) *Fourth round, July 2012 - October 2012*

Notes: The data for county level NRPS coverage are from State Council Leading Group Office of Poverty Alleviation and Development. The data are not public and the researchers need to apply for the data directly from the office.

Some developed regions such as Jiangsu and Zhejiang, the local governments established the individual bank account for the senior enrollees and automatically transfer the pension to these accounts independently; however, in some less developed regions, the seniors or their offsprings have to go to the designated place in local village to get the pension by themselves. The funding is under strict regulations to avoid corruption or benefit fraud.[23] By the end of 2014, 65 percent of rural *hukou* people participated in the pension scheme, and 140 million pensioners started to receive pension. By the end of 2012, the central and local governments have input more than 262 billion Yuan in NRPS, with more than 232 billion from the central government.

5.3.2 Data

China Family Panel Studies (CFPS) and China Health and Retirement Longitudinal Studies (CHARLS)

The main sample used in this study is from CFPS and CHARLS. CFPS is a biennial survey and is designed to be complementary to the Panel Study of Income Dynamics (PSID) in the United States. The first national wave was conducted in 2010. The five main parts of the questionnaire include communities, households, household members, adults and children data. The China Health and Retirement Longitudinal Study (CHARLS) is also

the accumulated total funds in the individual account, plus the basic pension benefits. The way funds in accumulated individual account are paid out, according to the formula, is as follows: when a beneficiary turns 60, he/she starts to receive a monthly benefit (1/139 of the total accumulation) from the individual account. At the same time, he/she receives a basic pension benefit (currently 55 RMB per month). For instance, one person who participates in the program at the age of 45 and chooses to pay a yearly premium of 100 RMB will have a total amount of 1,838 RMB accumulated in the individual account (assuming at one-year deposit rate at the age of 60) and will receive a monthly benefit of 68.22 RMB (1838/139+55). Those who are already 60 years old at the time the program starts automatically receive a basic pension benefit (i.e., 55 RMB per month) without paying any premiums. Therefore, this pension scheme was fully funded defined contribution plan with added attraction of government subsidy toward contributions, coupled with a minimum pension guarantee wholly funded by the government.

[23] To make sure the eligible enrollees to get the pension, the central government required the local governments to provide the personal information of each enrollee and then appropriate the corresponding funding after careful verification; and this information is needed to updated year by year. Because the offsprings of eligible pensioners can go and get the pension in case the seniors are ill or in bed, the evidence for aliveness of the pensioners has to be in presence whenever receiving pension. The evidence could be a recent video or a certification from a local government officer who personally visited the pensioner recently.

a biennial survey, and aims to collect a high quality nationally representative sample of Chinese residents ages 45 and older, and is designed to be complementary to the Health and Retirement Survey (HRS) in the United States. More details about the two datasets are provided in appendix. The baseline national wave of CHARLS is being fielded in 2011. This study used the 2010 and 2012 waves of CFPS, and the 2011 and 2013 waves of CHARLS.

Because of the consistency in variables and sampling, we pooled the CFPS and CHARLS data together to make a larger sample and to best exploit the regional and temporal variation in the NRPS expansion during 2009-2012. Note that both of CFPS and CHARLS are nationally representative, and each covers about 5 percent of the total counties in mainland China.[24] The main sample we use for this study have over 70 thousand observation (i.e., about 34 thousand from CFPS and 36 thousand from CHARLS) and 312 counties (162 counties from CFPS and 150 from CHARLS).[25]

Chinese Longitudinal Healthy Longevity Survey (CLHLS)

The CLHLS is a longitudinal survey with aims for a better understanding of the determinants of healthy longevity of human beings in China. The baseline survey of CLHLS was conducted in 1998, with follow-up surveys with replacements for deceased elders were conducted every three years in a randomly selected half of the total number of counties and cities in the 22 out of 31 provinces in mainland China. However, the earlier waves only surveyed people aged over 80 and had a smaller sample size, and thus we choose the sample started in 2005. Since the survey 2005, CLHLS followed the respondents in 2008, 2011 and 2014. Besides the information on basic demographic and socioeconomic status, the data also provide the survival status for all the seniors in each wave, as well as the date for the deaths.

[24]In our analysis, we include the data source dummy and interact it with the counties all the time. Because the number of counties covered by both CHARLS and CFPS is small (i.e., only 5 counties), we name "county dummies" short for the county dummies interacting with data source.

[25]The number of observations and counties are consistent with the population distribution in mainland China.

5.3.3 Methodology and Empirical Results

Who Received Pension from NRPS?

The first question we want to answer is who received money from the new social pension scheme. The answer is important to understand and interpret the results for the possible effects of NRPS provision. Above all, we would expect some effects of NRPS if only the rural enrollees received money from the program. We can also test the mechanical effects of NRPS provision and provide evidence for the policy effectiveness by doing so. We thus follow the strategy in Hoynes *et al.* (2012) and estimate the following equation:

$$Receipt_i^s = \alpha_0^s + \alpha_1^s NRPS_{ct}^s + \delta_c^s + \delta_t^s + X_{ict}^s + \epsilon_{ict}^s \tag{5.2}$$

The superscript s indicate a specific subsample, which can be a group of people with certain characteristics. The dependent variable $Receipt_i^s$ is an indicator for the household of individual i received any pension, which is consistently measured in both CFPS and CHARLS. $NRPS_{ct}$ is another indicator whether county c had the NRPS in year t. The covariates also include county dummies (δ_c), year dummies (δ_t), and other demographic controls (X_{ict}) such as gender, age and its square and education level. The coefficient on $NRPS_{ct}$, α_1^s, captures the short-term effects of new pension scheme provision on social pension receipt in subsample s. All the standard errors are clustered at county level (Bertrand *et al.*, 2004).

We first divide the sample by their *hukou* status and age in years to verify whether people aged over 60 of rural *hukou* are the only eligible group of pension receipts. The results are shown in Figure 5.3a. Each point and the corresponding intervals in the figure shows the coefficient, α_1^s and 90 percent CIs, derived by a separate regression in subsample s. The effects among urban people are always insignificant. Among the rural people, consistent with the policy design, the effects are positively significant for those aged over 60 but insignificant among those aged below 60. The pattern for rural people shows a significant jump at the threshold - age 60. We emphasize here the estimation identified *short-term* effects, which reflects how much the outcome variables change *just* after the NRPS coverage.[26]

Then we restrict the sample to those aged 60 with rural *hukou* to conduct further analysis. Panel A of Figure 5.3b presents the point estimate and 90 percent CIs for men and women, respectively. The effects are significant for both men and women, with insignificant difference in between. Panel B divides the sample by education level, and we find that the effects among the three groups are similar (i.e., all the coefficients are between 0.2 and 0.3). Panel C divides the sample by the county income level in 2005, and we find that the effect of NRPS on receipts in poorer regions are much larger than that in richer regions. This is consistent with expectation that people in the regions with more poverty have higher incentive to enroll.

Effects of NRPS on Income, Labor Supply and Expenditure

We also use the same framework to investigate behavior responses to the NRPS:

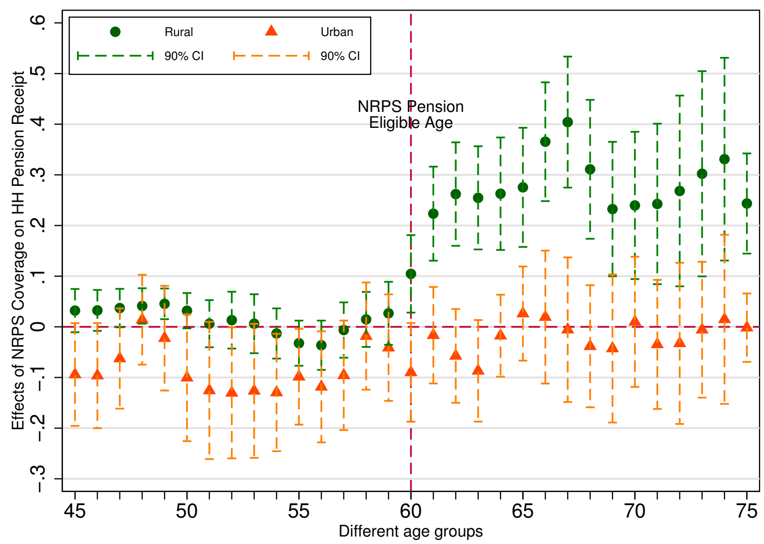
$$Y_{ict} = \beta_0 + \beta_1 NRPS_{ct} + \delta_c + \delta_t + X_{ict} + \epsilon_{ict} \tag{5.3}$$

The dependent variable Y_{ict} is the candidate outcomes to examine, which can be household income, health status and other interested outcomes. All the other variables are the same with those in equation 5.2. Same as above analysis, all the standard errors are clustered at county level. The estimation is based on the differences between before-after changes in outcomes of treated group and that in the same time period in control group. Since no evidence shows that the counties starting NRPS in different years were randomly selected, the DID estimator, β_1, is subject to a number of limitations. Most importantly, the estimation presumes that the trend of outcome variable Y_{ict} in treated group would be parallel to that in control group had the NRPS been not conducted. Therefore, we need to examine whether

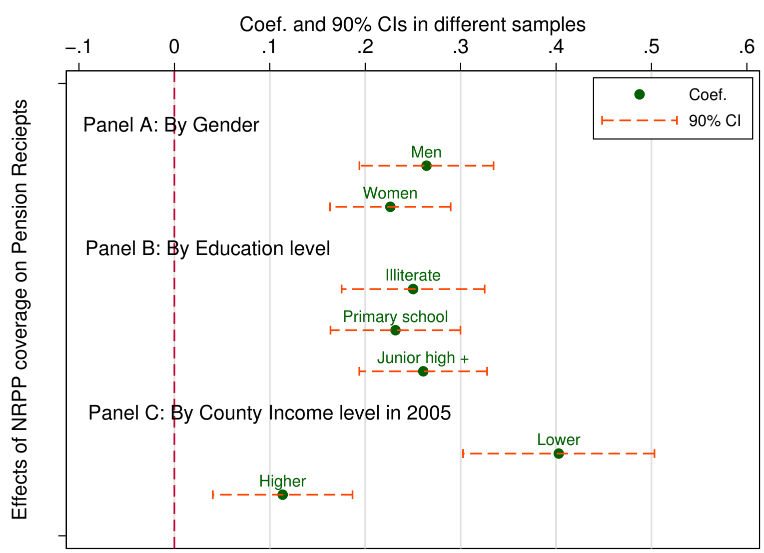
[26]There are some reasons why the older people may not fully participate in the program just after its implementation. First, they might not trust the policy in the early stage, especially peasants who have experienced the introduction and collapse of the old rural pension scheme; Second, some local governments need time to prepare documents and setup individual account; Third, information transition took some time because some potential enrollees may even not know the NRPS after the NRPS implementation and surveys. Fourth, their adult children may not want to enroll the NRPS, which is the prerequisite for them to receive pension benefit.

Figure 5.3: *Effects of NRPS introduction on Pension Receipt*

(a) *By type of* hukou *and age*



(b) *By gender, education and initial income level*



Notes: The data are from CFPS and CHARLS. Figure a above divides the sample by type of *hukou* and age in years. Figure b only uses the pension-eligible sample and divide it by gender (panel a), education level (panel b), and county income level (panel c), respectively. Each point and corresponding 90 percent confidential interval are based a separate regression of equation 5.2. The confidential intervals are calculated based on standard errors clustered at county level.

the counties covered in different waves have parallel trends in the outcome variables before NRPS coverage (i.e., pre-trend tests), and Section 3.4 provides details about this.

According to the policy design and based on the findings above, we have a couple of potential comparison groups to test the robustness and validity of our results, and alleviate the concern about the non-parallel trends as well.[27] The first is composed by the urban *hukou* people in the same counties, we would expect no effects among them due to *hukou*-ineligibility. The other comparison group is the people with rural hukou but ages below 60. However, we should bear in mind that the two comparison groups are not perfect. Urban people may not be comparable to the rural ones, and the differences between urban and rural people differ across region over time. For the other, those aged 60 in the same village may form different expectation because of introduction on NRPS since enrollees are eligible to receive pension once they reach age 60, and the enrollees also need to pay premium of NRPS because of enrollment.[28]

With the above considerations, we divide the whole sample by both age and *hukou* eligibility - rural people aged 60 and above, rural people aged below 60, urban people aged 60 and above, and urban people aged below 60. The first is the only group people who are eligible to both enroll in the pension scheme and receive 55 yuan per month once enrolled in the NRPS. The second group people are eligible to participate but not to receive pension. The third and the fourth group are not eligible to participate in NRPS.

Table 5.3 shows the results on the effects of NRPS on pension receiving and household income. Panel A and Panel B present the results for those aged 60 or above and those aged below 60, respectively. First two columns examine the effects for those with rural *hukou*. Consistent with Figure 5.3, the estimates suggest that NRPS coverage ($NRPS_{ct}$) significantly increased the probability of household pension receiving by 24.5 percentage points among rural people with ages 60 and above. The NRPS coverage also significantly increased the

[27]If the effects in treated group were mainly driven by the non-parallel pre-trends, we should expect the effects would also appear in these comparison groups.

[28]Previous literature such as Angelucci and De Giorgi (2009) found that cash transfer program also indirectly affects the consumption of the ineligible households in the same villages.

household income by 17.6 percent. In contrast, we do not find any significant evidence for the effects among rural but age-ineligible people in Panel B, whereas the coefficients are much smaller. The last two columns examine the effects for urban people. Consistently, the estimates do not present any significant effects of NRPS on pension receiving and household income in this group, no matter those aged above or below 60.

The NRPS-induced household income changes may not originate from the mechanical effects (i.e., pension receiving). Table 5.4 further examines the labor supply response to the NRPS coverage. Among rural people, the labor supply of those aged 60 and over reduced by about 3.0 percentage points (6.4 percent) significantly, and that of those aged below 60 also reduced by 2.6 percentage points (3.6 percent), though it is not statistically significant.

The next two columns further investigate the effects by decomposing the type of work into farm work and non-farm work. NRPS significantly reduced the proportion of farm work by 3.6 and 5.4 percentage points, for the age-eligible and age-ineligible people, respectively. For the age-ineligible people, however, the implementation of NRPS increased the proportion of non-farm work by 3.3 percentage points. One explanation for the "shift" is that farm work is generally more labor intensive and unfavorable, and thus people tend to "escape" from farm work in presence of stable income flow in the future. Our results are consistent with Angelucci and De Giorgi (2009), and also suggest that we need to be careful to interpret the results from econometric framework combining the age-ineligible people as control group. Consistent with expectation, the last column shows there is no significant effect among urban people.

The first two columns in Table 5.5 examine the effects of NRPS on received private transfer of the household. The estimates show no significant effects, suggesting that the provision of NRPS scheme might not have crowded out the private transfer to the elderly.[29] The last two columns examines the effects on total expenditure and the expenditure on

[29]Jensen (2004) used the pension expansion in South Africa and found that each rand of public pension income to the elderly leads to a 0.25–0.30 rand reduction in private transfers. One possible reasons is that the pension expansion in South Africa increased the individual income by almost 200 percent, which is much larger than NRPS.

Table 5.3: *Effects of NRPS on Pension Receipts and Household Income, by Type of hukou and Age-eligibility*

	(1)	(2)	(3)	(4)
Sample	Rural *hukou*		Urban *hukou*	
Variables	Household receiving pension (Yes = 1)	Log (Household income)	Household receiving pension (Yes = 1)	Log (Household income)
Panel A: Age-eligible group (60+)				
Mean	0.43	9.67	0.63	10.64
$NRPS_{ct}$	0.245***	0.176***	-0.023	0.041
	(0.039)	(0.068)	(0.016)	(0.055)
Observations	21,434	20,584	8,601	8,298
R-squared	0.448	0.219	0.644	0.303
F-statistics	–	–	86.6	15.8
P-value	–	–	0.00	0.00
Panel B: Age-ineligible group (45-59)				
Mean of Y	0.07	10.12	0.28	10.71
$NRPS_{ct}$	0.012	0.058	0.013	0.005
	(0.011)	(0.060)	(0.011)	(0.044)
Observations	28,795	27,575	10,145	9,822
R-squared	0.091	0.195	0.335	0.274
F-statistics	42.1	4.87	–	–
P-value	0.00	0.03	–	–

Notes: The data are from those ages 45 and above in CHARLS and CFPS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.
*** p<0.01, ** p<0.05, * p<0.1

Table 5.4: *Effects of NRPS on Labor Supply*

	(1)	(2)	(3)	(4)
		Rural *hukou*		Urban *hukou*
VARIABLES	Working now (Yes = 1)	Doing any farm work (Yes = 1)	Non-farm work (Yes = 1)	Working now (Yes = 1)
Panel A: Age-eligible group (60+)				
Mean of Y	0.477	0.424	0.054	0.121
$NRPS_{ct}$	-0.030*	-0.036**	0.006	0.017
	(0.018)	(0.018)	(0.006)	(0.011)
Observations	21,290	21,264	21,264	8,484
R-squared	0.284	0.246	0.092	0.267
F-statistics	–	–	–	6.69
P-value	–	–	–	0.01
Panel B: Age-ineligible group (45-59)				
Mean of Y	0.727	0.544	0.184	0.453
$NRPS_{ct}$	-0.026	-0.058**	0.033**	0.003
	(0.022)	(0.024)	(0.015)	(0.020)
Observations	28,376	28,334	28,334	9,797
R-squared	0.225	0.208	0.209	0.315
F-statistics	0.06	1.42	3.80	–
P-value	0.80	0.23	0.05	–

Notes: The data are from those ages 45 and above in CHARLS and CFPS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.
*** p<0.01, ** p<0.05, * p<0.1

Table 5.5: *Effects of NRPS on Received Private Transfer and Household Expenditure*

VARIABLES	(1) Received private transfer (Yes = 1)	(2) Log(Received private transfer)	(3) Log(HH total expenditure)	(4) Log(HH food expenditure)
Panel A: Age-eligible group (60+)				
Mean of Y	0.38	6.67	9.48	8.55
$NRPS_{ct}$	0.001	0.129	0.032	0.096*
	(0.028)	(0.101)	(0.044)	(0.058)
Observations	21,300	8,099	16,220	15,906
R-squared	0.148	0.196	0.189	0.262
Panel B: Age-ineligible group (45-59)				
Mean of Y	0.45	7.05	9.86	8.76
$NRPS_{ct}$	-0.015	-0.039	-0.012	0.036
	(0.025)	(0.093)	(0.033)	(0.051)
Observations	28,447	12,871	23,024	22,702
R-squared	0.264	0.240	0.196	0.284

Notes: The data are from those ages 45 and above in CHARLS and CFPS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.
*** p<0.01, ** p<0.05, * p<0.1

food. The results suggest that NPRS significantly increased food expenditure by 9.6 percent, respectively. The effect on total expenditure is positive but small and statistically insignificant. The magnitude is consistent with the NRPS-induced increase in food expenditure.

The effects on living arrangement and migration are important because the above results would be misleading had NRPS induced changes in living arrangement (e.g., the size of household increased and thus the total income increased).[30] Table 5.6 examines the effect of NRPS on household size and cross-county migration for those rural people. First, only 3 percent of people have different registered *hukou* county and current living county.

[30]The seminal work in this literature, Case and Deaton (1998), expected that the short-term effect of pension on living arrangement and migration decision should be small. However, Case and Deaton (1998) did not provide empirical evidence on this important presumption because of data limitation.

Table 5.6: *Effects of NRPS on living arrangement and migration*

VARIABLES	(1) Log(Household size)	(2) Cross-county migrants (Yes =1)
Panel A: Age-eligible group		
$NRPS_{ct}$	0.001	-0.021
	(0.014)	(0.018)
Observations	20,870	11,518
R-squared	0.265	0.133
Panel B: Age-ineligible group		
$NRPS_{ct}$	-0.001	-0.007
	(0.011)	(0.011)
Observations	28,240	16,445
R-squared	0.290	0.145

Notes: The data are from those ages 45 and above in CHARLS and CFPS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level.
*** p<0.01, ** p<0.05, * p<0.1

Consistent with the expectation in Case and Deaton (1998), the estimates does not show any evidence for the short-term effects of NRPS on household size or migration.

Effects of NRPS on Health and Healthcare Usage

Then we move forward estimate the effects of NRPS on health outcomes. We use self-reported fair/poor health, reported disability,[31] and underweight as triple-dimension measures for health status. In addition, we exploit the methodology in Poterba *et al.* (2013), use a principal component analysis (PCA) on the three dimensions, and get a unhealthiness score for a health measure for the full sample. This measure has zero mean and varies from

[31]The disability variable is constructed based on a set of activities, including walking, cooking, dining, traveling, shopping, and doing housework. The respondents are asked whether have difficulty in doing these activities in both CHARS and CFPS, we define them disabled if they have difficulty in doing any of these activities..

-1.37 to 3.35, with a standard deviation of 1.1. As pointed out in Poterba *et al.* (2013), the comprehensive measure is shown to be strongly associated with current health status and future mortality.

Table 5.7 shows the results. First column presents the results for unhealthiness score for rural people. First four columns show the results for rural people. Results in Panel A show that the mean value of the unhealthiness score is 0.33 for the age-eligible and -0.14 for the age-ineligible ones. Estimates in Panel A show that NRPS coverage significantly reduces unhealthiness score by 0.12 among age-eligible people, indicating that 0.1 standard deviation improvement in healthiness. In contrast, the estimate for age-ineligible group is insignificant and much smaller in magnitude, which is about one-third of that in Panel A.

The next three columns present the results for different health measures. We find significant effects of NRPS on health improvement for all measures except for self-reported health. Specifically, NRPS reduced the disability rate by 3.2 percentage points and underweight rate by 1.7 percentage points. In age-eligible group, although the estimates are negative suggesting the health is somehow improved, all the coefficients are about half or one-third of those in Panel A and insignificant. The F-tests suggest that the differences between the effects in age-eligible and those in age-ineligible group are significant at 10 percent level.

The final column shows the results for urban people. As expected, these people are healthier than their counterpart group with rural *hukou*. Investigation in the effects of NRPS in this groups also yield little and insignificant coefficients. In addition, the F-statistic and P-value suggest a significant difference between the effects of NRPS on health for rural and urban people.

Table 5.8 examines the effects of NRPS on individual behaviors such as health insurance participation, healthcare usage and smoking. Estimates in column 1 show that NRPS discourages people to participation health insurance, for both age-eligible and age-ineligible people.[32] The provision of social pension scheme significantly reduced the health insurance

[32]The health insurance program for rural people is mainly New Rural Cooperation Medical Insurance Scheme (NCMS) began in 2003 and reached a universal coverage in 2008, which is a heavily subsidized voluntary health insurance program targeting rural residents(Wagstaff *et al.*, 2009).The participation of NCMS is also voluntary

Table 5.7: *Effects of NRPS on Health Outcomes*

	(1)	(2)	(3)	(4)	(5)
	Rural *hukou*				Urban *hukou*
Variables	Unhealthiness score	Reported fair or poor health (Yes =1)	Reported disability (Yes = 1)	Underweight (Yes =1)	Unhealthiness score
Panel A: Age-eligible group (60+)					
Mean of Y	0.312	0.740	0.280	0.153	-0.00106
$NRPS_{ct}$	-0.117***	-0.015	-0.032*	-0.017*	0.030
	(0.045)	(0.020)	(0.017)	(0.010)	(0.041)
Observations	17,723	21,175	21,493	17,861	7,139
R-squared	0.167	0.071	0.197	0.120	0.160
F-statistics	–	–	–	–	13.9
P-value	–	–	–	–	0.00
Panel B: Age-ineligible group (45-59)					
Mean	*-0.139*	*0.713*	*0.108*	*0.0588*	-0.293
$NRPS_{ct}$	-0.042	-0.009	-0.010	0.000	-0.023
	(0.036)	(0.018)	(0.010)	(0.006)	(0.034)
Observations	24,568	28,647	28,899	24,611	8,316
R-squared	0.112	0.062	0.125	0.054	0.086
F-statistics	3.74	0.11	2.79	3.23	–
P-value	0.05	0.74	0.09	0.07	–

Notes: The data are from those ages 45 and above in CHARLS and CFPS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.
*** p<0.01, ** p<0.05, * p<0.1

participation rate by 5.2 percentage rate for the age-eligible group, but it also reduced the participation rate by 3.5 percentage points for the age-ineligible group, suggesting a crowd-out effects of NRPS on NCMS.[33] Although increased income would also enable rural people to purchase the insurance (i.e., increased income), the NRPS program also reduced insurance demand (i.e., reduced risk). People, no matter age-eligible or not, will face smaller income variation when the social pension program is in presence because of ensured pension at certain ages. This is especially true when most of the labor income is based on labor-intensive work such as farming in rural areas. The identified effects in column 1 can be interpreted as the net effects from both reduced demand and increased income. The reduced risk in future income also applies to those aged below 60 because expectations can be formed.

The next few columns in Table 5.8 examine the effects on other behaviors. All the estimates suggest that there are little effects of healthcare usage measured by inpatient and outpatient care. The estimates for both outcomes are insignificant and of small magnitude. The last column examines the smoking behaviors because previous literature suggests that higher income does not induce better health because of increased smoking (Chaloupka and Warner, 2000; Ruhm, 2000). However, the estimates for smoking in last column only yields insignificant effects of NRPS. Therefore, we conclude that NRPS improved individual health and crowded out health insurance participation, and there is no evidence for the effects of NRPS on other behaviors such as healthcare usage and smoking.

Note that we pool the two datasets and conduct the above regressions. Table 5.9 provides the regression results weighted by the represented population size of each datasets, which are fairly consistent with what have been represented.

---

and the enrollees need to pay for a yearly premium.

[33] The results are consistent with the findings in Bitler *et al.* (2005), who found the welfare program deterred people to participation in health insurance in the US.

Table 5.8: *Effects of NRPS on Healthcare Usage and Health Behaviors*

	(1)	(2)	(3)	(4)	(5)
Sample	Rural Hukou				Urban Hukou
Variables	Health Insurance Participation (Yes = 1)	Outpatient care (Yes = 1)	Inpatient care (Yes = 1)	Smoke currently (Yes =1)	Health Insurance Participation (Yes = 1)
Panel A: Age-eligible group (60+)					
Mean of Y	*0.915*	*0.258*	*0.167*	*0.282*	0.869
$NRPS_{ct}$	-0.043*	-0.012	0.004	-0.004	0.018
	(0.022)	(0.015)	(0.012)	(0.009)	(0.030)
Observations	21,310	21,457	17,336	19,887	8,672
R-squared	0.101	0.072	0.204	0.321	0.182
F-statistics	–	–	–	–	3.15
P-value	–	–	–	–	0.08
Panel B: Age-ineligible group (45-59)					
Mean of Y	*0.923*	*0.221*	*0.111*	*0.302*	0.817
$NRPS_{ct}$	-0.035**	-0.010	-0.013	-0.004	-0.020
	(0.016)	(0.012)	(0.009)	(0.007)	(0.022)
Observations	28,615	28,796	22,318	27,314	10,214
R-squared	0.119	0.057	0.246	0.432	0.203
F-statistics	0.28	0.01	1.64	0.00	–
P-value	0.60	0.91	0.20	1.00	–

Notes: The data are from those ages 45 and above in CHARLS and CFPS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.

*** p<0.01, ** p<0.05, * p<0.1

Table 5.9: *Effects of the NRPS in eligible group, weighted by represented population size in each dataset*

Dependent variables	(1) HH receiving pension (Yes = 1)	(2) Log (HH income)	(3) Received private transfer (Yes = 1)	(4) Log(private transfer)	(5) Log(HH expenditure)	(6) Log(Food expenditure)
Mean of Y	0.43	9.67	0.380	6.667	9.478	8.551
$NRPS_{ct}$	0.254***	0.178**	-0.000	0.127	0.032	0.096*
	(0.039)	(0.070)	(0.027)	(0.100)	(0.044)	(0.058)
Observations	21,434	20,584	21,300	8,099	16,220	15,906
R-squared	0.435	0.221	0.151	0.193	0.189	0.262

	(7) Working now (Yes = 1)	(8) Unhealthiness score	(9) Reported fair/poor health (Yes =1)	(10) Reported disable (Yes = 1)	(11) Underweight (Yes =1)	(12) Health Insurance participation (Yes = 1)
Mean of Y	0.477	0.312	0.740	0.304	0.153	0.882
$NRPS_{ct}$	-0.047***	-0.126***	-0.020	-0.044**	-0.017*	-0.050**
	(0.018)	(0.048)	(0.054)	(0.020)	(0.009)	(0.021)
Observations	21,357	17,723	21,175	21,164	17,726	21,508
R-squared	0.223	0.165	0.195	0.191	0.120	0.134

Notes: The data are from those ages 45 and above in CHARLS and CFPS. All the regressions are weighted by the represented population of the datasets. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, survey year and county. All the standard errors are clustered at county level.
*** p<0.01, ** p<0.05, * p<0.1

Effects of NRPS on Mortality

We transfer the CLHLS data to an individual balanced panel from 2006 to 2014, then use a dummy variable to depict the individual mortality status in the period. If the person is alive in 2014, then this variable is consistently equal to zero for the 9 years; and if the person died in year t, the value of this variable is set to zero for the years prior to year t, and is equal to one for year t and missing for the years afterwards. By doing so, we use best the time of the death and its variation. We then use this individual panel data to match the NRPS availability and conduct the following regression:

$$Die_{it} = \gamma_0 + \gamma_1 NRPS_{ct} + \delta_c + \delta_t + X_{ict} + \delta_{ia} + \epsilon_{ict} \tag{5.4}$$

The new independent variable is an indicator whether individual i died in year t. It equals to one if yes. All the other variables are the same with those in equation 5.3 except that we include a indicator δ_{ia} here to capture whether individual i is lost to follow during the years (i.e., attrition). All the standard errors are also clustered at county level. However, CLHLS does not provide information for *hukou* type. As a result, we use their residency type and the eligibility for retirement scheme instead.[34] In practice, we choose the people living in rural regions and having no retirement scheme as the treated group, and those living in urban regions and having retirement scheme as a comparison group. Column 1 of Table 5.10 presents the results. Panel A shows that NRPS reduced the mortality by 2.2 percentage points (14.4 percent of the mean value) among the treated group and had no significant effects in the comparison group.[35] Therefore, the estimates provide significant evidence for the effects of social pension on mortality.

The findings in previous literature are mixed: Jensen (2004) found 5 percent increase

[34]Only using residency type is incorrect because people living in urban region may have migrated from rural regions and have rural *hukou*. We additionally use whether the individual i is eligible for retirement scheme because of the fact that those who enjoyed retirement scheme generally have urban *hukou* and are not eligible for NRPS.

[35]Although the F-test cannot reject the null hypothesis for the coefficient difference due to large standard errors in the comparison group, the magnitude in the treated group is over three times larger than that in the comparison group.

Table 5.10: *Effects of NRPS on Mortality, CLHLS*

Variables	(1) Died in this year (Yes =1)	(2) Died due to severe disease (Yes =1)	(3) Died but have no severe disease (Yes =1)
Panel A: Living in rural area and having no retirement scheme			
Mean of Y	0.150	0.0541	0.0962
$NRPS_{ct}$	-0.0217**	-0.00426	-0.0174**
	(0.00952)	(0.00638)	(0.00793)
Observations	29,871	29,871	29,871
R-squared	0.139	0.060	0.122
Panel B: Living in urban area and having retirement scheme			
Mean of Y	0.102	0.0568	0.0456
$NRPS_{ct}$	-0.00678	-0.00195	-0.00483
	(0.0136)	(0.0107)	(0.00939)
Observations	9,047	9,047	9,047
R-squared	0.196	0.125	0.179
F-statistics	0.86	0.04	1.03
P-values	0.35	0.84	0.31

Notes: The data are from those ages 65 and above in CLHLS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, calendar year, county and whether the individual was lost in the years. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.

*** p<0.01, ** p<0.05, * p<0.1

in mortality after a pension system collapse in Russia in 1998 because of worse nutrition intake,[36] but Snyder and Evans (2006) found a significant drop in mortality when the elderly received less pension. Therefore, our findings build up the literature and provide consistent evidence for Jensen (2004). A natural question is whether the magnitude here is consistent with that in Jensen (2004). Since the oldest-old people in rural areas are the poorest group in China, the basic pension matters significantly to them, the back-of-envelop calculation of mortality-income elasticity in our study ranges from 0.18 to 0.60, which has the same scale as 0.21 in Jensen (2004).[37]

The next two columns divide the causes of death by whether the death is caused by a severe disease, and the results show that the NRPS induced mortality reduction is mainly contributed by less likelihood of deaths without severe disease. Since the deaths caused by severe disease usually took a relatively longer time than deaths of other causes, this is reasonable that the short-term effects of NRPS are mainly in present for the deaths without severe disease.

Although the individuals who were lost are only 8 percent of the full sample, it is not trivial when compared to the mortality. Because we have no information about whether the lost ones were dead or not, we drop the individuals who are lost in these years and conduct the same regressions as equation (5.4). Table 5.11 shows the results and they are fairly consistent with those in Table 5.10.

5.3.4 Pre-trends Tests

Our previous analysis uses the DID to identify the effects of NRPS on income, health and mortality and the estimates provide some evidence for them. However, the validity of the

---

[36]Jensen (2004) found that the income reduced by 24 percent and the two-year mortality increased by 5 percent, thus mortality-income elasticity = 0.21.

[37]Note that this sample is different from the CHARLS and CFPS sample since it overweights the people aged over 80 and those with lower income. In this sample, the median household income is 3,000 yuan per year in 2005, and the average household size in the CLHLS sample is 2.9. However, there is no information in CLHLS about the participation of new pension scheme. We thus conducted a back-of-envelope calculation suggesting the mortality-income elasticity ranges from 0.18 to 0.6 (i.e., the elasticity 0.18 is derived under assumption all the seniors participated, the elasticity 0.6 is derived suppose the NRPS participation rate just coverage is 0.3).

Table 5.11: *Effects of NRPS on Mortality in CLHLS, without those without mortality information*

Variables	(1) Died in this year (Yes =1)	(2) Died due to severe disease (Yes =1)	(3) Died but have no severe disease (Yes =1)
Panel A: Living in rural area and having no retirement scheme			
Mean of Y	0.157	0.0565	0.101
$NRPS_{ct}$	-0.0223**	-0.00417	-0.0181**
	(0.00981)	(0.00657)	(0.00817)
Observations	28,407	28,407	28,407
R-squared	0.137	0.060	0.124
Panel B: Living in urban area and having retirement scheme			
Mean of Y	0.124	0.0565	0.0690
$NRPS_{ct}$	-0.00798	-0.00359	-0.00438
	(0.0155)	(0.0122)	(0.0107)
Observations	7,456	7,456	7,456
R-squared	0.201	0.131	0.190

Notes: The data are from those ages 65 and above in CLHLS. The covariates in the regressions in each column include age and its square, and dummies for gender, education level, calendar year and county. All the standard errors are clustered at county level. The F-statistics in the bottom of each panel test whether differences with those for the rural people with ages 60 and over are significant or not.
*** p<0.01, ** p<0.05, * p<0.1

DID methodology cannot be taken for granted. For example, if the wave 1 counties have a more rapid improvement in health or development in economy prior to 2009, the effects identified by DID may just pick up the heterogenous trends rather than the actual effects of NRPS. The heterogenous trends may be caused by county-year level unobserved factors. We provide some evidence above to alleviate this concern that the NRPS-induced effects on income and health are much smaller and insignificant among the rural people aged below 60 and the urban people aged above 60 in the same village or county. Still, we plot the trends before the treatment (i.e., pre-trends) to test whether the presumption is true.

However, our analysis on effects of NRPS is based on the micro data in 2010 and afterwards, and thus it is impossible to plot and compare the pre-trends for both treated and control groups using data from CHARLS and CFPS. To shed some light on this, we collect prefecture-year panel data during 2003-2009 about the local economy, which including the local GDP, local salary level, government revenue, government expenditure, and sanitary conditions such as number of registered doctors and number of beds in local hospital.[38]

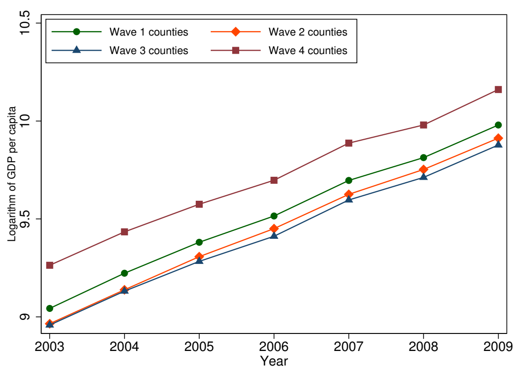
Then we match the data to the counties and plot the macro economy indexes over the calendar years in Figure 5.4, by which wave the county started the NRPS. Panel A shows the pattern for logarithm of GDP per capita. Consistent with our hypothesis that the NRPS started in poorer regions first, given the GDP level of first three waves counties is lower than that of the last wave counties. However, the trends are very parallel. We also conduct a regression with interactions between the year and county groups dummies, and the F-test cannot reject the null hypothesis for the interactions (F-statistic = 0.19, P-value = 0.99). The similar patterns are also found for the other outcomes, including salary, government revenue and expenditure, and quantity of doctors and beds in hospitals. These patterns suggest that the counties starting the NRPS in different years actually have no significant differences in trends of local economy statistics.

The mortality record data since 2005 provides four years before the starting year of

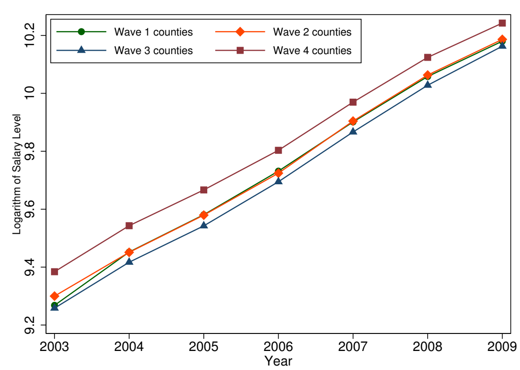
[38]Prefecture is one level higher than county according to administrative system in China. The data collect the balanced panel in 2003-2009 from local economy from 279 prefectures (97 percent of all counties) in mainland China.

Figure 5.4: *Pre-trends Examination in counties, by Starting Year of NRPS*

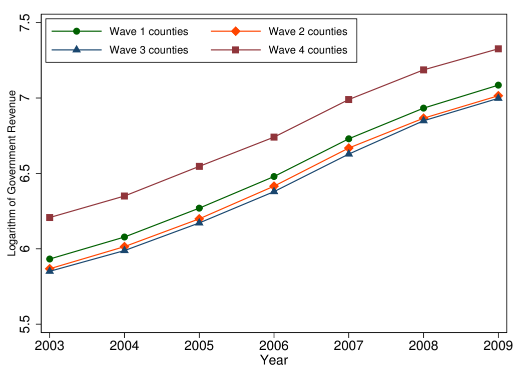
(a) *Log(GDP per capita)*



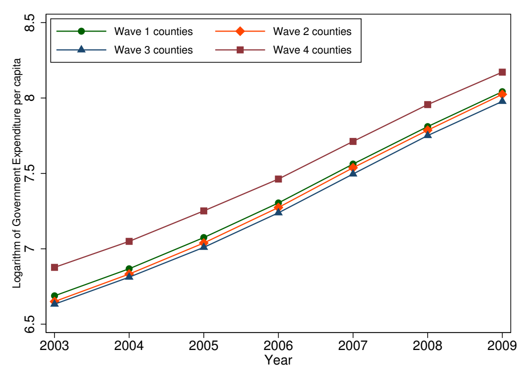
(b) *Log(Salary)*



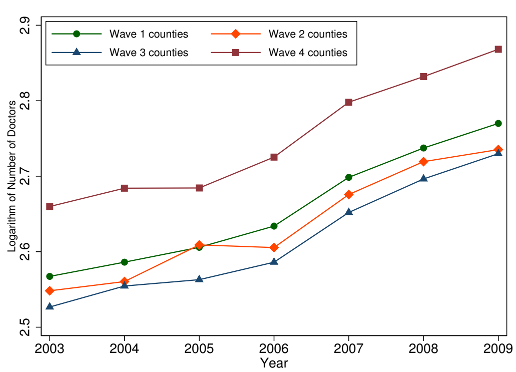
(c) *Log(Government Revenue)*



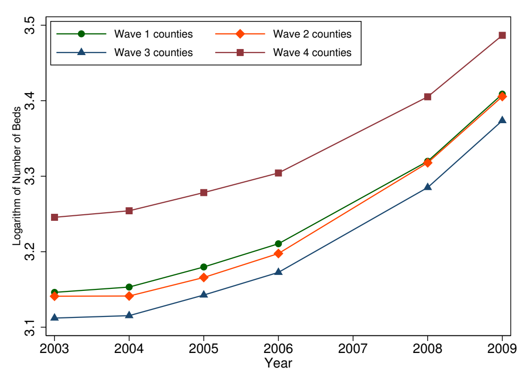
(d) *Log(Government Expenditure)*



(e) *Log(Number of Doctors)*

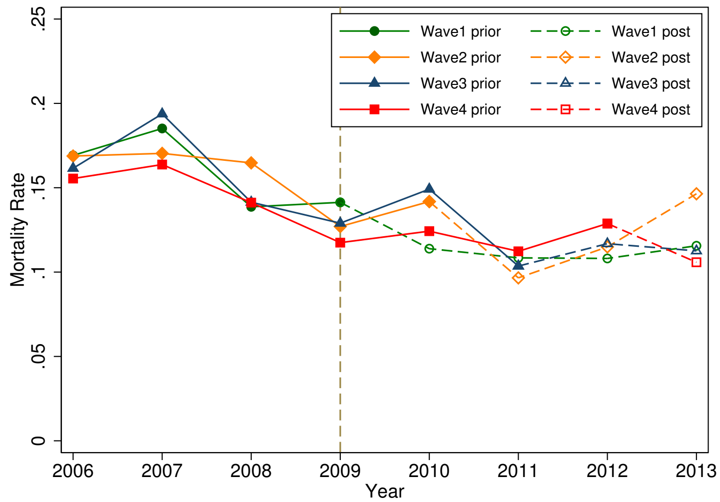


(f) *Log(Number of Beds in Hospitals)*



Notes: The economic indexes from different prefectures are from China City Statistics Yearbook 2004-2010. The prefectures are grouped by the different starting years of the NRPS. Each figure plots the mean values of the logarithm of economic indexes over the calendar years from 2003 to 2009.

Figure 5.5: *Mortality Time Trends in Counties, by NRPS Starting Year*



Notes: The mortality data are from CLHLS 2005-2014. The sample are divided by the different starting years of the NRPS. For each subsample, we plot the mean of mortality against the calendar year, with the solid lines for the period prior to the NRPS while dash lines for the period after the NRPS.

NRPS program. We thus also plot the mortality rates over each year since 2006 in Figure 5.5, by the counties with different starting years. Prior to 2009, there is no obvious different trends across the different groups. We also conduct a regression for the sample prior to 2009, and the joint F-test cannot reject the null hypothesis (F-statistic = 1.48, P-value = 0.15). After 2009, the mortality usually drops most when the county was just covered by NRPS. For example, the mortality of first wave counties dropped by 1.8 percentage points from 13.8 to 12.0 percent between 2009 to 2010, while the mortality of all the other counties actually increased during the same period. We emphasize that it is not just by accident because the second wave and fourth wave counties followed the similar pattern.[39]

[39]For wave 2 counties, the mortality dropped by 4.36 percentage points from 2010 to 2011 and those for three other counties dropped 1.35, 4.12 and 1.82 percentage points; for wave 4 counties, the mortality dropped by 2.1 percentage points from 2012 to 2013.

5.4 Conclusions and Discussion

This paper examines the effects of social pension provision on the lives of elderly in terms of income, expenditure, private transfers, labor supply, health and mortality. To quantify the effects of pension program is important since they are key parameters to evaluate and design *efficient* pension programs and retirement policies.

We first conduct a cross-country analysis to investigate the effects of introduction of social pension in 10 countries on mortality. The RD results suggest that the mortality of age-eligible seniors dropped by around 2.0 percent just after the introduction of social pension, while that of those with ages below the pension age changed little and insignificant. These results from aggregated data suggest that social pension potentially has remarkable effects on the lives of the elderly and call for evidence from individual level data.

Using the recent pension program in rural China - New Rural Pension Scheme, we present evidence for the NRPS-induced mechanical effects, individual behavior responses, and health consequences. For mechanical effects, we only find rural people with ages 60 and over were more likely to receive pension just after the NRPS. Among the pension-eligible group, the NRPS increased the proportion of household pension receipts by 24.5 percentage points. Meanwhile, it increased the household income by 17.6 percent, food expenditure by 9.6 percent, and reduced labor supply by 3.0 percentage points (6.2 percent) and health insurance participation by 5.0 percentage points. Furthermore, the rates of reported disability and underweight reduced by 3.2 percentage points (11.4 percent) and 1.8 percentage points (11.3 percent) after NRPS implementation, respectively. Finally, analysis of an individual-year panel composed of those aged 65 in CLHLS and above shows that the implementation of NRPS reduced the mortality by 2.2 percentage points (14.4 percent). In contrast, among the ineligible groups, we do not find any significant effects on pension receipt, income, labor supply, health and mortality. But the only exceptions are that, among the *hukou*-eligible people with ages below 60, the NRPS shifted their labor supply from farm work to non-farm work, and also reduced the health insurance participation. One possible explanation could be the higher and more stable expected income because of the NRPS.

Exploiting a natural experiment in China,this paper systematically examined the effects of a new pension program provision on a series of outcomes and shed some light on the several on-going literatures. It also has some pitfalls, too. The first one is about the potential heterogenous pre-trends or non-randomness of the NRPS counties selection. Although the results for ineligible groups provides some suggestive evidence to alleviate this concern, the identified effects may still be biased due to the heterogenous trends across different counties. Because of data limitation, we provide further supportive evidence on this by showing the parallel trends in a series of macro economy indexes for counties with different starting years of the NRPS. The second is about the measurement errors of reported income and expenditure. As mentioned in previous literature (e.g., Moore and Welniak 2000; Bound *et al.* 2001; Meyer and Sullivan 2003 etc.), the reported income and expenditure suffer serious measure errors and the coefficients should be interpreted carefully. The third one is that CLHLS data may not be nationally representative, and thus we should be careful when interpret or generalize the results to the full population or cases under other settings.

References

Aaronson, D. and Mazumder, B. (2011). The impact of rosenwald schools on black achievement. *Journal of Political Economy*, **119** (5), 821–888.

Acemoglu, D. and Angrist, J. (2001). How large are human-capital externalities? evidence from compulsory-schooling laws. In *NBER Macroeconomics Annual 2000, Volume 15*, MIT Press, pp. 9–74.

Aizer, A., Eli, S., Ferrie, J. and Lleras-Muney, A. (Forthcoming). The long run impact of cash transfers to poor families. *American Economic Review*.

Albouy, V. and Lequien, L. (2009). Does compulsory education lower mortality? *Journal of health economics*, **28** (1), 155–168.

Anderson, S. and Bidner, C. (2015). Property rights over marital transfers. *The Quarterly Journal of Economics*, p. qjv014.

Angelucci, M. and De Giorgi, G. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *American Economic Review*, **99** (1), 486–508.

Ardington, C., Case, A. and Hosegood, V. (2009). Labor supply responses to large social transfers: Longitudinal evidence from south africa. *American Economic Journal: Applied Economics*, **1** (1), 22–48.

Ashenfelter, O. C. and Krueger, A. B. (1994). Estimates of the economic returns to schooling from a new sample of twins. *American Economic Review*, **84** (5), 1157–73.

Atalay, K. and Barrett, G. F. (2015). The impact of age pension eligibility age on retirement and program dependence: Evidence from an australian experiment. *Review of Economics and Statistics*, **97** (1), 71–87.

Attanasio, O. P. and Brugiavini, A. (2003). Social security and households' saving. *The Quarterly Journal of economics*, pp. 1075–1119.

— and Rohwedder, S. (2003). Pension wealth and household saving: Evidence from pension reforms in the united kingdom. *American Economic Review*, pp. 1499–1521.

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

— (2010). " momma's got the pill": How anthony comstock and griswold v. connecticut shaped us childbearing. *American Economic Review*, **100** (1), 98–129.

BANISTER, J. (1991). *China's changing population*. Stanford University Press.

BAOCHANG, G., FENG, W., ZHIGANG, G. and ERLI, Z. (2007). China's local and national fertility policies at the end of the twentieth century. *Population and Development Review*, pp. 129–147.

BECKER, G. S. (1973). A theory of marriage: Part i. *The Journal of Political Economy*, pp. 813–846.

— (1974). A theory of marriage: Part ii. *The Journal of Political Economy*, pp. S11–S26.

BERNHEIM, B. D. *et al.* (2002). Taxation and saving. *Handbook of Public Economics*, **3**, 1173–1249.

BERTRAND, M., DUFLO, E. and MULLAINATHAN, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, **119** (1).

BITLER, M. P., GELBACH, J. B. and HOYNES, H. W. (2005). Welfare reform and health. *Journal of Human resources*, **40** (2), 309–334.

—, —, — and ZAVODNY, M. (2004). The impact of welfare reform on marriage and divorce. *Demography*, **41** (2), 213–236.

BORJAS, G. J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *The American Economic Review*, pp. 365–390.

BOTTICINI, M. and SIOW, A. (2003). Why dowries? *American Economic Review*, pp. 1385–1398.

BOUND, J., BROWN, C. and MATHIOWETZ, N. (2001). Measurement error in survey data. *Handbook of econometrics*, **5**, 3705–3843.

BRUNELLO, G., FABBRI, D. and FORT, M. (2013). The causal effect of education on body mass: Evidence from europe. *Journal of Labor Economics*, **31** (1), 195–223.

BURSZTYN, L. and JENSEN, R. (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics*, **130** (3), 1329–1367.

BUSSE, M. R., KNITTEL, C. R. and ZETTELMEYER, F. (2013). Are consumers myopic? evidence from new and used car purchases. *The American Economic Review*, **103** (1), 220–256.

CALONICO, S., CATTANEO, M. D. and TITIUNIK, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, **82** (6), 2295–2326.

CARD, D., DOBKIN, C. and MAESTAS, N. (2008). The impact of nearly universal insurance coverage on health care utilization: Evidence from medicare. *American Economic Review*, **98** (5), 2242–58.

—, — and — (2009). Does medicare save lives?*. *The Quarterly journal of economics*, **124** (2), 597–636.

CARLSSON, M., DAHL, G. B., ÖCKERT, B. and ROOTH, D.-O. (2012). The effect of schooling on cognitive skills. *Review of Economics and Statistics*, (00).

Case, A. (2004). *Does Money Protect Health Status? Evidence from South African Pensions*, University of Chicago Press, pp. 287–312.

— and Deaton, A. (1998). Large cash transfers to the elderly in south africa. *Economic Journal*, pp. 1330–1361.

— and Wilson, F. (2000). Health and well-being in south africa: evidence from the langeberg survey. *Research Program in Development Studies, Princeton, NJ (December). Processed*.

Chaloupka, F. J. and Warner, K. E. (2000). The economics of smoking. *Handbook of health economics*, **1**, 1539–1627.

Chan, S. and Stevens, A. H. (2004). Do changes in pension incentives affect retirement? a longitudinal study of subjective retirement expectations. *Journal of Public Economics*, **88** (7), 1307–1333.

Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy*, **116** (2), 173–234.

— (2009a). Is the taxable income elasticity sufficient to calculate deadweight loss? the implications of evasion and avoidance. *American Economic Journal: Economic Policy*, **1** (2), 31–52.

— (2009b). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, **1** (1), 451–488.

Chiappori, P.-A., Fortin, B. and Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. *Journal of Political Economy*, **110** (1), 37–72.

—, Iyigun, M. and Weiss, Y. (2009). Investment in schooling and the marriage market. *American Economic Review*, **99** (5), 1689–1713.

— and Oreffice, S. (2008). Birth control and female empowerment: An equilibrium analysis. *Journal of Political Economy*, **116** (1), 113–140.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, **114** (1), 175–201.

Ciccone, A. and Peri, G. (2006). Identifying human-capital externalities: Theory with applications. *The Review of Economic Studies*, **73** (2), 381–412.

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from britain. *American Economic Review*, **103** (6), 2087–2120.

Coile, C. and Gruber, J. (2000). *Social security and retirement*. Tech. rep., National Bureau of Economic Research.

Conti, G., Heckman, J. and Urzua, S. (2010). The education-health gradient. *The American economic review*, **100** (2), 234.

Currie, J. and Vogl, T. (2013). Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, **5** (1), 1–36.

CUTLER, D. M., GLAESER, E. L. and SHAPIRO, J. M. (2003). Why have americans become more obese? *The Journal of Economic Perspectives*, **17** (3), 93.

— and JOHNSON, R. (2004). The birth and growth of the social insurance state: explaining old age and medical insurance across countries. *Public Choice*, **120** (1-2), 87–121.

— and LLERAS-MUNEY, A. (2010). Understanding differences in health behaviors by education. *Journal of health economics*, **29** (1), 1–28.

— and — (2012). *Education and health: insights from international comparisons*. Tech. rep., National Bureau of Economic Research.

DEATON, A. (2003). Health, income and inequality. *National Bureau of Economic Research Reporter: Research Summary. Retrieved August*, **15**, 2009.

DUFLO, E. (2000). Child health and household resources in south africa: Evidence from the old age pension program. *American Economic Review*, pp. 393–398.

— (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review*, **17** (1), 1–25.

— (2012). Women empowerment and economic development. *Journal of Economic Literature*, **50** (4), 1051–79.

DUPUY, A. and GALICHON, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, **122** (6), 1271–1319.

EBENSTEIN, A. (2010). The missing girls of china and the unintended consequences of the one child policy. *Journal of Human Resources*, **45** (1), 87–115.

EDLUND, L. (2000). The marriage squeeze interpretation of dowry inflation: a comment. *Journal of Political Economy*, **108** (6), 1327–1333.

EDMONDS, E. V. (2006). Child labor and schooling responses to anticipated income in south africa. *Journal of Development Economics*, **81** (2), 386–414.

EVANS, W. N. and MOORE, T. J. (2011). The short-term mortality consequences of income receipt. *Journal of Public Economics*, **95** (11), 1410–1424.

— and — (2012). Liquidity, economic activity, and mortality. *Review of Economics and Statistics*, **94** (2), 400–418.

FANG, H., EGGLESTON, K. N., RIZZO, J. A., ROZELLE, S. and ZECKHAUSER, R. J. (2012). *The returns to education in China: Evidence from the 1986 compulsory education law*. Tech. rep., National Bureau of Economic Research.

FEI, X. (1979). *Modernization and National Minorities in China; China's National Minorities: An Introductory Survey; Ethnic Identification in China*. McGill University, Centre for East Asian Studies.

FIELD, E. and AMBRUS, A. (2008). Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of Political Economy*, **116** (5), 881–930.

Fischbein, S. (1977). Intra-pair similarity in physical growth of monozygotic and of dizygotic twins during puberty. *Annals of Human Biology*, **4** (5), 417–430.

Fletcher, J. M. (2015). New evidence of the effects of education on health in the us: Compulsory schooling laws revisited. *Social Science & Medicine*, **127**, 101–107.

Fogel, R. W. (1994). Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy1. *The American Economic Review*, **84** (3), 369–395.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, **72** (2), 395–427.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A. (2005). The causal effect of income on health: Evidence from german reunification. *Journal of Health Economics*, **5** (24), 997–1017.

Fryer, R. G. (2007). Guess who's been coming to dinner? trends in interracial marriage over the 20th century. *Journal of Economic Perspectives*, **21** (2), 71–90.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, **96** (2), 1–21.

— and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy*, **110** (4), 730–770.

— and — (2009). *The race between education and technology*. Harvard University Press.

—, — and Kuziemko, I. (2006). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, **20** (4), 133–156.

Greenhalgh, S. (1986). Shifts in china's population policy, 1984-86: Views from the central, provincial, and local levels. *Population and Development Review*, pp. 491–515.

— and Winckler, E. A. (2005). *Governing China's population: From Leninist to neoliberal biopolitics*. Stanford University Press.

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy*, **80** (2), 223–255.

Gustman, A. L. and Steinmeier, T. L. (2015). Effects of social security policies on benefit claiming, retirement and saving. *Journal of Public Economics*, **129**, 51–62.

Hanushek, E. A. (2013). Economic growth in developing countries: The role of human capital. *Economics of Education Review*, **37**, 204–212.

— and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of economic literature*, pp. 607–668.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the national Academy of Sciences*, **104** (33), 13250–13255.

— (2010). Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature*, **48** (2), 356–398.

Hendren, N. (2013). *The Policy Elasticity*. Tech. rep., National Bureau of Economic Research.

Hoynes, H., Miller, D. L. and Schaller, J. (2012). Who suffers during recessions? *The Journal of Economic Perspectives*, **26** (3), 27–47.

Huang, W., Lei, X., Ridder, G., Strauss, J. and Zhao, Y. (2013). Health, height, height shrinkage, and ses at older ages: evidence from china. *American Economic Journal: Applied Economics*, **5** (2), 86–121.

—, — and Zhao, Y. (Forthcoming). One-child policy and the rise of man-made twins. *Review of Economics and Statistics*.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, pp. 21–37.

Jacoby, H. G. and Mansuri, G. (2010). "watta satta": Bride exchange and women's welfare in rural pakistan. *The American Economic Review*, pp. 1804–1825.

Jayachandran, S. and Lleras-Muney, A. (2009). Life expectancy and human capital investments: Evidence from maternal mortality declines. *The Quarterly Journal of Economics*, **124** (1), 349–397.

Jensen, R. (2000). Agricultural volatility and investments in children. *American Economic Review*, **90** (2), 399–404.

— (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, **125** (2), 515–548.

— (2012). Do labor market opportunities affect young women's work and family decisions? experimental evidence from india. *The Quarterly Journal of Economics*, **127** (2), 753–792.

— and Lleras-Muney, A. (2012). Does staying in school (and not working) prevent teen smoking and drinking? *Journal of Health Economics*, **31** (4), 644–657.

Jensen, R. T. (2004). Do private transfers displace the benefits of public transfers? evidence from south africa. *Journal of Public Economics*, **88** (1), 89–112.

— and Richter, K. (2004). The health implications of social security failure: evidence from the russian pension crisis. *Journal of Public Economics*, **88** (1), 209–236.

Jia, R. and Persson, T. (2015). Individual vs. social motives in identity choice: Theory and evidence from china.

Kalmijn, M. (1991). Shifting boundaries: Trends in religious and educational homogamy. *American Sociological Review*, pp. 786–800.

Kaup, K. P. (2000). *Creating the Zhuang: ethnic politics in China*. Lynne Rienner Publishers.

Kemptner, D., Jürges, H. and Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: Evidence from germany. *Journal of Health Economics*, **30** (2), 340–354.

Lafortune, J. (2013). Making yourself attractive: Pre-marital investments and the returns to education in the marriage market. *American Economic Journal: Applied Economics*, **5** (2), 151–178.

Lager, A. C. J. and Torssander, J. (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million swedes. *Proceedings of the National Academy of Sciences*, **109** (22), 8461–8466.

Leisering, L., Gong, S. and Hussain, A. (2002). *People's Republic of China: Old-age Pensions for the Rural Areas: from Land Reform to Globalization*. Asian Development Bank.

Li, H., Yi, J. and Zhang, J. (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.

—, Zhang, J. and Zhu, Y. (2008). The quantity-quality trade-off of children in a developing country: Identification using chinese twins. *Demography*, **45** (1), 223–243.

Lleras-Muney, A. (2002). Were compulsory attendance and child labor laws effective? an analysis from 1915 to 1939*. *Journal of Law and Economics*, **45** (2), 401–435.

— (2005). The relationship between education and adult mortality in the united states. *The Review of Economic Studies*, **72** (1), 189–221.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. and Sanbonmatsu, L. (2012). Neighborhood effects on the long-term well-being of low-income adults. *Science*, **337** (6101), 1505–1510.

Ma, R. (2007). A new perspective in guiding ethnic relations in the twenty-first century: de-politicization of ethnicity in china. *Asian Ethnicity*, **8** (3), 199–217.

Madrian, B. C. and Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior*. *The Quarterly journal of economics*, **116** (4), 1149–1187.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, pp. 1329–1376.

McElroy, M. and Yang, D. T. (2000). Carrots and sticks: fertility effects of china's population policies. *American Economic Review*, pp. 389–392.

Meyer, B. D. and Sullivan, J. X. (2003). Measuring the well-being of the poor using income and consumption. *Journal of Human Resources*, **38** (4).

Miller, G. (2010). Contraception as development? new evidence from family planning in colombia*. *The Economic Journal*, **120** (545), 709–736.

— and Babiarz, K. S. (2014). *Family Planning: Program Effects*. Tech. rep., National Bureau of Economic Research.

Moore, J. C. and Welniak, E. J. (2000). Income measurement error in surveys: A review. *Journal of Official Statistics*, **16** (4), 331.

Oster, E., Shoulson, I. and Dorsey, E. (2013). Limited life expectancy, human capital and health investments. *American Economic Review*, **103** (5), 1977–2002.

Poston Jr, D. L. and Shu, J. (1987). The demographic and socioeconomic composition of china's ethnic minorities. *Population and Development Review*, pp. 703–722.

Poterba, J., Venti, S. and Wise, D. A. (2013). Health, Education, and the Postretirement Evolution of Household Assets. *Journal of Human Capital*, **7** (4), 297 – 339.

Qian, N. (2009). *Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural China*. Tech. rep., National Bureau of Economic Research.

Rao, V. (1993). The rising price of husbands: A hedonic analysis of dowry increases in rural india. *Journal of political Economy*, pp. 666–677.

Rosenzweig, M. R. and Wolpin, K. I. (1980a). Life-cycle labor supply and fertility: Causal inferences from household models. *The Journal of Political Economy*, pp. 328–348.

— and — (1980b). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica: journal of the Econometric Society*, pp. 227–240.

— and Zhang, J. (2009). Do population control policies induce more human capital investment? twins, birth weight and china's one-child policy. *The Review of Economic Studies*, **76** (3), 1149–1174.

Rossing, M. A., Daling, J. R., Weiss, N. S., Moore, D. E. and Self, S. G. (1994). Ovarian tumors in a cohort of infertile women. *New England Journal of Medicine*, **331** (12), 771–776.

Ruhm, C. J. (2000). Are recessions good for your health? *The Quarterly Journal of Economics*, **115** (2), 617–650.

Sautman, B. (1998). Preferential policies for ethnic minorities in china: The case of xinjiang. *Nationalism and Ethnic Politics*, **4** (1-2), 86–118.

Schultz, T. P. and Zeng, Y. (1995). Fertility of rural china. effects of local family planning and health programs. *Journal of Population Economics*, **8** (4), 329–350.

Segal, N. L. (1985). Monozygotic and dizygotic twins: A comparative analysis of mental ability profiles. *Child Development*, pp. 1051–1058.

— (1999). *Entwined lives: Twins and what they tell us about human behavior.* Dutton/Penguin Books.

Shi, S.-J. (2006). Left to market and family–again? ideas and the development of the rural pension policy in china. *Social Policy & Administration*, **40** (7), 791–806.

Smith, D. M., Nance, W. E., Kang, K. W., Christian, J. C. and Johnston Jr, C. C. (1973). Genetic factors in determining bone mass. *Journal of Clinical Investigation*, **52** (11), 2800.

Snyder, S. E. and Evans, W. N. (2006). The effect of income on mortality: evidence from the social security notch. *The Review of Economics and Statistics*, **88** (3), 482–495.

Starr, B. (2008). Ask a geneticist. *Available from Stanford School of Medicine The Tech: http://www. thetech. org/genetics/ask. php*.

Staubli, S. and Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers? *Journal of public economics*, **108**, 17–32.

Stephens, M. and Yang, D.-Y. (2014). Compulsory education and the benefits of schooling. *American Economic Review*, **104** (6), 1777–1792.

Stunkard, A. J., Foch, T. T. and Hrubec, Z. (1986). A twin study of human obesity. *JAMA: The Journal of the American Medical Association*, **256** (1), 51–54.

Thomas, D., Strauss, J. and Henriques, M.-H. (1991). How does mother's education affect child height? *Journal of Human Resources*, pp. 183–211.

UN, P. D. (2006). *World Contraceptive Use 2005.* New York: United Nations. ISBN 92-1-151418-5.

van Kippersluis, H., ODonnell, O. and Van Doorslaer, E. (2011). Long-run returns to education does schooling lead to an extended old age? *Journal of human resources*, **46** (4), 695–721.

Wagstaff, A., Lindelow, M., Jun, G., Ling, X. and Juncheng, Q. (2009). Extending health insurance to the rural population: An impact evaluation of china's new cooperative medical scheme. *Journal of Health Economics*, **1** (28), 1–19.

Wang, D. (2006). China's urban and rural old age security system: Challenges and options. *China & World Economy*, **14** (1), 102–116.

Wang, F. *et al.* (2012). Family planning policy in china: Measurement and impact on fertility. *Munich Personal RePEC Archive*, (42226).

Wantchekon, L., Klašnja, M. and Novta, N. (2015). Education and human capital externalities: Evidence from colonial benin. *The Quarterly Journal of Economics*, **130** (2), 703–757.

Wasserman, J., Manning, W. G., Newhouse, J. P. and Winkler, J. D. (1991). The effects of excise taxes and regulations on cigarette smoking. *Journal of health economics*, **10** (1), 43–64.

Wei, S.-J. and Zhang, X. (2011). The competitive saving motive: Evidence from rising sex ratios and savings rates in china. *Journal of Political Economy*, **119** (3), 511–564.

Zimmermann, A. C. and Easterlin, R. A. (2006). Happily ever after? cohabitation, marriage, divorce, and happiness in germany. *Population and Development Review*, **32** (3), 511–528.

Appendix A

Appendix to Chapter 1

A.1 Background of One-Child Policy and Fertility Fines

At the very beginning of China's one-child policy, vice premier Chen Muhua proposed that it would be necessary to pass new legislation imposing an extra child "tax" on excess children. Considering the "local difficulties", the central government authorized the provincial governments to determine their "tax rates" according to their localities. The provincial leaders were facing strong pressures from both the central government and the masses. A high rate can decrease fertility more effectively but also would invoke more mass resistance and complaints. The subnational leaders, rather than the central government, would face the difficulties in collecting high penalties and bear the costs associated with resistance and complaints directly. For example, only in the year 1984, Guangdong province received more than 5000 complaint letters. The choices of penalty rate mainly reflected the trade-offs made by the provincial leaders. If they cared more about the evaluation by central government than mass approval, they would prefer high penalty rates. Therefore, a change in fine rate can be thought as the move from original equilibrium to new equilibrium when the attitudes of provincial leaders or external conditions have significantly changed.

The central government did not fully realize potential agency problems in the enforcement of the OCP until late 1980s. The there were very few changes in the fine rates in these years. One promoting mechanism was formally established in the year 1989, when the central government was designing the Eighth Five-Year Plan and effective birth control was among the major objectives. As Greenhalgh and Winckler (2005) wrote in the book *Governing China's Population: Addressing governors in spring 1989 Li Peng (current premier) said that population remained in a race with grain, the outcome of which would affect the survival of the Chinese race. To achieve subnational compliance, policy must be supplemented with more detailed management by objectives (ME 890406). At a meeting on birth policy in the premier's office, Li Peng explained that such targets would be neither "mandatory" nor "guidance" but "evaluative".* Therefore, the national leaders employed a management-objective "responsibility system" to induce subnational leaders to induce them to set high fine rates and compel local cadres to enforce. In this system, a promoting mechanism was created for making that compliance central to provincial leaders' career prospects. In the later few years, there are many provinces started to increase their fines.

The newly appointed provincial governors had stronger incentives to implement tough measures of the OCP because of delay in policy implementation and age limits in promotion. Other factors affecting promotion concerns include gender, education and connection to senior leaders. Take the event of Li Peng's address as an example, we find that 12 provinces increased their fine rates in a period of three years after the address, on average from 0.84 yearly household incomes to 3.07 yearly household incomes. The increases in fine rate in this period were significantly larger than any other period. Moreover, 7 out of these 12 province governors, who increased fine rates in this period, made such a policy change exactly in the first years of their tenures . The average age of these 12 province governors was 55.75 years old, which was lower than the average age of other provincial governors, 57.88 years old. In general, these evidences show that the personal interest of provincial governor plays a crucial role in the change of fine rate.

Although the promotion of local provincial governors is relevant to the performance of OCP, we did not find significant difference in tenure lengths of the provincial governors across provinces with different levels of fertility fine rates.

Appendix B

Appendix to Chapter 2

B.1 Solving the equilibrium

Based on the equilibrium conditions in Section 3.1, plug in the two ethnicities H and M, and explicitly express the equations by ethnicities. We have the following set of equations:

$$\begin{cases} ln\mu_{HH} - ln\mu_{H0} = \dfrac{\alpha_{HH} + \gamma_{HH}}{2} \\ ln\mu_{HM} - \frac{ln\mu_{H0} + ln\mu_{M0}}{2} = \dfrac{\alpha_{HM} + \gamma_{HM}}{2} \\ ln\mu_{MM} - ln\mu_{M0} = \dfrac{\alpha_{MM} + \gamma_{MM}}{2} \\ \mu_{H0} + \mu_{HH} + \mu_{MH} = \overline{H} \\ \mu_{M0} + \mu_{MH} + \mu_{MM} = \overline{M} \end{cases}$$

For simplicity, we define $\theta_{HH} = \dfrac{\alpha_{HH} + \gamma_{HH}}{2}$, $\theta_{HM} = \dfrac{\alpha_{HM} + \gamma_{HM}}{2}$ and $\theta_{MM} = \dfrac{\alpha_{MM} + \gamma_{MM}}{2}$, which are the expected marriage gains for the H-H, H-M and M-M couples, respectively. Then we translate the equations above into proportions and rates:

$$
\begin{cases}
ln(h_m r_H^H) - lnh_0 = \theta_{HH} \\
ln\overline{H}h_m r_H^M - \frac{1}{2}(ln\overline{H}h_0 + ln\overline{M}m_0) = \theta_{HM} \\
ln(m_m r_M^M) - lnm_0 = \theta_{MM} \\
\overline{H}h_m r_H^M = \overline{M}m_m r_M^H \\
h_m + h_0 = m_m + m_0 = r_H^H + r_H^M = r_M^M + r_M^H = 1
\end{cases}
$$

where h_m, h_0 are the married and unmarried rates for Han ethnicity; and m_m, m_0 are married and unmarried rates for minorities. Similarly, r_H^H and r_H^M are the proportion of married Han people marrying to Han and minorities, respectively; r_M^H and r_M^M are the proportion of married minority people marrying to Han and minorities, respectively. The first three equations are directly from the first three in (1). The fourth one means that the number of Han people involved in H-M marriages are the same with that of Minorities involved.

Then we take derivatives with f and note that $\frac{\partial\theta_{HH}}{\partial f} = u'_{HH}$, $\frac{\partial\theta_{HM}}{\partial f} = u'_{HM}$, $\frac{\partial\theta_{MM}}{\partial f} = u'_{MM}$, and $\frac{dr_M^M}{df} = -\frac{dr_M^H}{df}$, $\overline{H}h_m r_H^M = \overline{M}m_m r_M^H$, we have

$$
\begin{cases}
(\frac{1}{h_m} + \frac{1}{h_0})e_h - \frac{1}{r_H^H}e_H^M = u'_{HH} \\
-\frac{r_M^H}{h_m r_M^M}e_h - \frac{r_M^H}{r_H^M r_M^M}e_H^M + (\frac{1}{m_0} + \frac{1}{m_m} + \frac{r_M^H}{m_m r_M^M})e_m = u'_{MM} \\
(\frac{1}{h_m} + \frac{1}{2h_0})e_h + \frac{1}{r_H^M}e_H^M + \frac{1}{2m_0}e_m = u'_{HM}
\end{cases}
$$

where $e_h = \frac{dh_m}{df}$, $e_m = \frac{dm_m}{df}$, and $e_H^M = \frac{dr_H^M}{df}$. The first two are the responses of married rates of Han and Minorities to one unit increase in the fertility fines; the last one represents the response of the H-M marriage rate among the Han ethnicity with respective to the fertility fines. We can solve these three equations above to derive the expressions in terms of u'_{HH}, u'_{HM} and u'_{MM} for the three unknowns.

We first define $\alpha_1 = (\frac{1}{h_m} + \frac{1}{h_0})$, $\alpha_2 = \frac{1}{r_H^H}$, $\alpha_3 = \frac{r_M^H}{h_m r_M^M}$, $\alpha_4 = \frac{r_M^H}{r_H^M r_M^M}$, $\alpha_5 = (\frac{1}{m_0} + \frac{1}{m_m} +$

$\dfrac{r_M^H}{m_m r_M^M})$, $\alpha_6 = (\dfrac{1}{h_m} + \dfrac{1}{2h_0})$, $\alpha_7 = \dfrac{1}{r_H^M}$ and $\alpha_8 = \dfrac{1}{2m_0}$. Obviously, $\alpha_i > 0$, $\forall i$.

By solving the the equations, we have,

$$e_h = \frac{Au'_{HH} + \alpha_2 C}{\alpha_1 A + \alpha_2 B} \tag{B.1}$$

$$e_H^M = \frac{-Bu'_{HH} + \alpha_1 C}{\alpha_1 A + \alpha_2 B} \tag{B.2}$$

$$e_m = \frac{u'_{MM} + \alpha_3 e_h + \alpha_4 e_H^M}{\alpha_5}\ (or\ = \frac{u'_{HM} - \alpha_6 e_h - \alpha_7 e_H^M}{\alpha_8}) \tag{B.3}$$

where $A = \alpha_5\alpha_7 + \alpha_4\alpha_8$, $B = \alpha_5\alpha_6 + \alpha_3\alpha_8$, and $C = \alpha_5 u'_{HM} - \alpha_8 u'_{MM}$.

Because $\alpha_5 > \alpha_8 > 0$ and $u'_{HH} \leq u'_{HM} \leq u'_{MM} \leq 0.$, we have $C \leq 0$.

B.2 Proof of Predictions

Proof of Predictions 1: We have found that $e_h = \frac{Au'_{HH}+\alpha_2 C}{\alpha_1 A+\alpha_2 B}$ and thus it's easy to find that $e_h < 0$. Without the loss of generality, we can reasonably assume that, in the preferential-policy regions, the One-Child policy has very little impact on the welfare of H-M marriage and M-M marriage. That is, $u'_{HM} = u'_{MM} = 0$. Thus, the absolute value of e_h will be lower in the preferential-policy regions because $C = 0$ when $u'_{HM} = u'_{MM} = 0$.

From (3), we have $e_m = \frac{u'_{HM}-\alpha_6 e_h-\alpha_7 e_H^M}{\alpha_8}$. In the preferential-policy regions, the expression of e_m can be simplified as follow:

$$e_m = \frac{(\alpha_7 B - \alpha_6 A)u'_{HH}}{\alpha_8(\alpha_1 A + \alpha_2 B)} \tag{B.4}$$

By substituting $A = \alpha_5\alpha_7 + \alpha_4\alpha_8$ and $B = \alpha_5\alpha_6 + \alpha_3\alpha_8$, we have $\alpha_7 B - \alpha_6 A = (\alpha_3\alpha_7 - \alpha_4\alpha_6)\alpha_8$. Because $\alpha_3\alpha_7 - \alpha_4\alpha_6 = -\dfrac{r_M^H}{2h_0 r_H^M r_M^M} < 0$, $e_m > 0$ holds in the preferential-policy regions. That is, in these regions, the One-Child policy may have a positive effect on the marriage rate of minority people.

However, in the non-preferential-policy regions, whether e_m is positive or negative is inconclusive.

Proof of Prediction 2: From $\overline{H}h_m r_H^M = \overline{M}m_m r_M^H$, we have the expression of e_M^H as follow:

$$e_M^H = r_M^H(\frac{1}{h_m}e_h + \frac{1}{r_H^M}e_H^M - \frac{1}{m_m}e_m) \tag{B.5}$$

According to the formula (2), the sign of e_H^M is not generally determinate. The sign of e_M^H is indeterminate also because it's linear combination of e_h, e_m and e_H^M.

However, in the preferential-policy regions, we have $e_H^M = \frac{-Bu'_{HH}}{\alpha_1 A + \alpha_2 B} > 0$ because C=0. That is, in these regions, an increase of OCP penalty rate would increase the probability that a Han people choose to marry a minority people.

Moreover, in these regions, we can express e_M^H as follow by substituting formulas (1), (2) and (4):

$$e_M^H = \frac{r_M^H(\frac{1}{h_m}A - \frac{1}{r_H^M}B - \frac{1}{m_m}D)u'_{HH}}{\alpha_1 A + \alpha_2 B} \tag{B.6}$$

where $D = (\alpha_3\alpha_7 - \alpha_4\alpha_6)$. By substituting the values of a_i(s), we find that $(\frac{1}{h_m}A - \frac{1}{r_H^M}B - \frac{1}{m_m}D) = -\frac{1}{2h_0}(\frac{1}{m_m} + \frac{1}{m_0})\frac{1}{r_H^M} < 0$. Thus, $e_M^H > 0$ holds inthe preferential-policy regions.

Proof of Prediction 3: By definition, $\tau_{HM} = \frac{\ln \mu_{H0} - \ln \mu_{M0} + \alpha_{HM} - \gamma_{HM}}{2}$. We take derivatives and then have:

$$\frac{d\tau_{HM}}{df} = -\frac{1}{h_0}e_h + -\frac{1}{m_0}e_m \tag{B.7}$$

Here, $\dfrac{d\tau_{HM}}{df}$ is a linear combination of e_h and e_m. In non-preferential-policy regions, it's difficult to see the sign of $\dfrac{d\tau_{HM}}{df}$. However, in preferential-policy regions, it's obvious that the transfer from the Han spouse to the minority spouse is increasing in the fine rate because $e_h < 0$ and $e_m > 0$.

B.3 Welfare Implications

From the social welfare expressed as below,

$$\Pi = \sum_i \overline{m}_i ln(\sum_j exp(\tilde{\alpha}_{ij})) + \sum_j \overline{n}_j ln(\sum_i exp(\tilde{\gamma}_{ij})) + \sum_{i,j\neq 0} \mu_{ij} c_{ij} f. \tag{B.8}$$

We take derivative with respective to the fertility penalty f to the equation above. Denote that $P_{ij} = \frac{exp(\tilde{\alpha}_{ij})}{\sum_k exp(\tilde{\alpha}_{ik})}$ is the proportion of type i men married to type j women; correspondingly, $Q_{ij} = \frac{exp(\tilde{\gamma}_{ij})}{\sum_k exp(\tilde{\gamma}_{kj})}$ the proportion of type j women married to type i men. Then we have:

$$\frac{d\Pi}{df} = \sum_i \overline{m}_i \sum_j P_{ij} \frac{d\tilde{\alpha}_{ij}}{df} + \sum_j \overline{n}_j \sum_i Q_{ij} \frac{d\tilde{\gamma}_{ij}}{df} + \sum_{i,j\neq 0} \mu_{ij} c_{ij} + (\frac{d\mu_{ij}}{df} c_{ij} + \mu_{ij} \frac{dc_{ij}}{df}) f \tag{B.9}$$

Assuming the gains of being unmarried is not changed by the penalties, and considering that $\overline{m}_i P_{ij} = \overline{n}_j Q_{ij} = \mu_{ij}$ for given i, j, and $\frac{d\tilde{\alpha}_{ij}}{df} + \frac{d\tilde{\gamma}_{ij}}{df} = \frac{du_{ij}}{df} = -c_{ij}$, we have

$$\begin{aligned}
\frac{d\Pi}{df} &= -\sum_{i,j\neq 0} c_{ij}\mu_{ij} + + \sum_{i,j\neq 0} \mu_{ij} c_{ij} + \sum_{i,j\neq 0} (\frac{d\mu_{ij}}{df} c_{ij} + \mu_{ij} \frac{dc_{ij}}{df}) f \\
&= \sum_{i,j\neq 0} (\frac{d\mu_{ij}}{df} c_{ij} + \mu_{ij} \frac{dc_{ij}}{df}) f
\end{aligned} \tag{B.10}$$

Divide the both sides by $\overline{H} + \overline{M}$, we can have the equation in the main text.

Appendix C

Appendix to Chapter 3

Twins Height Difference and One-Child Policy

In this part, we derive testable implications for two different hypotheses, the real "man-made" twins hypothesis and the false twins hypothesis. For simplicity, we only use financial penalties *(fines)* to measure OCP and assume it equals one when a financial penalty policy is established in the local province and zero otherwise.

C.1 Real "Man-Made" Twins Hypothesis

The man-made twins hypothesis indicates that individuals are motivated to use technology, such as fertility drugs, to give birth to twins under OCP. Taking fertility drugs should be reasonable, but embryo technologies did not appear in China until the late 1990s. Fertility drugs are usually used to induce ovulation in women with an infertility problem (Rossing *et al.*, 1994). When a woman takes a fertility drug, the possibility of multiple ovulations and thus the likelihood of having twins is raised (Starr, 2008). Fertility drugs are classified as prescription medicines, but some people may purchase them in certain private hospitals or obtain prescriptions from certain doctors in an illegal way, such as bribing the doctors.

Under Man-Made Twins Hypothesis, individuals are more likely to take fertility drugs (*Take* = 1) to have more children while avoid being punished under one-child policy, that is, $Pr(Take = 1|Fine = 1) > Pr(Take = 1|Fine = 0)$. According to the medical literature and realized facts, we know that taking certain fertility drugs or using some technologies really increase the probability of giving birth to twins:$Pr(Twins = 1|Take = 1) > Pr(Twins = 1|Take = 0)$. We also assume that conditional on individuals' behaviors (*Take*), giving birth to twins is independent of one-child policy. In addition, the biological results of fertility drugs or embryo technologies is to make multiple zygotes developed in the uterus at the same time, rather than to stimulate a single fertilized egg in the mother's body to divide into two or more embryos. Thus, these actions only raise probability of DZ twins rather than that of MZ twins, that is, $Pr(DZ = 1|Take = 1) > Pr(DZ = 1|Take = 0)$ and $Pr(MZ = 1|Take = 1) = Pr(MZ = 1|Take = 0)$.

According to the medical literature, MZ twins are genetically nearly identical and they are always of the same sex unless there has been a mutation during development. But it is possible that same-gender twins are DZ Certain characteristics of MZ twins become more alike as twins age, such as IQ and personality (Segal, 1999). DZ twins, however, like any other siblings, have an extremely small chance of having the same chromosome profile. DZ twins may look very different from each other, and may be of different sexes or the same sex. The above also holds for brothers and sisters from the same parents, meaning that DZ twins can be viewed as siblings who happen to be of the same age. Therefore, we have$Pr(DZ = 1|DG = 1) = 1$, $Pr(SG = 1|MZ = 1) = 1$, $0 < Pr(DZ = 1|SG = 1) < 1$ and $0 < Pr(MZ|SG = 1) < 1$.

An established strand of literature has proved that, DZ twins tend to have more differences than MZ ones as they grow up because of genetic disparity, including height and bone mass (Smith *et al.*, 1973), weight (Stunkard *et al.*, 1986), mental ability profiles (Segal, 1985), and so on. Specifically, Fischbein (1977) found that MZ twins have a significantly higher concordance in height than for DZ pairs during puberty, for both boys and girls, and yearly height increments are also more similar for the MZ pairs, indicating that the height spurt occurs more simultaneously for MZ twins in comparison to DZ twins. Thus, we presume that the height difference (HD) within a DZ (same-gender) pair should be larger than that of a MZ pair if other factors are equalized. Thus, $E(HD|SG=1,DZ=1) > E(HD|SG=1,MZ=1)$.

Additionally, we also assume that the actions people may take do not influence twins height difference conditional on these twins type (MZ or DZ). Based on the facts or assumptions above, it can be shown that

(1) $E(HD|Fine=1,Twins=1) > E(HD|Fine=0,Twins=1)$,

(2) $E(HD|Fine=1,SG=1) > E(HD|Fine=0,SG=1)$,

and (3) $E(HD|Fine=1,DG=1) = E(HD|Fine=0,DG=1)$.

First, twins height difference would be larger under one-child policy because there will be more DZ twins due to the methods individuals take as response to one-child policy. This response will increase DZ proportion in same-gender twins, and thus the height difference in this group will enlarge. However, different-gender twins height difference will not change because they themselves are DZ. Because we do not know whether a woman took fertility drugs or not, or whether a pair of twins is MZ or DZ, the above findings suggest that we can use one-child policy fine and twins gender composition to conduct a series of empirical tests.

C.2 Fake Twins Hypothesis

Fake twins hypothesis means that parents report real siblings as twins so as to except from the punishment of one-child policy. It is somehow feasible under some special circumstances in earlier China. Firstly, many pregnant women gave birth to babies at home in the 1980s and population administration would not be noticed until the parents report the infants though they were required to do so. Second, birth certification did not launch until 1997, and the children's birth date was easy to revised before that. Third, children especially siblings look alike especially when the age difference is not large enough. Though parents would face more harsh punishment once they are found to report fake twins, many parents may still choose to do so because of strong children or boy preference.

Under fake twins hypothesis, one child policy stimulate people's incentive to report fake twins, that is,

$Pr(Twins^*|Fine = 1) > Pr(Twins^*|Fine = 0)$,

in which $Twins^*$ denotes the observed twins, including real ones and fake ones. For real twins ($Twins$), I assume all of them are reported, that is, $Pr(Twins^*|Twins) = 1$.

For simplicity, I do not consider the gender factor in height difference in different-gender twins here,[1] then because of age difference, the height difference within fake twins is supposed to be larger, so $E(HD|Twins^*) > E(HD|Twins)$ if $Pr(Twins|Twins^*) < 1$.

The condition $Pr(Twins|Twins^*) < 1$ ensures that there exist some fake twins. If parents have strong children preference and do not care about the gender, then the gender composition of fake twins should be random, so the height difference in both (observed) same-gender twins and different-gender ones should be larger. However, if parents have strong boy preference, and they report siblings as twins only if the first baby is girl and try to give birth to a boy in the next, then the height difference within (observed) different-gender twins is expected to be larger. No matter which case it is, under fake twins hypothesis, we must have

(4) $E(HD|Twins^*, Fine = 1) > E(HD|Twins^*, Fine = 0)$,

and (5) $E(HD|DG^*, Fine = 1) > E(HD|DG^*, Fine = 0)$,

in which DG^* denotes the observed different-gender twins. Same as before, equations (4) and (5) are all based on observables so that can be tested in empirical analysis. Same to man-made twins hypothesis, the height difference also is expected to be larger under one-child policy according to fake twins hypothesis. However, fake twins hypothesis predict that height difference within different-gender twins should be larger under one child policy, which is different from the results in C.1. Such a difference provides us an identification strategy to differentiate the two hypotheses.

[1]Conditional that boys may be taller than girls conditional on the same age, height gap, when defined as the taller one's height minus that of shorter one's, is possible to be larger or narrower if the proportion of fake twins increase, when those with boy preference are more likely to report fake twins if they gave birth to a girl firstly.

Appendix D

Appendix to Chapter 4

D.1 China Health and Nutrition Survey (CHNS)

The China Health and Nutrition Survey (CHNS) was designed to examine the effects of the health, nutrition, and family planning policies and programs implemented by national and local governments and to see how the social and economic transformation of Chinese society is affecting the health and nutritional status of its population. The survey takes place over a 3-day period using a multistage, random cluster process to draw a sample of about 4,400 households with a total of 26,000 individuals in nine provinces that vary substantially in geography, economic development, public resources, and health indicators. The CHNS data collection began in 1989 and has been implemented every two to four years since. The CHNS uses a multistage cluster sample design to survey individuals and households in 218 neighborhoods in nine provinces in China. These nine provinces contain approximately 56 percent of the population of China. The baseline sample was representative of each province, but over time loss-to-follow-up has occurred.

D.2 Chinese Family Panel Studies (CFPS)

CFPS is a biennial survey and is designed to be complementary to the Panel Study of Income Dynamics (PSID) in the United States. The first national wave was conducted under the collaboration of the Institute of Social Science Survey at Peking University and the Survey Research Center at the University of Michigan from April 2010 to August 2010. The five main parts of the questionnaire include communities, households, household members, adults and children data.

The 2010 round covered approximately 14,000 households in 25 provinces, in which 95% of the Chinese population reside. The population is divided into six subpopulation, i.e. five large provinces (Guangdong, Gansu, Liaoning, Henan, and Shanghai) and the other 20 provinces. The final sample is made to be representative of 25 provinces through careful weighting.

The survey sample was obtained by three-stage cluster sampling with unequal probabilities. In the first stage, 16 counties were sampled from each of the four large provinces and 32 township-level units from Shanghai, and 80 counties from 20 other provinces, with probability proportional to population size (PPS). In total there were 144 counties and 32 township-level units. In the second stage, 2 or 4 administrative villages or resident committees were sampled with PPS in each county or town. Together there were 640 villages or resident committees. In the third stage, 28-42 households were sampled from each village or resident committee, and in all there were about 16,000 households.

The national representative final sample covers 14,960 households and 33,600 adults (age 16+). The follow-up survey of CFPS was conducted in 2012, covers 13,448 households and 35729 adults, 12,724 households and 26,385 adults out of which originally covered in the baseline survey.

D.3 Chinese Household Income Project Series (CHIPS)

The purpose of the Chinese Household Income Project was to measure and estimate the distribution of personal income in the rural and urban areas of the People's Republic of China. Data were collected through a series of questionnaire-based interviews conducted in rural and urban areas in 1988, 1995, 2002, and 2007. Individual respondents reported on their economic status, employment, level of education, sources of income, household composition, and household expenditures. The study was interview-based. For each year, there are three different data sets for urban residents, rural residents, and migrants, separately. This study only uses the data for the residents. On average, each year there are more than over 20,000 individuals in the urban or rural survey. The data are coded in on-site observations through face-to-face interviews.

Appendix E

Appendix to Chapter 5

E.1 China Health and Retirement Longitudinal Studies (CHARLS)

The China Health and Retirement Longitudinal Study (CHARLS) aims to collect a high quality nationally representative sample of Chinese residents ages 45 and older to serve the needs of scientific research on the elderly. The baseline national wave of CHARLS is being fielded in 2011. The individuals will be followed up every two years. This study used the 2011 and 2013 two waves. In the base line survey, the sample was drawn in four stages. County-level units (counties or urban districts) were sampled directly. All county-level units in all provinces except for Tibet were stratified by 8 regions, by whether they were urban districts or rural counties, and by county GDP. They were sorted based on this stratification and 150 were randomly chosen proportional to population size. These counties cover 28 out of 30 provinces, other than Tibet. After the county units were chosen, the National Bureau of Statistics helped the CHARLS team to sample villages and communities within county units using recently updated village level population data. CHARLS sample used administrative villages in rural areas and neighborhoods, which comprise one or more formal resident committees, in urban areas as primary sampling units (PSUs). CHARLS then sampled three PSUs within each county-level unit, using PPS sampling, for a total of 450 PSUs. In each PSU, the CHARLS team constructed sampling frame using Google Earth base maps and a CAPI (computer assisted personal interview) program was then used to sample households and to conduct the interviews using laptops. All age-eligible sample households who were willing to participate in the survey were interviewed, with 10,257 households containing 18,245 respondents aged 45 and over and their spouses ultimately interviewed. The follow-up survey covers 10,979 households containing 19,666 respondent, with 16,159 (9,185) out of 18,245 (10,257) individuals (households) in the baseline survey successfully re-interviewed and 3507 individuals in 2,053 households newly interviewed. The main questionnaire includes information on basic demographics, family, health status, health care and health insurance, work, retirement and pension, and household economy (income, consumption and wealth).

E.2 Chinese Family Panel Studies (CFPS)

CFPS data have been described in Appendix to Chapter 4.

E.3 Chinese Longitudinal Healthy Longevity Survey (CLHLS)

The CLHLS is a longitudinal survey conducted by the Center for Healthy Aging and Family Studies in Peking University, sponsored and supported by the National Institute on Aging, United Nations, Duke University and Max Planck Institute for Demographic Research. Demographic and statistical methods are used to analyze data in the longitudinal surveys with the research goal of determining which factors, out of a large set of social, behavioral, biological, and environmental risk factors play an important role in healthy longevity.

The baseline survey was conducted in 1998, with follow-up surveys with replacements for deceased elders were conducted in 2000, 2002, 2005, 2008, 2011 and 2014 in a randomly selected half of the total number of counties and cities in the 22 out of 31 provinces in mainland China. The survey areas covered 1.1 billion people, 85 percent of the total population in China. An enumerator and a nurse or a medical school student conducted the interview and performed a basic health examination at each interviewee's home. We use data from the longitudinal datasets staring from the 2005 wave. The 2005 wave interviewed 15,638 individuals, with 4,955 young elderly aged 65-79 and 10,658 oldest-old aged 80+ (including 2,797 centenarians, 3,952 nonagenarians and 3,909 octogenarians), and another 25 elders who are younger than age 65.