Netspar

Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Early Life Adversities and Well-being Later in Life

Evidence from China

Jiyuan Wang

PhD 10/2020-002

Early Life Adversities and Well-being Later in Life: Evidence from China

Wang, Jiyuan

DOI:
10.33612/diss.134514701

Early Life Adversities and Well-being Later in Life: *Evidence from China*

Jiyuan Wang

Publisher: University of Groningen
Groningen, The Netherlands

Printed by: Ipskamp Printing
Enschede, The Netherlands

ISBN: 978-94-034-2896-3 (printed)
978-94-034-2897-0 (electronic)

university of groningen

Early Life Adversities and Well-being Later in Life: Evidence from China

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 22 October 2020 at 12.45 hours

by

Jiyuan Wang

born on 19 October 1990
in Gansu, China

Supervisors

Prof. R.J.M. Alessie
Prof. V. Angelini

Assessment Committee

Prof. G. Pasini
Prof. M.G. Knoef
Prof. B.W. Lensink

Dedicated to my dear mother Rongmei Li, and my beloved girlfriend Ya Gao

Acknowledgements

Once I was 23 years old, I fell in love with Econometrics and had a dream to become an econometrician/economist. Then I made a decision to start my training at the University of Groningen. Most of the time, the training was enjoyable because I could feel that I was making progress everyday. Nevertheless, it was also a tough journey. I had once doubted my goal I was pursuing and the meaning of my life. Luckily, when I felt low, there were many people cheering me up and encouraging me to move on.

First of all, I want to express my great gratitude to my supervisors Rob Alessie and Viola Angelini. We have had a lot of formal meetings and lunch talks about my curriculum studies, research projects, job seeking, and so on. They are very knowledgeable in both theoretical and empirical economics. I have gained a lot of research experience under their supervision. They always encourage me to think outside the box, ask critical questions to develop my data and economic intuition, and guide me to be a good economist. They are also wonderful people personally. They both care about my work-life balance and help me relax when I experience stressful moments. When I visited the job market, they provided a lot of support and encouragement. Rob inspires me with his insights into many economic and econometric models, and he always encourages me to explore beyond that. One of my favourite things in the past few years was to ask him economic/econometric questions. If he was not busy, he would ask me to sit beside him, take out a notepad under several empty coffee cups, and enthusiastically show me how to play around with the models. This experience has deeply influenced my research attitude. Now, when I have interesting ideas, I can quickly immerse myself into the world of models. Rob is also very humorous and approachable. It often made my day when he greeted me on the seventh floor. When I felt stressed, Rob often made friendly jokes to help me relax. Viola has taught me how to develop research stories convincingly and cohesively. More importantly, she has profoundly improved my way of conducting research. When I got distracted by extra information aside from the main story, she

would directly point it out that I should focus on the main research question. The word 'focus' is now the top word in my mind when I conduct research, and it is the most valuable word I have learned in the past six years. She has also taught me many useful presentation and collaboration skills. Besides research, Viola is also a wonderful life advisor. She always recommends excellent food and places. When I receive good news or make progress, she is always the first person to congratulate me and encourage me to move on. They are my role models, and I feel extremely lucky to have both of them as my supervisors.

I would like to thank my assessment committee: professor Robert Lensink, professor Marike Knoef, and professor Giacomo Pasini. Thank you very much for your valuable comments on my thesis. The thesis has also benefited from many useful comments and suggestions by Arthur van Soest (Tilburg), Hanz Bloemen (UV), Justina klimaviciute (Vilnius), Mauro Mastrogiacomo (UV), Max Groneck (RuG), Sandór sóvagó (RuG), Shiqi Guo (SHJTU), and many others during various seminars and conferences.

My gratitude also goes to many people in the EEF department. Max gave me a lot of helpful comments on my research and job market preparations. My job market visit would not be as smooth as it was without his help. Raun provided me with a lot of insights and suggestions during both seminars and our daily talks. It was also a pleasure to talk with him about our life experience. Jochen shared a lot of nice job information with me, and I have also learned a lot from him via many activities at AJSPH he is directing. Niels and Michaela also helped me with my job market preparations. I have also learned a lot from many talks with Agnieszka, Eduard, Jeroen, Laura, Sarah, and Uli. I want to thank you all. I have many good memories with a lot of departmental members in various social activities. I am very happy to be a PhD student in the EEF department.

I also want to thank my officemates Matthijs, Gianmaria, and Bart. All of them are smart, open-minded, and easygoing. Matthijs, even though I am not a macro person, I really enjoyed chatting with you about our research when we shared the "fish tank" office on the ninth floor. It was also my pleasure when we were drinking beer together. Gianmaria, it has always been my pleasure when we were sharing the 709 office, doing sports (fencing and bouldering), playing video games, and watching horror movies together. Bart, you are such an excellent person to talk with and you are always warm-hearted and humorous. It was really a pity that I have lost quite some time sharing office with you because of COVID-19. I cherish all the

moments I had with all of you.

SOM and Netspar have funded my research and many conference visits. My job placement officer professor Harry Garretson helped arrange the funding for my job market visits to Rotterdam and San Diego. I wish to thank you all for your support. My Chinese supervisor professor Shouyang Wang and professor Xinyu Zhang helped me in many aspects. It was my great honour to have such an opportunity to be your student. I also want to thank Adriaan and Sandór for their encouragement when I was having my interviews in San Diego. I still remember that we had a wonderful dinner at an Indian restaurant. It made me feel supported during that stressful time on the job market. My gratitude extends to Arthur, Ellen, Grietje, Hanneke, Kimberley, Kristian, Richard, and Rina. Thank you so much for your administrative support during my research.

I have many happy moments with my friends in Groningen. Especially I want to thank David, Yimin, Linyang, Yan, Jingjing, Ye, Tobi, Maite, Gianmaria, Wentao, Lina, Zhongbo, Yuwan, Huan, Chengyong, Li, and many more. I really enjoyed our occasional gatherings in Groningen.

Lastly, I want to thank my family, who always stand by me throughout this journey. Especially, I want to thank my girlfriend Ya for her company. It is my greatest luck to be your boyfriend. We had a lot of wonderful memories together in the past several years, and I am looking forward to having more with you in the future. Last but not least, my gratitude goes to my mother, Rongmei Li, who is the best person in my world and the best reason I live in this world. Her altruistic and dedicated investment in my human capital is the major determinant of all the achievements I have made.

Jerry
Groningen
Summer 2020

Contents

List of Figures

List of Tables

Chapter 1

Introduction

As living conditions and healthcare technologies improve, individual life expectancy increases. The pattern is not only evident in developed countries but also in developing countries such as China. According to the *World Population Prospect 2017* report by the United Nations, average life expectancy at birth in China has risen from 69.7 between 1990 and 1995 to 76.5 between 2015 and 2020, and it is projected to rise to 81.1 in the period 2045-2050. Yet, the goals of public health are not only about quantity but also about quality of life. Even though people now live longer, they wish to live in good health and financial well-being.

An expanding body of literature has documented that conditions early in life are important determinants of well-being at older ages. Therefore, if we want to understand how to improve financial, physical and mental health of older individuals, it is crucial to take a life-cycle perspective. In this thesis, I study the early life determinants of well-being at older ages in China.

I mainly focus on two factors that might affect old-age well-being: children's support and early life adversities. Covered by less generous social insurance plans, poor households in China have to seek alternative support to finance their old-age consumption. One natural solution for parents is to find supports from their children. Children's economic support for their parents might be further determined by parental human capital investment in children when they were at schooling ages. The second factor is the role of early life adversities. In China, the current cohort of older people might have experienced the Chinese Great Famine and/or the Cultural Revolution, potentially resulting in income losses during life and long-term negative health effects.

The main data is drawn from the China Health and Retirement Longitudinal Study (CHARLS). CHARLS is a nationally representative survey focusing on individuals aged above 45 years old, including data on intergenerational transfers, individual lifetime work histories, and blood-based biomarker information.

1.1 Why do Parents Underinvest in their Children's Education?

In chapter 2, we examine children's old-age support to their parents. Children's support to their old-age parents is believed to be a good supplement to public social insurance programs, especially in developing countries (see, e.g., Cai et al., 2006; Oliveira, 2016). To what extent children support their old parents depends on the children's income capacity which had been determined by educational human capital invested by the parents when their children were at schooling ages (see, e.g., Alessie et al., 2014; Cai et al., 2006; Raut & Tran, 2005). The relationship between parental human capital investment in children and children's support to their parents can be also affected by certain socio-economic and environmental conditions when parents had to make decisions about investing in their children's human capital. Liquidity constraints and the presence of educational fixed costs are two main conditions (see, e.g., Keane & Wolpin, 2001; Lochner & Monge-Naranjo, 2011). For example, even if the parents are willing to invest, they invest insufficiently as they are financially too weak to borrow from banks or neighbours. Or because there are high entry costs for parents to invest in children's human capital. As a consequence, their children's human capital is underinvested. Incorporating these two conditions into the intergenerational transfer framework and using Chinese data to test the predictions are very important, as we can learn which conditions hinder parents from investing in their children's education. Then the government can implement targeted interventions. Therefore, the first research question is:

To what extent are the parental investment in children's human capital and the old-age support by children affected by the presence of liquidity constraints and educational fixed costs?

To study this research question, this chapter employs a two-sided altruism intergenerational transfer framework proposed by Raut & Tran (2005). We extend the model by incorporating the parental liquidity constraint and educational fixed

costs, and we empirically test the predictions using the 2013 China Health And Retirement Longitudinal Study (CHARLS) survey data.

The results first show that, even with non-binding liquidity constraints and in the absence of fixed costs, children's human capital is underinvested because of the bargaining process between parents and children who are both imperfectly altruistic. Second, the binding liquidity constraint on the parental side reduces the human capital investment even further. Finally, we show that the presence of educational fixed costs increases the children's transfer to their parents during old-age.

1.2 The Impact of the Cultural Revolution on Lifetime Income

In chapter 3, I look at the effect of exposure to adverse political events early in life (aged 5-19) on lifetime income using the Chinese Cultural Revolution (1966-1976) as a natural experiment. The Cultural Revolution hit prefectures differently, hence I can make use of the variation in violence across prefectures to identify the long run effect of the Cultural Revolution. There are also different parties involved in the Cultural Revolution: the general population as victims and the individuals who created violence by joining the so-called Red Guard Army. Therefore, it is also important to investigate whether the Red Guard Army members are affected similarly to those non-Red Guard Army members. In terms of measuring income, while existing studies use cross-sectional income information (see, e.g., Meng & Gregory, 2007), I consider lifetime income which measures human capital over the life cycle. In chapter 3, my research question is:

What is the effect of exposure to the Cultural Revolution during schooling ages (5-19) on individual lifetime income? Are there any different effects between the Red Guard Army members and the non-Red Guard Army members?

To study this research question, I use data from the CHARLS life history survey data set. The CHARLS life history survey collects information on all life events such as work and education episodes from the respondents who have been initially interviewed in 2011 and 2013. The data allow constructing two measures of lifetime income: average lifetime income per worked year and income growth between the first and last jobs in the working life. I then merge the prefecture-level variations in violence during the Cultural Revolution (1966-1976) with the individual level

lifetime income data using the residence information in 1966 when the Cultural Revolution began.

The results show that males who were exposed to a 1% higher violence experience an 8 percentage points decrease in their average lifetime income per worked year, whereas this effect is insignificant for females. In addition, not everyone was affected to the same extent. Males who actively joined the Red Guard Army actually benefited from the Cultural Revolution, resulting in a 16% increase in their average lifetime income per worked year. I also find that individual income growth over the life cycle is not affected by exposure to the Cultural Revolution.

1.3 Exposure *in utero* to Adverse Events and Health Late-in-life

In this chapter, we look at the effects of *in utero* exposure to adverse events on health late in life. We employ two natural experiments in this study: the Chinese Great Famine, which mainly affected rural areas, and the Cultural Revolution, which mainly affected urban areas. To measure health, we use blood-based biomarker information, and we construct risk scores for diabetes and cardiovascular diseases. This is important as it allows us to capture the underlying risks of getting diabetes and cardiovascular diseases, especially when these chronic diseases have not yet manifested. In addition to physical health, we also construct measures of mental health such as depression and cognition. As the biomedical literature suggests that boys develop tissues and organs *in utero* differently than girls (see, e.g., Eriksson et al., 2010), we also conduct analyses separately by gender. Hence, the third research question is:

What are the effects of exposure in utero to the Chinese Great Famine and the Cultural Revolution on physical and mental health at older ages? Do the effects have gender-specific patterns?

To prepare the data set, we measure the severity by the province-level excessive death rates which are constructed by subtracting the average death rates in the non-famine years from the death rates during the famine years, and we use the prefecture-level violence measures which are also employed in chapter 3. Then we merge these data with individual health outcomes constructed from the CHARLS 2011 and 2013 surveys using information on the place and year of birth. In terms of health outcomes,

we construct an 8-year risk score for diabetes using the algorithm in Wilson et al. (2007), a 10-year risk score for cardiovascular diseases using the algorithm in D'agostino et al. (2008), and several measures of cognition and depression.

The results show that female babies who were exposed *in utero* to a 1% higher famine-induced excessive death rate have a 0.06 percentage point higher chance of getting type-2 diabetes. The effects are not significant for male babies. Moreover, male babies who were exposed to a higher level of violence during the Cultural Revolution are shown to have worse cognitive abilities.

The thesis contributes to the literature in several ways. Chapter 2 extends the model in Raut & Tran (2005) and studies two crucial mechanisms: liquidity constraint and fixed costs. The empirical results support both mechanisms, so focusing on these mechanisms and designing proper interventions will be beneficial to the society. Chapter 3 contributes to the literature by considering lifetime income as the measure of economic well-being. This measure captures income over the lifecycle rather than at one point in time. Besides, this chapter also finds that not everyone exposed to the CR was affected in the same way. Red Guard Army members have benefited from this event. Chapter 4 uses biomarkers to compute risk scores as health outcome measures, which are better health proxies when chronic diseases such as diabetes and CVD are not entirely developed. We also find that the Cultural Revolution and the Chinese Great Famine affect different health outcomes.

1.4 Policy Recommendations

To meet the challenges imposed by population ageing and the associated rising healthcare costs, governments need to propose and implement proactive policies. My thesis offers some insights in the design of relevant policies. The main take-away message is simple: interventions at early stages of life generate more long run benefits, especially for people who have experienced adversities early in life.

1.4.1 General Policy Recommendation

In chapter 2 of my thesis, we find that both binding liquidity constraints and the presence of educational fixed costs can explain the underinvestment of parents in their children's education. Therefore relaxing the binding liquidity constraint on the parental side and removing the fixed costs can help improve the educational

investment in the children. In terms of easing the household liquidity constraint, the current poverty alleviation programs (see, e.g., Meng, 2013) implemented nationwide might increase the educational attainment of children from poor families. Governments could, however, do more to reduce educational fixed costs. These costs could be tangible barriers such as transportation costs and tuition fees, which can be mitigated by expanding low-interest rate student loan programs or other policy instruments. Some developed countries have implemented certain policies already. The Netherlands, for example, provides college students with discounted public transportation (*Studentenreisproduct*).[1] The fixed costs can also be intangible such as inequality of opportunity for students from low-income families to receive a high-quality college education. For example, coastal provinces provide more opportunities for local students to enter good universities, while students from western or central provinces have to compete fiercely for limited opportunities (see, e.g., Zhang & Kanbur, 2005). If the human capital investment level increases towards optimum, the results of chapter 2 suggest that old-age transfers to parents will also increase. This is important as in a country with a not well developed pension system such as China, older people strongly rely on the financial support from their children.

From chapter 3 of my thesis, I find that males who were exposed to the Cultural Revolution during their schooling ages have experienced a lifetime income loss. The results imply that, the effect due to the Cultural Revolution on income persists over the life cycle. In other words, once the human capital investment process has been interrupted, it is difficult to "pick up the losses" during work after the Cultural Revolution. Therefore, if children are experiencing interrupted education, either due to domestic conflicts or school shutdowns, it is important to devote effort and resources for children who could benefit the most from on-site teaching. Chapter 3 also finds that individuals who joined the Red Guard Army and who were already more likely to come from an advantaged socio-economic background have much higher lifetime income. This result suggests that events like the Cultural Revolution might increase long run inequality.

In chapter 4, we find females exposed *in utero* to the famine have a higher chance of getting type-2 diabetes, and individuals who were exposed to the Cultural Revolution have lower cognitive abilities later in life. The general conclusion is that adverse conditions experienced around birth can have an important negative effect on health at older ages. This suggests that prevention should start very early in life,

[1] See, for example, https://www.duo.nl/particulier/aanvragen-studentenreisproduct.jsp.

even during pregnancy. In some developing countries which are now experiencing domestic conflicts or famines, it is therefore important to first sustain food supply and maintain safe shelters to pregnant mothers.

1.4.2 Policy Recommendation Regarding COVID-19

The recent global pandemic outbreak of coronavirus disease (COVID-19) has imposed severe health and economic threats to many countries. Most European and Eastern Asian countries have lost many lives due to this pandemic. After implementing lockdowns and social-distancing policies, the situations in these countries seem to be under control.

If we switch our attention to the other part of the globe, the situation is becoming worse. Currently, millions of civilians in African and the Middle Eastern (developing) countries are experiencing not only health pandemic but also severe hunger pandemic, according to a recent report by the World Food Programme.[2] In addition to lockdowns and social-distancing policies, it is extremely important for the governments to implement immediate interventions to sustain the food supply, especially for the pregnant mothers. This is because, according to the results in chapter 4, when the mothers experience severe food shortage, their babies will be more likely to develop chronic diseases such as diabetes once they become old.

The potential hunger pandemic will not only affect newborns but also has the potential to produce lifetime income losses for the current generations of school-aged children. As this World Food Programme report points out, the countries which have higher risks of shaping large-scale famines are mostly conflict-riven. According to the result in chapter 3, individuals who were exposed to higher conflicts during their early lives are more likely to have lower lifetime income. Therefore, when the governments are handling food shortages, they should also provide safe living and studying environments for the young people aged between 5 and 19.

[2] https://www.wfp.org/news/wfp-chief-warns-hunger-pandemic-covid-19-spreads-statement-un-security-council.

Chapter 2

Why do Parents Underinvest in their Children's Education?*

2.1 Introduction

Human capital accumulation is of crucial importance for both individual welfare and economic growth (Schultz, 1961). Besides the governments' efforts in increasing human capital investment, parental investment in children's human capital is also important (see, e.g., Barro, 1974). The traditional theory predicts that the optimal amount of human capital investments is determined by equating its marginal returns to the market return (see, e.g., Becker, 2009). Many papers, however, found that parents often underinvest in the human capital of their children, especially in developing countries such as China (Psacharopoulos, 1985; Heckman, 2005).

Raut & Tran (2005) model the human capital investment behaviour of (extended) families living in developing countries. In such countries with rudimentary pension systems and not well developed capital markets, there exists a link between the parental investment in human capital of children and old-age financial support provided by children. Raut & Tran (2005) take this link explicitly into account by formulating two alternative two-sided altruism models of parental investment in children in which it is assumed that the parents are altruistic towards children and *vice versa*.[1] According to the first model, parents and children have a pure loan

*This chapter is based on Wang et al. (2017).

[1] Alternative explanations have been offered in the literature. Brown et al. (2011), for example, find that the parents will rationally underinvest in their children's human capital strategically because the

contract in which the parental investment in education is paid back by means of old-age support by the child. Children will take out the loan of the parents, as long as they are not made worse-off by the terms of the contract. The first model assumes that the intergenerational contract is enforced by means of social norms. By contrast, the second model is based on reciprocity and the child autonomously decides how much (s)he transfers to parents. In the second model, the level of parental investment in children's education and the amount of transfers from children to old parents are determined simultaneously in a Nash equilibrium. In the empirical part of their paper using data from the Indonesian Family Life Survey, Raut & Tran (2005) find that the two-sided altruistic model fits the data in a better way.

In this chapter, we extend the theoretical model of Raut & Tran (2005) by including parental liquidity constraints and fixed costs of investment as two potential explanations for the underinvestment in education. First, binding liquidity constraints might prevent the parents from investing sufficiently in the education of their children (see Barham et al., 1995). Second, parents might face fixed costs when they perform such investments. One can think of direct costs, such as tuition fees and transportation costs, but also opportunity costs in terms of foregone wages because a full-time student cannot contribute to the income of the household. We construct a two-period theoretical model in which children and parents are both altruistic. In the first period, parents decide how much to invest in their children's education, subject to liquidity constraints and a fixed cost of investment. In the second period, children can provide old-age financial support to their parents. We then test the theoretical predictions of our model on data from the China Health and Retirement Longitudinal Study (CHARLS). The 2013 wave and the retrospective life-history survey provide us with very rich information on 17,311 children and their older parents, aged 45 and over. Our results support the idea that both fixed costs and binding liquidity constraints play an important role in explaining human capital underinvestment.

The rest of the chapter is structured as follows. In section 2.2, the theoretical model and its corresponding predictions will be presented. Section 2.3 provides information on the CHARLS data and on the variables of interest. Section 2.4 summarizes our estimation strategies, and section 2.5 presents the empirical results. Section 2.6 concludes.

parents might not be able to claim returns from strategic-concerned children. Parental risk aversion is also revealed to have a negative impact on the children's human capital investment (see Checchi et al., 2014).

2.2 Model

In this subsection we extend the two-period model of Raut & Tran (2005) by allowing for liquidity constraints and fixed costs. The model distinguishes two types of agents: parents and children who are both altruistic. No distinction is made between father and mother so that they are treated as a single representative parent. Moreover, the parent has n children which are assumed to be identical in preferences, endowment of abilities and in altruism towards parents.

Parents and children both live for three periods but we focus on two overlapping periods between the parent and the child. Figure 2.1 provides a description of the timing structure of the model. In the first period, the parent is in the labour market and the child is at school age. In the second period the parent is retired and the child has joined the labour market. In the first period, the parent earns E_{p1}, consumes c_{p1}, and invests T_1 in each child's schooling. If $T_1 > 0$, the parent faces also fixed costs C of education. In the second period, the parent receives E_{p2} as pension benefits and consumes c_{p2}, and each child is liquidity constrained, i.e. (s)he earns E_{k2}, consumes c_{k2}, and transfers T_2 to support his/her parent and does not save. The parameters

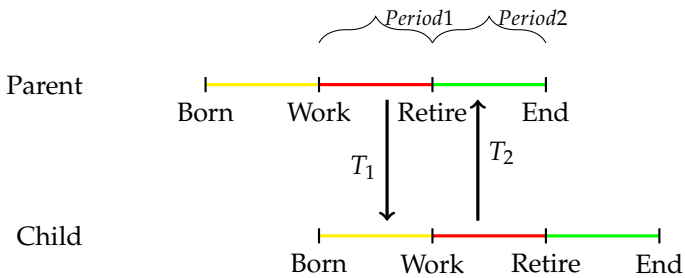


Figure 2.1. Parent-Child Transfer Scheme

γ_p and γ_k measure the degree of the parent's and child's altruism respectively. Due to imperfect altruism, the parameters γ_p and γ_k are strictly positive but smaller than one.

Raut & Tran (2005) construct two models where in the first model the parent is the dominant decision maker: she makes every choice, including parental and child's consumption across two periods, parental human capital investment to children (T_1) and the children's old-age transfer to parents (T_2). The second model allows for some bargaining power for the children in the second period so they can decide on T_2 and c_{k2} in a Nash equilibrium framework. By using a data set from Indonesia,

they find that the second model fits the data better. We therefore only consider the second model in this study. We extend this model by allowing for fixed costs and a liquidity constraint. In other words, we assume that the parent takes the child's decision on T_2 and c_{k2} as given, and that she solves the following optimization problem:

$$\max_{c_{p1}, c_{p2}, T_1} \quad u(c_{p1}) + \beta[u(c_{p2}) + \gamma_p \nu_p(c_{k2})] \tag{2.1a}$$

$$\text{s.t.} \quad c_{p1} = E_{p1} - (s + nT_1 + nC\mathbf{1}(T_1 > 0)) \tag{2.1b}$$

$$c_{p2} = (1 + r)s + E_{p2} + nT_2 \tag{2.1c}$$

$$s \geq 0 \tag{2.1d}$$

$$T_1 \geq 0, \tag{2.1e}$$

where β is the time preference parameter and $\mathbf{1}(T_1 > 0)$ is an indicator function being equal 1 if the parent invests and 0 if she does not, and C is the fixed cost. s denotes net worth at the end of period 1. The parameters γ_p and γ_k measure the degree of the parent's and child's altruism respectively. Due to imperfect altruism, the parameters γ_p and γ_k are strictly positive but smaller than one. $\nu_p(c_{k2})$ represents the parent's perception of child's utility from his/her consumption in period 2. Liquidity constraint (2.1d) says that the parent cannot borrow without collateral. Constraint (2.1e) says that human capital investment cannot be negative.

Taken her parent's decision on T_1 and c_{p1} as given, each child solves the utility maximization problem subject to the budget constraint only in the second period:

$$\max_{c_{k2}, T_2 \geq 0} \quad \nu(c_{k2}) + \gamma_k u_k(c_{p2}) \tag{2.2a}$$

$$\text{s.t.} \quad c_{k2} = E_{k2}(T_1, \tau) - T_2, \tag{2.2b}$$

where $E_{k2}(T_1, \tau)$ denotes a Mincerian earnings function and τ is the talent of the child. $u_k(c_{p2})$ denotes the child's perception of her parent's utility of consumption at old age. The model above represents a sequential game and we solve it by using the backward induction approach, namely the child first optimizes her problem with respect to his/her consumption and the transfer amount to the parent and then the parent solves his/her optimization problem given the child's optimal decisions.

2.2.1 Closed Form Solution

To provide the intuition behind the model, we follow Raut & Tran (2005) in making several additional assumptions. First, we assume that both the children and the parents have the perfect perception on each others' utilities, namely $u_k(c_p) = u(c_p)$ and $v_p(c_k) = v(c_k)$. Second, we impose the restriction that parents and children have the same within-period utility function of the Cobb-Douglas type: $u(c) = v(c) = \alpha \ln c$, with preference parameter $\alpha > 0$.

With this assumption, we restrict the degree of CRRA (Constant Relative Risk Aversion) of all households to be one.[2] Third, we assume a Mincerian earnings function of the children, that is, the log earnings is a function of school-related human capital investment and has a hump-shaped relationship with working experience. Contrary to Raut & Tran (2005), we do not impose the Inada conditions. Consequently, we allow for corner solutions for the human capital investment T_1. Intuitively, those with no human capital investment could also work and earn some money.

If the liquidity constraint (2.1d) is not binding in the first period and no fixed costs appear ($C = 0$) in (2.1b), then we will obtain the benchmark predictions as Raut & Tran (2005) on the optimal human capital investment and the optimal old-age transfers. The optimal investment amount is reflected in the corresponding marginal returns of T_1, which states that

$$\frac{\partial E_{k2}}{\partial T_1} = \frac{1+r}{\gamma_k \gamma_p}, \tag{2.3}$$

where the numerator denotes the market returns. The general marginal returns to human capital investment are higher since both the parent and the child are assumed to be imperfectly altruistic ($0 < \gamma_k \gamma_p < 1$). The marginal returns are negatively associated with the level of human capital investment due to the rule of diminishing rate of returns, see Figure 2.2. The underinvestment is a consequence of the bargaining between the child and the parent. If the parent's bargaining power is large enough to determine both T_1 and T_2, then the marginal returns are equal to $1 + r$ (see appendix 2.A).

Another prediction is that the optimal level of old-age transfer from the child to

[2] We leave the general risk averse cases in future research. Some research such as Checchi et al. (2014) has shown that the parental risk aversion has a negative impact on the children's human capital investment.

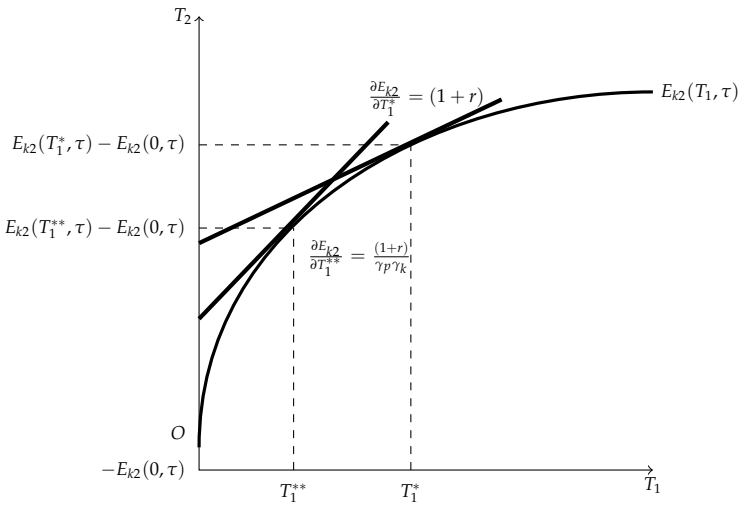


Figure 2.2. Rate of Returns to Human Capital Investment

the parent is equal to:

$$T_2 = \left[\frac{\gamma_k}{\alpha\beta + \gamma_k}\right] E_{k2} + \left[\frac{(1+r)\alpha\beta}{\gamma_k + \alpha\beta}\right]\left[T_1 - \frac{E_{p1} + \frac{E_{p2}}{1+r}}{n}\right]. \tag{2.4}$$

This equation deserves some explanation. First, notice that the child's transfer to the parent does not depend on the degree of parental altruism γ_p. Second, equation (2.4) also says that T_2 is positively related to the income of the child E_{k2}. The impact of E_{k2} is larger if the child is more altruistic. There is also a positive effect of human capital investment T_1 on T_2: parents who invested in their children's human capital are more likely to receive financial support from their children during retirement. This effect is negatively related to γ_k.

Third, the higher the lifetime income of the parents ($E_{p1} + \frac{E_{p2}}{1+r}$), the less will be transferred. Finally, the model predicts zero transfers by the child if $\gamma_k = 0$.[3]

2.2.2 Binding Liquidity Constraint

We now add an extra liquidity constraint (2.1d) to the optimization problem which says that the parents cannot borrow against future income due to imperfect capital

[3] In appendix 2.A we also derive the analogy to (2.4) if the parent decides both T_1 and T_2. The signs of the variables are the same, but the coefficients are functions of parental altruism γ_p instead of γ_k.

market.

The extended model produces the following prediction on human capital investment T_1 (see appendix 2.A for details):

$$\frac{\partial E_{k2}(T_1, \tau)}{\partial T_1} = \frac{1+r}{\gamma_p \gamma_k} + \frac{\mu}{\nu'(c_{p2})\beta\gamma_p\gamma_k}, \tag{2.5}$$

where μ denotes the Kuhn-Tucker multiplier for the restriction (2.1d). The second term on the right hand side is therefore positive if the liquidity constraint is binding. This implies that the optimal human capital investment level is even lower for liquidity constrained households.[4]

As we will argue in section 2.4, we can empirically test for the existence of binding liquidity constraints by following a similar approach as in Zeldes (1989). Equation (2.5) forms the basis of this test.

2.2.3 Fixed Costs

If human capital investment involves fixed costs, i.e. $C > 0$ in equation (2.1b), then the choice problem is not standard due to the fact that the choice set is non-convex. Finding the optimal solution requires several steps (see, e.g., Hausman, 1980 and Cogan, 1981). First, one solves the Nash Bargaining problem presented above under the restriction that the parent does not invest in the human capital of the child ($T_1 = 0$). In that case the parent does not incur any fixed costs and obtains a maximum intertemporal utility level of U_p^0. Then one solves the same problem imposing the condition that the parent invests in the schooling of the child ($T_1 > 0$) and as a consequence faces fixed costs. Let U_p^1 and T_1^* the optimal solution in case of investment. The decision to invest in the education of the children involves a utility comparison: if $U_p^1 > U_p^0$, the parent invests T_1^*, otherwise not.

The child's decision on the old-age transfer T_2 depends the parent's decision on human capital investment in the following way:

$$T_2 = \left[\frac{\gamma_k}{\alpha\beta + \gamma_k}\right] E_{k2} + \left[\frac{(1+r)\alpha\beta}{\gamma_k + \alpha\beta}\right]\left[T_1 - \frac{E_{p1} + \frac{E_{p2}}{1+r}}{n} + C \cdot \mathbf{1}(T_1 > 0)\right], \tag{2.6}$$

[4] Actually the optimal transfer decision can also be derived, but it cannot be tested empirically because the variables entering into the equations are the same as in equation (2.4). The derivations and complete results can be found in appendix 2.A.

and it holds if the parental liquidity constraint (2.1d) is not binding. If the parents face a binding liquidity constraint, T_2 only depends on income of the children and parents in the second period and not on fixed costs C (see equation (2.A.19) in appendix 2.A). Compared with (2.4), equation (2.6) contains one extra term which reflects the impact of fixed costs associated with human capital investment. Notice that T_2 is positively related to those fixed costs provided that the child is altruistic ($\gamma_k > 0$). As before, $T_2 = 0$ if the child is not altruistic. Empirically, we can just test whether the estimated coefficient corresponding to the dummy for positive human capital investment ($I(T_1 > 0)$) is positive and significant. Since transfers from children to parents are often observed with corner solutions and children also might face fixed costs if they transfer money to the parents, it is worthwhile to consider both the intensive and extensive margins. In section 2.4 we will discuss the empirical strategy in more detail.

2.3 Data

We draw data from the China Health and Retirement Longitudinal Study (CHARLS), which collects information on a nationally representative sample of Chinese residents aged 45 and older. The baseline wave of CHARLS was fielded in 2011 and includes about 10,000 households and 17,500 individuals in 150 counties/districts and 450 villages/resident communities.[5] The individuals are then followed every two years. Currently, three regular waves of CHARLS are available, in addition to a life history survey which was administered to the households interviewed in the first two waves. In this chapter, we use data from the second wave[6] and the life history survey of CHARLS, which were collected in 2013 and 2014 respectively. From the second wave we obtain information on parental investment in children's human capital, cash transfers that parents receive from their children, and socio-demographic characteristics, while we use the life history of CHARLS to construct a measure of permanent income and gather data on the financial situation of the household when the children were at college ages.

We first construct a child-level data set from CHARLS 2013 where each observation is a child and then we merge it with data from the life history survey. After dropping observations with missing values on some key variables, and keeping only

[5] Detailed information on the data and sampling procedure can be found in Zhao et al. (2012).

[6] We also conduct a sensitivity analysis using data from the first wave (results available upon request).

children with completed education, we obtain a sample of 17,311 children.[7]

The main variables of interest for our analysis are parental investment in children's human capital (T_1), financial transfers from children to their ageing parents (T_2), children's current income, and parental lifetime income. The human capital investment variable is obtained by asking the main respondent[8] how much money (s)he and his(her) spouse have invested on each of their children's college education. As an alternative, we also include the child's number of years of college education as a proxy for the human capital investment. Old-age transfers are measured in net values and are obtained by asking the parents how much they received in total from each non-coresiding child in the last year. Another proxy for the transfer term is a dummy ("Depend on This Child") which is equal to 1 if the respondent plans to rely on the child as a source of financial support during retirement. The income of the child is obtained by directly asking the parents how much each of their children earned last year. The permanent income of the parents is more difficult to measure.

We consider four measures to proxy for it: net financial wealth, net durable wealth, net income, and lifetime income. Net financial wealth is the sum of cash, deposits, stocks, bonds and other financial holdings subtracted by financial debts. Net durable wealth includes housing assets minus debts, gold, treasures and antiques. We only include those housing assets which can be sold immediately on the market,[9] and we do not include other durable goods such as vehicles and televisions since the depreciation rates of these assets are extremely high, hard to evaluate, and are determined by the complexity of the local second-hand markets. Yearly net income consists of the wage income and the pension income for employed and retired workers respectively, and the net income of the farmers and self-employers. Lifetime income is obtained from the life history of CHARLS and are calculated by summing up the inflation-adjusted wages during the working years (see, e.g., Alessie et al., 2013).[10] The correlations between different measures of lifetime income are displayed in Table 2.1.

The core explanatory variables are the liquidity constraint dummies.[11] Liquidity

[7] The detailed sample selection strategies can be found in appendix 2.B.

[8] The main respondent is randomly selected among all age eligible respondents in the household.

[9] In the last century, many housings are allocated by the employers in urban areas such as governments and state-owned enterprises. Some of them can be traded while some cannot.

[10] We use urban/rural-specific CPI data from 1950-2015 across 31 provinces from Chinese National Statistics Bureau as the earnings deflater. Some data points in early years are missing and we use local general CPI or nationwide rural/urban CPI in those cases.

[11] When constructing liquidity constraint indicators, we employ the basic idea from Zeldes (1989) who proposes splitting the whole sample into liquidity constrained sub-sample and the unconstrained sub-

Table 2.1. Cross-correlation between Wealth Variables

Variables	Net Durable Wealth	Net Total Wealth	Net Income	Lifetime Income	Net Financial Wealth
Net Durable Wealth	1.000				
Net Total Wealth	1.000	1.000			
Net Income	0.025	0.035	1.000		
Lifetime Income	0.003	0.006	0.241	1.000	
Net Financial Wealth	-0.078	-0.061	0.082	0.014	1.000

Note: In the estimation process we use mainly financial wealth and total wealth instead of durable wealth, since as the table shows, the correlation between total wealth and real wealth is almost perfect. This also implies that the durable assets are the majority of the total assets in elderly Chinese households. The net financial assets are more fluctuated and therefore negatively correlated with the total wealth.

constraints are measured by six indicators, namely (1) whether the family lived in a shack when the child was at college ages; (2) whether the income of the respondent was in the lowest 25% group of the sample when the child was at college ages; (3) whether the family had a shared toilet or a private one; (4) whether the family had a water closet or not; (5) whether the family was using clean water or not; (6) whether the family used electricity or not. The correlations among the liquidity dummies are shown in Table 2.2.

Table 2.2. Cross-correlation between Liquidity Variables

Variables	Income in Bottom Quartile	Living in a Shack	Private Toilet	Water Closet	Clean Water	Electricity
Income in Bottom Quartile	1.000					
Living in a Shack	0.088	1.000				
Private Toilet	-0.058	-0.139	1.000			
Water Closet	-0.118	-0.264	0.398	1.000		
Clean Water	-0.136	-0.271	0.178	0.395	1.000	
Electricity	-0.058	-0.184	0.027	0.058	0.185	1.000

Note: Living in a shack or earning the income in bottom quartile show the cases where liquidity constraints are binding while having private toilet, water closet, using clean water and electricity shows the cases where liquidity constraints are non-binding.

In addition, we also construct several socio-demographic controls, such as the number of children of both the main respondents and his or her children, their genders, the household registration types (*Hukou*), marriage status of both the main respondents and their children, their ages, and cohort dummies. The descriptive statistics can be found in Table 2.3.

sample using whether households with low savings or financial wealth as the classification criterion. Jappelli et al. (1998) criticize this approach and introduce whether the households were refused loans or discouraged from borrowing as the proxy for liquidity constraints. Guiso et al. (1996) provide a similar proxy, namely the credit status of the individuals, to test the impact of liquidity constraints on income uncertainty.

Table 2.3. Descriptive Statistics of Child-Level Data

	VARIABLES	N	Mean	S.D.	Min	Max	Quartile 1	Median	Quartile 3
Parent (Every Respondent Answer)	Female	17,311	0.511	0.500	0	1	0	1	1
	Age	17,311	64.53	10.29	36	95	57	64	72
	Agric. Hukou	17,311	0.801	0.399	0	1	1	1	1
	Married or Cohabited	17,311	0.785	0.411	0	1	1	1	1
	Num. of Children	17,311	3.384	1.616	0	16	2	3	4
	Years of Schooling[1]	17,311	5.758	4.303	0	19	2	6	9
	HH Income[2]	17,311	1.166	4.578	-299.7	92	0	0.192	1.595
	Net Financial Wealth[3]	17,311	-0.0215	8.854	-289.9	210.5	0	0.0500	0.500
	Net Durable Wealth	17,311	73.13	539.4	-2,999	28,000	0	0.01000	0.150
	Net Total Wealth	17,311	74.27	539.1	-3,035	27,992	0.0902	0.635	3.530
	HH Income Children at College Age	17,311	1.175	4.585	-299.7	92	0	0.190	1.600
	HH Lifetime Income (Life History CHARLS)	17,311	31.93	91.53	0	4,932	4.481	12.22	30.32
Child (Family Respondent Answer)	Female	17,311	0.461	0.498	0	1	0	0	1
	Age	17,311	35.75	9.682	16	77	28	35	42
	No Work	17,311	0.161	0.363	0	1	0	0	0
	Urban Jobs	17,311	0.148	0.355	0	1	0	0	0
	Public Servants	17,311	0.243	0.429	0	1	0	0	0
	Rural Jobs	17,311	0.445	0.497	0	1	0	0	1
	Num. of Children	17,311	1.280	0.915	0	6	1	1	2
	Depend on This Child	17,311	0.496	0.500	0	1	0	0	1
	Coresiding with Family[4]	17,311	0.266	0.442	0	1	0	0	1
	Years of Schooling	17,311	8.544	4.096	0	23	6	9	12
	Agric. Hukou	17,311	0.759	0.428	0	1	1	1	1
	Married or Cohabited	17,311	0.810	0.392	0	1	1	1	1
	Biological Child	17,311	0.991	0.0933	0	1	1	1	1
	Work Experience[5]	17,311	21.20	11.41	0	67	12	20	29
	Have Received Transfer	17,311	0.500	0.500	0	1	0	0	1
	Positive Net Transfer	8,654	0.213	0.613	1.00e-04	30	0.0350	0.0900	0.200
	Net Transfer from Child	17,311	0.0131	3.152	-394.6	30	0	0	0.0900
	Have Received Investment	17,311	0.102	0.303	0	1	0	0	0
	Hum. Cap. Cons.[6]	1,770	0.0689	0.0754	1.23e-06	1.514	0.0362	0.0582	0.0849
	Years of College Schooling	17,311	0.263	0.860	0	10	0	0	0
	Income	17,311	3.040	3.449	0	30	0.750	1.500	3.500
Marginal Returns of Human Capital(Predicted)	Schooling Based	17,311	0.542	0.615	0	5.350	0.134	0.267	0.624
	Investment Based	17,311	0.0475	0.0566	0	0.616	0.0116	0.0308	0.0719
Liquidity Constraints (Life history CHARLS)	No Access to Electricity	17,311	0.743	0.437	0	1	0	1	1
	Private Toilet	17,311	0.140	0.347	0	1	0	0	0
	Water Closet	17,311	0.118	0.322	0	1	0	0	0
	Clean Water	17,311	0.279	0.449	0	1	0	0	1
	Income in Bottom Quartile	17,311	0.211	0.408	0	1	0	0	0
	Living in a Shack[7]	17,311	0.425	0.494	0	1	0	0	1

Note: 1. Years of schooling is calculated based on the legislated years at each education level. 2. All the monetary amounts are at 2010 constant price of ten thousand China yuan which is approximately 1296.546 euros. 3. All wealth are in net terms and at the household level. 4. Coresiding means that the child is living with the parents, regardless whether the child is economic independent or not. 5. Work experience is calculated based on Mincer (1974), namely age minus years of schooling and 6. Human capital investment at constant price of ten thousand 2010 China yuan. 7. This dummy equals 0 if the family lived in houses with concrete structure, built with bricks and wood, while equals 1 if lived in adobe, thatched house, cave, Mongolian yurt or boat house.

2.4 Empirical Strategy

To test against the prediction of the presence of binding liquidity constraint, we first need to predict the marginal returns to human capital based on the estimation of children's earnings function (see equations (2.3) and (2.5)). For the earnings function, we estimate the following model by gender separately:

$$ln(ChildInc_i) = \alpha_0 + \alpha_1 HCinvest_i + \alpha_2 Ability_i + \alpha_3 Exp._i + \alpha_4 Exp.Sq._i + \mathbf{x}_i'\delta + \varepsilon_i, \tag{2.7}$$

where $ChildInc_i$ is the current income of a child and $HCinvest_i$ is the human capital investment during school for which we use two measures: the logarithm of one plus the actual money invested in college education and the number of years of college schooling. To mitigate the selection bias, we control for the innate ability $Ability_i$ of the child, which is proxied by the parental years of schooling. In addition, we employ Mincer (1974)'s idea to incorporate the working experience $Exp._i$ which is equal to the current age minus years of schooling and 6, and $Exp.Sq._i$ is the squared term of experience. Other demographic controls are contained in $\mathbf{x}_i$ such as job categories and *hukou* status. Notably, the children's earnings function is for approximation and prediction instead of causal inference purpose. The estimated parameter $\hat{\alpha}_1$ is employed to predict the marginal returns to human capital investment. It can be shown that the marginal return in specification (2.7) is $MR_i = \frac{\hat{\alpha}_1 \times ChildInc_i}{1+HCinvest_i}$ when we use the logarithm of one plus the actual amount invested as a measure of human capital. When we use the number of years of schooling instead, the marginal return is $MR_i = \hat{\alpha}_1 \times ChildInc_i$. The summary statistics on the predicted marginal returns for each individual can be found in Table 2.3. Then we first test whether the binding liquidity constraint induces underinvestment of human capital. According to equation (2.5), the liquidity constrained child has a positive excess return. We therefore estimate the following model:

$$\begin{aligned} MR_i = \beta_0 + \beta_1 HHInc_i + \beta_2 Num._i + \beta_3 Num.Sq._i + \mathbf{z}_i'\theta + \\ \beta_4 * Cons_i + Cons_i \times (\beta_5 HHInc_i + \beta_6 Num._i + \beta_7 Num.Sq._i + \mathbf{z}_i'\delta) + u_i, \end{aligned} \tag{2.8}$$

where MR_i represents the marginal returns to college human capital investment, $HHInc_i$ is the household income when the child was at college age, $Num._i$ is the number of children in the household and $Num.Sq._i$ is the squared term which

proxies for the altruistic parameters $\gamma_p \gamma_k$, and the vector $\mathbf{z}_i$ contains demographic and social-economic control variables such as age, gender, *hukou* and marriage status of both parent and child, and parental education. The liquidity constraint dummy is represented by $Cons_i$ which is measured by the six indicators in the previous section which equals to one if the individual is liquidity constrained and zero otherwise.

The final step is to estimate the transfer equation with fixed costs namely equation (2.6). The old-age transfer term allows for both positive and negative values, but in this section we care more about the children's decisions of positive transfers to their parents, hence we transform the transfer variable into a corner solution response which is censored at zero.[12] Theoretically, the fixed costs will prevent the parents from making investing decisions, while whether the fixed costs matter for both the intensive and extensive margins is unclear. Therefore, empirically, it is important to investigate which decision is affected by the fixed costs. We will use a double hurdle (two-part) model as our main model for the transfer equation, demonstrating the marginal effects at both the extensive margins (children deciding on whether to transfer to parents) and the intensive margins (if the children decide to transfer, then how much to transfer). The participation equation is a Probit specification, and the second part is a log-linear specification since the positive transfer is highly skewed.[13]

According to equation (2.6), the estimated coefficients for the child's income and the human capital investment are expected to be positive, and the one for household lifetime income is expected to be negative. The coefficient on the fixed costs term, both at the extensive and intensive margins, is expected to be positive if fixed costs are relevant in explaining the underinvestment in human capital.

[12] Raut & Tran (2005) employ a Tobit specification considering the sample selection where $T_1 > 0$, since in their data, only positive transfers are observed. We censored the data below zero but no sample selection problem appears.

[13] Even though we take the log transformation of the positive transfer, we do not assume the log-terms to be normally distributed. Therefore, in estimating the marginal effects, we use the smearing estimator of Duan (1983) to overcome the potential bias problem. When estimating standard errors of the smearing estimator, we use the bootstrapping method with 200 replications.

2.5 Results

2.5.1 Binding Liquidity Constraint

We first estimate equation (2.7) by gender and use $\hat{\alpha}_1$, which is the estimated coefficient on human capital investment, to predict the marginal returns MR_i. We use two measures of human capital investment, namely the number of years of college schooling and the logarithm of one plus the money invested by the parents in the college education of their children. By using the six liquidity indicators selected in the Life history of CHARLS, we further test the role of liquidity constraints in explaining the underinvestment of human capital. The six liquidity constraint indicators are employed to split the whole sample into the liquidity constrained and unconstrained subsamples as Zeldes (1989) suggests.

We estimate model (2.8) by assuming that the selection of liquidity constrained behaviour is uncorrelated with the error term. We test for the presence of liquidity constraints by testing the difference between the estimates of the constrained and unconstrained sample. If the estimates are significantly different, then binding liquidity constraints matter. Therefore, Chow tests are constructed to test whether the slope coefficients of the two subsamples are the same. The Chow statistics and corresponding p-values can be found in the last two rows of Table 2.4 and Table 2.5. The results in Table 2.4 are obtained by using the marginal returns which are predicted based on the human capital investment amount while the results in Table 2.5 employ the years of college schooling. When marginal returns are predicted on the basis of the investment measure, liquidity constraints play an important role when using the indicators "Income in Bottom Quartile", "Water Closet" and "Clean Water". When the marginal returns are predicted through the schooling measure, liquidity constraints matters for the indicators "Income in Bottom Quartile", "Living in a Shack" and "Private Toilet". The evidence on the importance of liquidity constraints is therefore somewhat mixed.

2.5.2 Fixed Costs

In this section, we discuss whether fixed costs affect the human capital investment decisions of parents on their children. In theory, fixed costs affect human capital investment through the old-age transfer equation. The expected sign of the coefficient on the fixed costs is positive. As usual, we show the results for two different

measures of the human capital investment: the actual years of college schooling and the money that parents invested in their children's college education. Table 2.6 shows the estimation results using investment amount as the measure of human capital, while Table 2.7 provides the results using years of college schooling.

The first columns in Tables 2.6 and 2.7 show the marginal effects of the Probit estimation using the "Depend on This Child" indicator. Column (2) reports the coefficient estimates from the Tobit model in which the dependent variable is the value of the monetary transfers received by parents from their children. Columns (3-4) exhibit the marginal effects of the participation equation and the amount equation, and the unconditional marginal effects in the two-part models. Notably again, the transfers are in log forms.

Table 2.4. Testing for Binding Liquidity Constraint, Investment Measure for T_1

	VARIABLES	(1) Income in Bottom Quartile	(2) Living in a Shack	(3) Private Toilet	(4) Water Closet	(5) Clean Water	(6) Electricity
	Poor	-5.898	1.791	4.513	2.288	-6.362	-9.050
		(8.108)	(8.361)	(11.91)	(14.19)	(9.711)	(8.520)
	Age	0.637***	0.781***	0.715***	0.960***	0.683***	0.644***
		(0.0966)	(0.108)	(0.201)	(0.255)	(0.163)	(0.102)
	Age Squared	-0.00759***	-0.00903***	-0.00917***	-0.0122***	-0.00832***	-0.00734***
		(0.00115)	(0.00135)	(0.00268)	(0.00366)	(0.00226)	(0.00133)
	Female	-0.298	-0.374	-0.904*	0.0596	-0.508	-0.426**
		(0.214)	(0.256)	(0.519)	(0.666)	(0.373)	(0.215)
	Agric. Hukou	-1.257***	-0.970**	-1.585	0.0195	-0.253	-1.050**
		(0.444)	(0.486)	(0.965)	(1.063)	(0.712)	(0.412)
	Married or Cohabited	5.145***	4.977***	4.994***	5.721***	5.230***	4.952***
		(0.237)	(0.266)	(0.474)	(0.549)	(0.380)	(0.237)
Child	Coresiding with Family	-2.244***	-2.214***	-2.371***	-2.862***	-1.996***	-2.088***
		(0.224)	(0.255)	(0.503)	(0.563)	(0.341)	(0.223)
	Biological Child	1.680*	0.706	2.878**	3.181	0.379	0.910
		(0.875)	(1.034)	(1.374)	(1.989)	(1.270)	(0.856)
	Age × Poor	0.0233	-0.207	-0.0340	-0.300	0.0469	-0.0281
		(0.166)	(0.161)	(0.220)	(0.269)	(0.189)	(0.178)
	Age Squared × Poor	-4.65e-05	0.00222	0.00116	0.00451	-0.000109	0.000119
		(0.00184)	(0.00187)	(0.00287)	(0.00380)	(0.00250)	(0.00204)
	Female × Poor	-0.270	-0.0326	0.631	-0.461	0.185	0.213
		(0.400)	(0.365)	(0.554)	(0.693)	(0.432)	(0.413)
	Agri-Hukou × Poor	-0.982	-1.009	0.110	-1.524	-1.578*	-1.536
		(0.906)	(0.800)	(1.057)	(1.149)	(0.843)	(1.039)
	Married × Poor	-1.382***	-0.430	-0.143	-1.055*	-0.484	-0.282
		(0.417)	(0.401)	(0.524)	(0.588)	(0.439)	(0.445)
	Num. Children Squared × Poor	-0.0537	-0.213	0.120	0.701	-0.394	-0.709*
		(0.420)	(0.370)	(0.543)	(0.599)	(0.409)	(0.416)
	Biological × Poor	-0.635	1.792	-1.734	-1.824	1.602	2.539*
		(1.433)	(1.326)	(1.615)	(2.142)	(1.564)	(1.508)

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Chow tests (F statistics) and corresponding p-values are given in the last two rows. 2. The dependent variable is the marginal returns to human capital investment and it is predicted by using the actual investment amount, which is obtained by directly asking the parents, as the human capital investment T_1. 3. *Poor* indicates whether the household faced binding liquidity constrained when the children were at college age (age of 18). 4. Six liquidity dummies are employed to proxy for the liquidity situations, namely whether the household income was in the bottom quartile of the sample (Column (1)), whether the family lived in a shack (Column (2)), whether the family had a private toilet (Column (3)), whether the family had a water closet (Column (4)), whether the family had access to clean water (Column (5)), and whether the family had electricity (Column (6)).

Table 2.4 Continued

	VARIABLES	(1) Income in Bottom Quartile	(2) Living in a Shack	(3) Private Toilet	(4) Water Closet	(5) Clean Water	(6) Electricity
	Age	-0.0791	-0.0879	-0.0448	-0.302	-0.321	-0.231
		(0.159)	(0.185)	(0.393)	(0.497)	(0.308)	(0.171)
	Age Squared	0.000633	0.000658	0.000529	0.00297	0.00281	0.00188
		(0.00127)	(0.00151)	(0.00329)	(0.00433)	(0.00263)	(0.00138)
	Female	-0.964***	-1.052***	0.171	0.810	-0.245	-0.660**
		(0.283)	(0.315)	(0.591)	(0.713)	(0.413)	(0.274)
	Agric. Hukou	0.410	0.168	0.915	1.171	0.542	0.0973
		(0.501)	(0.551)	(1.055)	(1.079)	(0.766)	(0.506)
	Married or Cohabited	0.395	0.548	1.351*	1.669	0.524	0.434
		(0.406)	(0.502)	(0.778)	(1.221)	(0.674)	(0.416)
	HH Income	0.184**	0.170*	0.342***	0.315***	0.414***	0.174*
		(0.0897)	(0.0888)	(0.0817)	(0.0810)	(0.136)	(0.0994)
	Years of Schooling	0.130***	0.0873**	0.0424	0.0368	0.119*	0.108***
		(0.0381)	(0.0438)	(0.0865)	(0.109)	(0.0631)	(0.0371)
	Num. of Children	-0.546*	-0.591*	-0.931	-0.192	-0.215	-0.382
		(0.316)	(0.352)	(0.882)	(1.255)	(0.378)	(0.280)
Parent	Num. of Children Squared	0.0578	0.0657	0.134	0.0599	0.0502	0.0265
		(0.0412)	(0.0486)	(0.144)	(0.221)	(0.0515)	(0.0368)
	Age × Poor	0.127	-0.0161	-0.0676	0.236	0.247	0.282
		(0.264)	(0.279)	(0.419)	(0.516)	(0.346)	(0.291)
	Age Squared × Poor	-0.000844	0.000144	0.000323	-0.00254	-0.00238	-0.00235
		(0.00202)	(0.00219)	(0.00348)	(0.00446)	(0.00290)	(0.00223)
	Female × Poor	0.652	0.387	-1.292**	-1.922**	-0.948*	-0.917
		(0.491)	(0.477)	(0.645)	(0.755)	(0.501)	(0.558)
	Agri-Hukou × Poor	1.424	1.321	-0.726	-0.666	-0.0999	1.420
		(1.155)	(0.930)	(1.168)	(1.185)	(0.939)	(1.030)
	Married × Poor	0.209	-0.393	-1.042	-1.308	-0.135	-0.170
		(0.653)	(0.645)	(0.859)	(1.264)	(0.768)	(0.688)
	HH Income Child College Age × Poor	1.024***	0.408**	-0.176	-0.156	-0.252	0.175
		(0.247)	(0.196)	(0.130)	(0.128)	(0.164)	(0.126)
	Schooling × Poor	-0.0942	0.0457	0.0940	0.0753	-0.0171	0.0146
		(0.0693)	(0.0637)	(0.0927)	(0.113)	(0.0726)	(0.0782)
	Num. Children × Poor	-0.417	-0.126	0.213	-0.608	-0.623	-0.696
		(0.428)	(0.494)	(0.922)	(1.280)	(0.474)	(0.456)
	Num. Children Squared × Poor	0.0212	0.00338	-0.0644	0.0214	0.0308	0.0942*
		(0.0483)	(0.0634)	(0.148)	(0.223)	(0.0623)	(0.0539)
	Admission Rate	-1.394	-0.802	-2.942	-6.059	-1.695	-0.496
		(1.476)	(1.748)	(3.668)	(4.514)	(2.609)	(1.460)
	Admission Rate × Poor	2.984	-0.0424	2.652	5.785	1.487	-0.573
		(2.782)	(2.523)	(3.898)	(4.698)	(3.012)	(2.877)
	Constant	-3.118	-4.379	-6.508	-5.417	2.077	1.365
		(4.737)	(5.393)	(11.10)	(13.58)	(8.394)	(5.151)
	Observations	17,311	17,311	17,311	17,311	17,311	17,311
	R-squared	0.088	0.088	0.084	0.088	0.086	0.086
	lnL	-66425	-66431	-66466	-66423	-66445	-66444
	Chow Test	2.868	1.308	0.933	3.056	1.499	1.299
	Prob ≥ F	6.84e-05	0.176	0.534	2.19e-05	0.0848	0.182

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Chow tests (F statistics) and corresponding p-values are given in the last two rows. 2. The dependent variable is the marginal returns to human capital investment and it is predicted by using the actual investment amount, which is obtained by directly asking the parents, as the human capital investment T_1. 3. *Poor* indicates whether the household faced binding liquidity constrained when the children were at college age (age of 18). 4. Six liquidity dummies are employed to proxy for the liquidity situations, namely whether the household income was in the bottom quartile of the sample (Column (1)), whether the family lived in a shack (Column (2)), whether the family had a private toilet (Column (3)), whether the family had a water closet (Column (4)), whether the family had access to clean water (Column (5)), and whether the family had electricity (Column (6)).

Table 2.5. Testing for Binding Liquidity Constraint, Years of College Schooling Measure for T_1

	VARIABLES	(1) Income in Bottom Quartile	(2) Living in a Shack	(3) Private Toilet	(4) Water Closet	(5) Clean Water	(6) Electricity
	Poor	0.108	0.322	0.678	0.938	0.450	-0.211
		(0.417)	(0.421)	(0.600)	(0.727)	(0.492)	(0.437)
	Age	0.0415***	0.0547***	0.0504***	0.0809***	0.0565***	0.0468***
		(0.00485)	(0.00555)	(0.0106)	(0.0139)	(0.00854)	(0.00505)
	Age Squared	-0.000478***	-0.000624***	-0.000574***	-0.000928***	-0.000662***	-0.000531***
		(5.62e-05)	(6.71e-05)	(0.000131)	(0.000187)	(0.000113)	(6.39e-05)
	Female	-0.0364***	-0.0428***	-0.0425	-0.0163	-0.0279	-0.0423***
		(0.0109)	(0.0133)	(0.0261)	(0.0360)	(0.0193)	(0.0110)
	Agric. Hukou	-0.286***	-0.302***	-0.317***	-0.226***	-0.260***	-0.294***
		(0.0249)	(0.0272)	(0.0569)	(0.0611)	(0.0399)	(0.0238)
	Married or Cohabited	0.224***	0.212***	0.250***	0.255***	0.221***	0.211***
		(0.0130)	(0.0148)	(0.0293)	(0.0343)	(0.0210)	(0.0129)
Child	Coresiding with Family	-0.169***	-0.174***	-0.182***	-0.252***	-0.172***	-0.155***
		(0.0122)	(0.0139)	(0.0282)	(0.0331)	(0.0192)	(0.0119)
	Biological Child	0.0861*	0.0409	0.106	0.0624	0.0357	0.0490
		(0.0441)	(0.0529)	(0.0704)	(0.0954)	(0.0755)	(0.0434)
	Age × Poor	-0.00423	-0.0260***	-0.0106	-0.0440***	-0.0178*	-0.0174*
		(0.00824)	(0.00810)	(0.0115)	(0.0146)	(0.00978)	(0.00894)
	Age Squared × Poor	5.53e-05	0.000288***	0.000116	0.000509***	0.000221*	0.000175*
		(9.08e-05)	(9.12e-05)	(0.000140)	(0.000193)	(0.000124)	(9.94e-05)
	Female × Poor	-0.0144	0.00437	0.00330	-0.0248	-0.0174	0.00951
		(0.0199)	(0.0184)	(0.0279)	(0.0372)	(0.0221)	(0.0206)
	Agri-Hukou × Poor	-0.0135	0.0456	0.0309	-0.0611	-0.0369	0.0287
		(0.0510)	(0.0449)	(0.0618)	(0.0657)	(0.0476)	(0.0518)
	Married × Poor	-0.0482**	-4.76e-05	-0.0407	-0.0468	-0.0100	0.00822
		(0.0210)	(0.0213)	(0.0316)	(0.0361)	(0.0242)	(0.0244)
	Num. Children Squared × Poor	0.0349*	0.0272	0.0229	0.103***	0.0126	-0.0222
		(0.0211)	(0.0195)	(0.0302)	(0.0347)	(0.0225)	(0.0222)
	Biological × Poor	-0.0337	0.0837	-0.0341	0.0215	0.0695	0.135*
		(0.0712)	(0.0688)	(0.0826)	(0.104)	(0.0864)	(0.0754)

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Chow tests (F statistics) and corresponding p-values are given in the last two rows. 2. The dependent variable is the marginal returns to human capital investment and it is predicted by using the years of college schooling as the human capital investment T_1. 3. *Poor* indicates whether the household faced binding liquidity constrained when the children were at college age (age of 18). 4. Six liquidity dummies are employed to proxy for the liquidity situations, namely whether the household income was in the bottom quartile of the sample (Column (1)), whether the family lived in a shack (Column (2)), whether the family had a private toilet (Column (3)), whether the family had a water closet (Column (4)), whether the family had access to clean water (Column (5)), and whether the family had electricity (Column (6)).

Table 2.5 Continued

	VARIABLES	(1) Income in Bottom Quartile	(2) Living in a Shack	(3) Private Toilet	(4) Water Closet	(5) Clean Water	(6) Electricity
	Age	-0.00210	-0.00656	0.00691	0.000452	-0.0103	-0.0125
		(0.00793)	(0.00965)	(0.0193)	(0.0264)	(0.0160)	(0.00873)
	Age Squared	1.96e-05	5.25e-05	-5.50e-05	-1.60e-05	9.33e-05	9.84e-05
		(6.29e-05)	(7.90e-05)	(0.000158)	(0.000227)	(0.000137)	(7.06e-05)
	Female	-0.0559***	-0.0645***	-0.00676	0.0332	0.00146	-0.0478***
		(0.0147)	(0.0168)	(0.0322)	(0.0391)	(0.0217)	(0.0144)
	Agric. Hukou	0.117***	0.128***	0.141**	0.153**	0.141***	0.112***
		(0.0283)	(0.0313)	(0.0620)	(0.0614)	(0.0416)	(0.0294)
	Married or Cohabited	0.0125	0.0272	0.0774*	0.0833	0.0303	0.0224
		(0.0198)	(0.0250)	(0.0406)	(0.0629)	(0.0343)	(0.0203)
	HH Income	0.0108*	0.00984*	0.0176***	0.0181***	0.0277***	0.0104
		(0.00575)	(0.00570)	(0.00395)	(0.00467)	(0.00858)	(0.00668)
	Years of Schooling	0.0123***	0.0111***	0.0107**	0.0153***	0.0139***	0.0112***
		(0.00198)	(0.00231)	(0.00449)	(0.00544)	(0.00315)	(0.00193)
	Num. of Children	-0.0634***	-0.0676***	-0.134***	-0.118**	-0.0524***	-0.0508***
		(0.0143)	(0.0156)	(0.0462)	(0.0593)	(0.0175)	(0.0144)
Parent	Num. of Children Squared	0.00586***	0.00628***	0.0154**	0.0158*	0.00545***	0.00405**
		(0.00177)	(0.00202)	(0.00708)	(0.00935)	(0.00197)	(0.00182)
	Age × Poor	-0.00487	0.00128	-0.0140	-0.00354	0.00433	0.0168
		(0.0140)	(0.0140)	(0.0206)	(0.0273)	(0.0178)	(0.0147)
	Age Squared × Poor	3.81e-05	-1.29e-05	0.000111	3.90e-05	-5.42e-05	-0.000126
		(0.000109)	(0.000110)	(0.000168)	(0.000233)	(0.000149)	(0.000112)
	Female Parent × Poor	0.0355	0.0259	-0.0584*	-0.100**	-0.0793***	-0.0226
		(0.0254)	(0.0244)	(0.0348)	(0.0411)	(0.0261)	(0.0277)
	Agri-Hukou × Poor	0.109*	0.0202	-0.0272	-0.0208	-0.0265	-0.00263
		(0.0603)	(0.0506)	(0.0681)	(0.0673)	(0.0513)	(0.0530)
	Married × Poor	0.0222	-0.0255	-0.0685	-0.0700	-0.0176	-0.0240
		(0.0324)	(0.0316)	(0.0442)	(0.0648)	(0.0387)	(0.0333)
	HH Income × Poor	0.0490***	0.0216*	-0.00742	-0.00900	-0.0192*	0.00861
		(0.0120)	(0.0111)	(0.00809)	(0.00793)	(0.0101)	(0.00773)
	Schooling × Poor	-0.00732**	-0.00134	0.000960	-0.00548	-0.00466	0.000311
		(0.00342)	(0.00326)	(0.00481)	(0.00564)	(0.00366)	(0.00391)
	Num. Children × Poor	0.00935	0.0221	0.0778	0.0617	-0.00730	-0.0315
		(0.0196)	(0.0229)	(0.0478)	(0.0605)	(0.0226)	(0.0226)
	Num. Children Squared × Poor	-0.00159	-0.00210	-0.0104	-0.0106	-1.28e-05	0.00398
		(0.00213)	(0.00281)	(0.00723)	(0.00945)	(0.00261)	(0.00256)
	Admission Rate	0.218***	0.288***	0.280	0.235	0.361**	0.260***
		(0.0782)	(0.0941)	(0.217)	(0.282)	(0.152)	(0.0772)
	Admission Rate × Poor	-0.0374	-0.208	-0.0817	-0.0468	-0.209	-0.233
		(0.140)	(0.131)	(0.227)	(0.289)	(0.169)	(0.149)
	Constant	-0.322	-0.388	-0.784	-1.132	-0.525	-0.0860
		(0.238)	(0.278)	(0.559)	(0.697)	(0.428)	(0.261)
	Observations	17,311	17,311	17,311	17,311	17,311	17,311
	R-squared	0.122	0.122	0.118	0.124	0.121	0.118
	lnL	-15014	-15008	-15051	-14989	-15022	-15046
	Chow Test	2.267	1.647	0.951	2.974	1.396	0.849
	Prob ≥ F	0.00211	0.0452	0.512	3.61e-05	0.127	0.636

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Chow tests (F statistics) and corresponding p-values are given in the last two rows. 2. The dependent variable is the marginal returns to human capital investment and it is predicted by using the years of college schooling as the human capital investment T_1. 3. *Poor* indicates whether the household faced binding liquidity constrained when the children were at college age (age of 18). 4. Six liquidity dummies are employed to proxy for the liquidity situations, namely whether the household income was in the bottom quartile of the sample (Column (1)), whether the family lived in a shack (Column (2)), whether the family had a private toilet (Column (3)), whether the family had a water closet (Column (4)), whether the family had access to clean water (Column (5)), and whether the family had electricity (Column (6)).

Table 2.6. Estimation on T_2 Equation Using Investment as T_1

		(1)	(2)	(3)	(4)
		Probit	Tobit	ME Two-part	
	VARIABLES	ME DependOnChild	Parameter Estimates	$P(transfer > 0\|x)$	$E(\ln(transfer)\|x, transfer > 0)$
Parent	Age	-0.00194	0.518***	0.0418***	-0.00751
		(0.00820)	(0.0851)	(0.00704)	(0.0259)
	Age Squared	-8.49e-06	-0.00366***	-0.000289***	-3.89e-05
		(6.14e-05)	(0.000625)	(5.28e-05)	(0.000188)
	Female	0.0491***	0.0809	0.00833	0.0171
		(0.0136)	(0.124)	(0.0112)	(0.0359)
	Years of Schooling	-0.00647***	-0.0436***	-0.00473***	0.0209***
		(0.00173)	(0.0168)	(0.00150)	(0.00482)
	HH Net Income/Num. of Children	-0.0721***	-0.204**	-0.0257***	0.00375
		(0.00992)	(0.0799)	(0.00726)	(0.00697)
	HH Net Financial Wealth/Num. of Children	-3.88e-05	-0.000216	-1.86e-05	9.10e-05*
		(2.87e-05)	(0.000228)	(1.79e-05)	(4.94e-05)
	HH Lifetime Income/Num. of Children	2.52e-06	-0.00199	-0.000191	0.00101*
		(0.000337)	(0.00264)	(0.000205)	(0.000562)
Child	Age	-0.0125***	0.247***	0.0211***	-0.0167
		(0.00442)	(0.0499)	(0.00413)	(0.0139)
	Age Squared	0.000111**	-0.00301***	-0.000254***	0.000142
		(5.25e-05)	(0.000595)	(4.99e-05)	(0.000160)
	Female	-0.189***	1.321***	0.134***	-0.236***
		(0.00901)	(0.101)	(0.00875)	(0.0298)
	Num. of Children	0.0301***	0.265***	0.0306***	-0.130***
		(0.00687)	(0.0694)	(0.00642)	(0.0206)
	Agric. Hukou	0.159***	0.0774	0.0246*	-0.262***
		(0.0149)	(0.147)	(0.0129)	(0.0442)
	Income	0.00914***	0.231***	0.0168***	0.0839***
		(0.00154)	(0.0156)	(0.00176)	(0.00437)
	Human Capital Investment	0.00887**	-0.0548	-0.00256	-0.00621
		(0.00451)	(0.0590)	(0.00443)	(0.00824)
	Fixed Costs ($T_1 > 0$)	0.0140	0.288	-0.00325	0.485***
		(0.0277)	(0.338)	(0.0267)	(0.0662)
	Observations	12,700	12,700	12,700	8,056
	lnL	-8049.102	-28363.358	-7805.221	-12575.384

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. 2. Clustered robust standard errors are at the household level in parentheses. 3. Column (1) contains the marginal effects using the dummy of "whether a household member would like to rely on children" as dependent variable. Column (2) contains the parameter estimates of the Tobit model using log transfer as dependent variable. Column (3-4) are the marginal effects of the participation decision and the amount decision of transfers.

We first consider the marginal effects using the parental investment measure, which is shown in Table 2.6. As expected, the probability of depending on children for support during retirement decreases with parental income and increases with the income of the children. Interestingly, parents who invested in their children' human capital are significantly more likely to receive financial transfers from the children at older ages. The marginal effect for fixed costs, however, is not significant which does not support the fixed costs argument. The Tobit estimation is shown in Column (2), and we find that marginal effects of both the human capital investment and the fixed costs are not significant. One possible explanation for the insignificant effect of fixed costs on T_2 is that many parents faced a binding liquidity constraint. As equation (2.A.19) in appendix 2.A shows, T_2 will only depend on child and parental income in the second period if the parental liquidity constraint (2.1d) was binding in the first period.

Table 2.7. Estimation on T_2 Equation Using Years of College Schooling as T_1

		(1) Probit	(2) Tobit	(3) ME Two-part	(4)
	VARIABLES	ME DependOnChild	Parameter Estimates	$P(transfer > 0\|x)$	$E(\ln(transfer)\|x, transfer > 0)$
	Age	-0.00108	0.532***	0.0423***	0.00133
		(0.00819)	(0.0850)	(0.00705)	(0.0256)
	Age Squared	-1.46e-05	-0.00374***	-0.000292***	-0.000102
		(6.14e-05)	(0.000624)	(5.29e-05)	(0.000186)
	Female	0.0480***	0.0718	0.00787	0.0124
		(0.0136)	(0.123)	(0.0112)	(0.0354)
	Years of Schooling	-0.00826***	-0.0587***	-0.00542***	0.0117**
		(0.00176)	(0.0171)	(0.00152)	(0.00487)
Parent	HH Net Income/Num. of Children	-0.0681***	-0.212***	-0.0262***	0.00206
		(0.0110)	(0.0817)	(0.00727)	(0.00655)
	HH Net Financial Wealth/Num. of Children	-4.51e-05	-0.000228	-1.88e-05	7.05e-05
		(2.96e-05)	(0.000234)	(1.81e-05)	(4.70e-05)
	HH Lifetime Income/Num. of Children	-2.15e-05	-0.00212	-0.000197	0.00104*
		(0.000330)	(0.00267)	(0.000207)	(0.000557)
	Age	-0.0127***	0.245***	0.0206***	-0.0133
		(0.00441)	(0.0499)	(0.00412)	(0.0138)
	Age Squared	0.000119**	-0.00291***	-0.000243***	0.000116
		(5.25e-05)	(0.000595)	(4.98e-05)	(0.000160)
	Female	-0.181***	1.435***	0.141***	-0.184***
		(0.00917)	(0.103)	(0.00888)	(0.0303)
	Num. of Children	0.0330***	0.297***	0.0314***	-0.107***
		(0.00692)	(0.0693)	(0.00643)	(0.0206)
Child	Agric. Hukou	0.178***	0.251*	0.0315**	-0.159***
		(0.0151)	(0.151)	(0.0133)	(0.0455)
	Income	0.00839***	0.220***	0.0163***	0.0798***
		(0.00154)	(0.0155)	(0.00175)	(0.00430)
	Human Capital Investment	0.00815***	0.0405**	-0.000267	0.0566***
		(0.00173)	(0.0172)	(0.00156)	(0.00504)
	Fixed Costs ($T_1 > 0$)	0.0348	1.198***	0.118***	-0.0869
		(0.0271)	(0.254)	(0.0238)	(0.0721)
	Observations	12,700	12,700	12,700	8,056
	lnL	-8035.951	-28339.656	-7788.606	-12525.309

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. 2. Clustered robust standard errors are at the household level in parentheses. 3. Column (1) contains the marginal effects using the dummy of "whether a household member would like to rely on children" as dependent variable. Column (2) contains the parameter estimates of the Tobit model using log transfer as dependent variable. Column (3-4) are the marginal effects of the participation decision and the amount decision of transfers.

However, the Tobit specification is arguably too restrictive in explaining the difference between extensive and intensive margins. Therefore we focus on the double hurdle model in Columns (3) and (4). The results for the probability that parents receive financial transfers from their children in old age are broadly consistent with those obtained using "Depend on this Child" as the dependent variable. The effect of human capital investment, though, is not significant and fixed costs only play a role when looking at the amount transferred. These results follow the theoretical prediction on fixed costs only at the intensive margin.

In Table 2.7, we report the results using the number of years of college schooling measure. The structure and specifications are the same to Table 2.6. In Column (1), we find that, except for fixed costs, all the other marginal effects follow the theory.

In the double hurdle model, we find that, in the participation equation, the marginal effect of the fixed costs is positive and significant. The human capital investment is negative and insignificant, which does not follow the theory. However, human capital investment has a positive and significant effect on the amount of money transferred, while fixed costs do not play a significant role. The results support the theoretical prediction with fixed costs at the extensive margin only when using the years of college schooling measure.

2.6 Conclusion

At both the macro and micro levels, there is evidence of substantial underinvestment in human capital for the cohorts who are now the backbone of the labour market in China. In this chapter, we investigate the role of parental liquidity constraints and fixed costs of education in explaining this underinvestment. We first extend the theoretical model by Raut & Tran (2005), and then we empirically test the model predictions using data from the 2013 China Health and Retirement Longitudinal Study (CHARLS).

From the benchmark theoretical models, we know that there is always human capital underinvestment in presence of bargaining between the parent and the children. The binding liquidity constraint would further reduce the optimal investment amount, and it is reflected through an increase in the marginal returns on the investment. To empirically test whether liquidity constraint matter for the parents, we employ the methods in Zeldes (1989) by testing the difference in the marginal returns between the constrained and unconstrained sample. We measure the liquidity constraints by six indicators constructed from the life history survey of CHARLS. The results show some evidence that binding liquidity constraints can affect the decision of the parents to invest in their children's college education.

We further examine the role of fixed costs, such as tuition fees and children's foregone wages, in explaining the underinvestment in human capital. From the theory we find that if fixed costs affect the parental investment decision in the first period, then this is also reflected in the children's old-age transfer decisions during the second period. Empirically we then test the relevance of fixed costs in the old-age transfer decision at both the extensive and intensive margins. When using actual years of college schooling as measure for human capital investment, we find that fixed costs only affect the extensive margin. However, if we employ the actual

amount of money invested by the parents, then the intensive margin is significantly affected. The sensitivity of the effects using various measures might be explained by the fact that some parents did not face binding liquidity constraints when their children were at schooling ages. In general, the results provide support for the importance of fixed costs in human capital investment decisions.

Overall our findings point to the importance of policy measures aimed at reducing barriers to the private investment in human capital in developing countries, such as poverty-alleviation policies and student-loan programs.

In this chapter, we assume that all children have identical levels of altruism toward their parents. In reality, however, this might not always hold. For example, Bonsang (2009) finds that the share of daughters in the household could serve as an instrument for childrens informal care to the parents. Kalwij et al. (2014) suggest that different children have different tastes for providing care to parents. In this thesis, we consider the number of children in the household as the measure of total altruism to parents. We could further examine whether the share of daughters (as used by Bonsang (2009)) in the household could affect the results. We leave this as future research.

2.A Model Derivations

In this appendix, we solve the models discussed in section 2.2 and derive for the optimal human capital investment decision T_1 (from parent to child) and the optimal old-age transfer decision T_2 (from child to parent). We extend the models in Raut & Tran (2005) by incorporating the parental liquidity constraint and educational fixed costs. The rest of this appendix is organized as follows. In section 2.A.1, we solve the so-called "parent as dictator" model which serves only as a benchmark model in this chapter. Section 2.A.2 solves the main model employed in this chapter, which is called "children with bargaining power". The major difference between these models is who decides the amount of T_2: in the "parent as dictator" model, parents decide T_2; while in the "children with bargaining power" model, children decide T_2.

2.A.1 Parent as Dictator

In the first "parent as dictator" model, the parent chooses consumption c_{p1}, c_{p2}, human capital investment in child T_1, and the old-age transfer T_2. By substituting the constraints (2.1b), (2.1c) and (2.2b) into the parental lifetime utility function, the optimization problem can be expressed as follows:

$$\max_{s,T_1,T_2} \quad u(E_{p1} - (s + nT_1 + nC\mathbf{1}(T_1 > 0)))$$

$$+\beta[u((1+r)s + E_{p2} + nT_2) + \gamma_p u(E_{k2}(T_1,\tau) - T_2)] \tag{2.A.1a}$$

$$\text{s.t.} \quad u(E_{k2}(T_1,\tau) - T_2) - u(E_{k2}(0,\tau))$$

$$+ \gamma_k[u((1+r)s + E_{p2} + nT_2) - u(c^o_{p2})] \geq 0 \tag{2.A.1b}$$

$$s \geq 0 \tag{2.A.1c}$$

$$T_1 \geq 0, \tag{2.A.1d}$$

where c^o_{p2} denotes the optimal level of parental consumption if she does not make an educational investment in her child. All variables are defined in section 2.2. Equation (2.A.1b) describes the child's participation compatibility constraint, and the child obeys the arrangement by the parent if this constraint is not binding. Constraint (2.A.1c) is the parental liquidity constraint.

If human capital investment involves fixed costs, i.e. $C > 0$ in equation (2.1b), then the choice problem is not standard due to the fact that the choice set is non-convex. As we stated in subsection 2.2.3, finding the optimal solution requires

several steps (see, e.g., Hausman, 1980; Cogan, 1981). First, one solves the choice problem presented above under the restriction that the parent does not invest in the human capital of the child ($T_1 = 0$). In that case the parent does not incur any fixed costs and obtains a maximum intertemporal utility level of U_p^0. Then one solves the same problem under the condition that the parent invests in the schooling of the child ($T_1 > 0$) and as a consequence faces fixed costs. Let U_p^1 and T_1^* be the optimal solution in case of investment. The decision to invest in the education of the children involves a utility comparison: if $U_p^1 > U_p^0$, the parent invests T_1^*, otherwise not.

In case that $U_p^1 > U_p^0$, the first order conditions (F.O.Cs) imply the following equalities:

$$u'(c_{p1}) - \mu = (1+r)(\beta + \theta\gamma_k)u'(c_{p2}) \tag{2.A.2}$$

$$\frac{\partial E_{k2}(T_1, \tau)}{\partial T_1} = (1+r) + \frac{\mu}{(u'(c_{p1}) - \mu)} \geq (1+r) \tag{2.A.3}$$

$$(\beta + \theta\gamma_k)u'(c_{p2}) = \frac{\beta\gamma_p + \theta}{n} u'(c_{k2}), \tag{2.A.4}$$

where the Kuhn-Tucker multipliers θ and μ are associated with the participation constraint of the children and the parental liquidity constraint, respectively.

If we assume that the liquidity constraint is not binding (μ=0), equation (2.A.3) simplifies to:

$$\frac{\partial E_{k2}(T_1, \tau)}{\partial T_1} = (1+r). \tag{2.A.5}$$

This is the result reported by Raut & Tran (2005): the parent's education decision for each child is market efficient as the marginal rate of return of investment in child's education equals the market rate of return on other assets $(1 + r)$. Notice the level of investment in each child's education does not depend on other variables, like the degree of parental altruism and the total number of children. This result does not hold if the child has bargaining power. To discriminate between the dictator model and the bargaining model, the identification strategy of Raut & Tran (2005) relies on the sensitivity of T_1 to the number of children n. Equation (2.A.3) says that the marginal rate of return is higher than (1+r) and the human capital investment level is lower if liquidity constraint (2.A.1c) is binding. Notice that in that case T_1 might be negatively related with n. Please also note that equation (2.A.3) holds irrespective of whether the participation constraint (2.A.1b) is binding or not. But from now onwards, we assume like Raut & Tran (2005) that the child obeys the

parental arrangement and his(her) participation constraint is not binding ($\theta = 0$). In that case, the first order conditions (2.A.2) and (2.A.4) simplify to:

$$u'(c_{p1}) - \mu = (1+r)\beta u'(c_{p2}) \tag{2.A.6}$$

$$u'(c_{p2}) = \frac{\gamma_p}{n} u'(c_{k2}), \tag{2.A.7}$$

One can show that these first order conditions also hold if we impose that the parent doesn't invest in the education of their children ($T_1 = 0$). To have a explicit consumption path, we follow Raut & Tran (2005) and assume that the utility function is of logarithm form ($u(x) = \alpha \ln(x)$) and that $\alpha + \alpha\beta = 1$. Then one can show that

$$c_{p2} = \begin{cases} (1+r)\frac{\beta}{1+\beta}\mathrm{Y}(T_1, T_2) = \alpha\beta(1+r)\mathrm{Y}(T_1, T_2) \text{ if } \mu = 0 \ (s > 0) \\ E_{p2} + nT_2 \text{ if } \mu > 0 \ (s = 0), \end{cases} \tag{2.A.8}$$

where $\mathrm{Y}(T_1, T_2) = E_{p1} + \frac{E_{p2}}{1+r} - nT_1 + \frac{nT_2}{1+r} - nC \cdot \mathbf{1}(T_1 > 0)$. Combining (2.A.7-2.A.8) and rearranging yield an explicit solution for T_2:

$$T_2 = \begin{cases} \left[\frac{1}{1+\alpha\beta\gamma_p}\right] E_{k2} + \left[\frac{(1+r)\alpha\beta\gamma_p}{1+\alpha\beta\gamma_p}\right]\left[T_1 - \frac{E_{p1}+\frac{E_{p2}}{1+r}}{n} + C \cdot \mathbf{1}(T_1 > 0)\right] \text{ if } \mu = 0 \ (s > 0) \\ \frac{1}{\gamma_p+1} E_{k2} - \frac{\gamma_p}{\gamma_p+1}\frac{E_{p2}}{n} \text{ if } \mu > 0 \ (s = 0), \end{cases} \tag{2.A.9}$$

where the coefficients are functions of parental altruism γ_p.

2.A.2 Children with Bargaining Power

In the second model "children with bargaining power", the parent first makes investment T_1 in the first period, then the child reacts by making T_2 in the second period. The parent first solves the following optimization problem:

$$\max_{s,T_1} u(E_{p1} - (s + nT_1 + nC\mathbf{1}(T_1 > 0)))$$

$$+\beta[u((1+r)s + E_{p2} + nT_2) + \gamma_p \nu_p(E_{k2}(T_1, \tau) - T_2)] \tag{2.A.10a}$$

$$\text{s.t. } s \geq 0 \tag{2.A.10b}$$

$$T_1 \geq 0, \tag{2.A.10c}$$

then the child solves the following problem in the second period:

$$\max_{T_2} \; \nu(c_{k2}) + \gamma_k u(c_{p2}) \tag{2.A.11}$$

$$s.t. \; c_{k2} = E_{k2}(T_1, \tau) - T_2 \tag{2.A.12}$$

$$c_{p2} = (1+r)s + E_{p2} + nT_2. \tag{2.A.13}$$

Using backward induction, we first solve the optimization problem for the child. We can derive the optimal consumption allocation between the parent and the child which is decided by the child:

$$u'(c_{p2}) = \frac{u'(c_{k2})}{n\gamma_k}. \tag{2.A.14}$$

Equation (2.A.14) is similar to equation (2.A.7), but the difference here is that equation (2.A.7) is obtained from the child's problem rather than the parent's. The parent foresees this result when (s)he solves his(her) optimization problem (2.A.10). Like in the dictator model, the parent faces a non-convex choice set because investment in children's human capital involves fixed costs. Again, the parental decision whether or not to invest in human capital is made on basis of a comparison of lifetime utilities U_p^1 and U_p^0.

In the case that the parents decide to invest in their children's human capital, the first order conditions (F.O.Cs) imply the following equalities:

$$u'(c_{p1}) - \mu = (1+r)\beta u'(c_{p2}) \tag{2.A.15}$$

$$\frac{\partial E_{k2}(T_1, \tau)}{\partial T_1} = \frac{(1+r)}{\gamma_p \gamma_k} + \frac{\mu}{(nu'(c_{p2})\beta\gamma_p\gamma_k)} \geq \frac{(1+r)}{\gamma_p \gamma_k}. \tag{2.A.16}$$

The Euler equation (2.A.15) is exactly the same as in the dictator model if the participation constraint is not binding (cf. equation 2.A.6). If the liquidity constraint is not binding, equation (2.A.16) simplifies to

$$\frac{\partial E_{k2}(T_1, \tau)}{\partial T_1} = \frac{(1+r)}{\gamma_p \gamma_k}. \tag{2.A.17}$$

According to the bargaining model, parents invest less in their children's human capital than according to the dictator model because parents and children are typically imperfectly altruistic ($0 < \gamma_p < 1, 0 < \gamma_k < 1$). Similar to equation (2.A.3) in the dictator model, we observe that a binding liquidity constraint causes parents to

invest even less in their children.

Then we need to derive an explicit solution for T_2. Notice that equations (2.A.15) and (2.A.14) also hold if T_1=0. Again we assume logarithmic utility and $\alpha + \alpha\beta = 1$. Under those assumptions, the closed form solution for c_{p2} can still be described by equation (2.A.8), and the first order condition (2.A.14) can be rewritten as:

$$c_{k2} = \frac{c_{p2}}{n\gamma_k}. \tag{2.A.18}$$

After combining this result with equation (2.A.8) and by rearranging, we obtain the following explicit solution for T_2:

$$T_2 = \begin{cases} \frac{\gamma_k}{\alpha\beta+\gamma_k} E_{k2} + \frac{(1+r)\alpha\beta}{\gamma_k+\alpha\beta}\left[T_1 - \frac{E_{p1}+\frac{E_{p2}}{1+r}}{n} + C \cdot \mathbf{1}(T_1 > 0)\right] \text{ if } \mu = 0 \ (s > 0) \\ \frac{\gamma_k}{\gamma_k+1} E_{k2} - \frac{1}{\gamma_k+1}\frac{E_{p2}}{n} \text{ if } \mu > 0 \ (s = 0). \end{cases} \tag{2.A.19}$$

Different from equation (2.A.9) whose coefficients are only determined by the parent's altruism, the coefficients in equation (2.A.19) are only determined by the children's altruism.

2.B Detailed Sample Selection

We first construct a child-level data set from CHARLS 2013 where each observation is a child.[14] After deleting the dead children and those observations with missing values on the gender (174 missing), date of birth (3,030 missing), education (2,551 missing), *Hukou* status[15](1,132 missing), marriage (1,506 missing), coresiding with family (1,584 missing), income (9,003 missing[16]) and the number of alive children (6,141 missing), in total 19540 observations are left (Originally 30,051 alive children with missing values in covariates are available).

Then we merge the child-level data with the household information. We choose the family respondent as the representative of the household and consider also the information on his(her) spouse if married. For instance, the gender and age information of the household is from the family respondent while the education level of the household is taken from the highest one of the couple,[17] and the household wealth is the total wealth of the couple. At this stage, 10,787 out of 10,803 households are retained. After merging, we further conduct a sample selection as follows. First, the children from those household members who are not family respondent or spouse are dropped,[18] retaining 19,540 observations in total. Second, we drop those observations with missing values on some important variables of the households and drop those children who are still at school. At this stage, 18,995 observations are left. We do not drop those children who are still coresiding with the family even though the directions of the transfers could be ambiguous.[19]

[14] We choose the second wave because it is more updated. Nevertheless, we also conduct robustness check using the first wave and the third wave. We do not choose the third wave namely CHARLS 2015 as our main data set because we also need information in the life history survey in 2014 therefore the retrospective information on the new households in the third wave are not available.

[15] *Hukou* is a special institutional arrangement to identify the origins of households. In general, two kinds of *Hukou* are frequently discussed: the agricultural one and the non-agricultural one. It is a very important institution for Chinese households, since non-agricultural *Hukou* brings about more social advantages than the agricultural one, and the costs of altering the *Hukou* are substantial.

[16] The size of the missing values in income is huge, and it deserves more investigation. Actually, no information on the reasons of the missing is given in CHARLS 2013 or 2015. However, in CHARLS 2011, the number of missing values is 4,212, among which 4,032 are due to "don't know" and 69 due to "refuse to answer. Therefore, the majority of the missing values here are also from the "don't know" type non-response.

[17] We just consider the main respondent and spouse in the households and delete the other household members.

[18] In CHARLS, it is possible to have children from multiple couples in a household. The sampling strategy employs a screening method at the household level, namely asking all the age eligible household members aged 45 above and their spouses.

[19] No consensus has been reached on this issue. Raut & Tran (2005) only have data from non-coresiding parents and children, so they don't consider coresiding children (page 397, the last sentence is section 2.1). Oliveira (2016), however, assumes coresiding is the transfer from children to parents, therefore he

Finally, in order to obtain information on liquidity constraints and alternative measures of permanent income, we merge the data with the life history data set[20] based on household ID and the year when each child reached the college age of 18. At last we reach a sample size of 17,311.

includes coresiding children in the sample (page 5, footnote 25). Ham & Song (2014) drop the coresiding children, since they argue that the directions of transfers in coresiding families are not clear (page 81, line 3 of paragraph 2 in Data section).

[20] We first transform the life history data into an informative unbalanced panel data, then construct the liquidity variables and calculate lifetime income.

Chapter 3

The Impact of the Cultural Revolution on Lifetime Income*

3.1 Introduction

The long-term effect of exposure to adverse events early in life on individual income has important implications for policy makers. A precise evaluation of the total costs throughout one's lifetime caused by these adverse events enables governments to propose more effective remedies. The extant literature employs various events (e.g. wars, regional violence, and earthquakes) to estimate the long-term effects of those events on individual income (see, e.g., Almond et al., 2018). The Chinese Cultural Revolution (CR hereinafter), an adverse political event, is often treated as a targeted event of interest in studies of Chinese populations. The literature using the CR, however, reports mixed results. For example, Zhang et al. (2007), Zhou (2013) and Li et al. (2010) find that the CR had positive effects on individuals' current income. They argue that individuals who lived through the CR may have developed a mental fortitude that has led to the observation of this positive effect. Alternatively, it might be the result of a selection effect–that is, those who were affected were originally from elite social classes. In contrast, another stream of research (see, e.g., Xie et al., 2008) shows that the CR had no significant effects on earnings. A third stream of

*This chapter is based on Wang (2019).

research such as Meng & Gregory (2007) reports negative effects. Negative effects are explained by interrupted human capital accumulation due to the CR. The mixed findings are the result of different research streams focusing on different specific contexts, such as the "sent-down movement" or interrupted education during the CR. However, the most important feature of the CR has not been well studied, namely, the violence. Violence can affect individuals significantly over their life cycles. In addition, due to data limitations, many well-established studies, such as Xie et al. (2008), employ the data sets that cover only certain cities rather than the entire country. Assessing the general impact of the CR on income using a nationally representative sample is worthwhile because it is important to evaluate the cost of this event on the population in this cohort.

In this chapter, we estimate the general impact of the CR on individual lifetime income, giving special attention to two critical aspects. The first is the measure of income. Most traditional literature employs individual income from a single year as the measure (see, e.g., Li et al., 2010). This measure is also often observed in research using population census data. Using only one year of observation fails to capture the general life-cycle effect of the adverse event (see, e.g., Haider & Solon, 2006) and suffers from temporary fluctuations. To overcome this issue, we consider two measures for a lifetime earnings profile as the dependent variables. To capture both the levels and curvature of the profile, we consider the average lifetime income per worked year and the difference of a person's first and last wages before retirement. By using average lifetime income, we smooth the temporary fluctuations. By using the first-last wage differentials, we capture the steepness of the age-earnings profile and thus the effect of on-the-job training. The second problem arises with respect to measures of exposure to the CR. Existing literature characterizes the CR by considering certain isolated events or phenomena that occurred during the CR, such as the sent-down movement or interrupted education (see, e.g., Li et al., 2010; Xie et al., 2008; Zhou & Hou, 1999). In this chapter, we consider two types of exposure to the CR. The first exposure measure is regional violence during the first five years of the CR (1966-1971), which is a prefecture-level measure, and the second exposure measure is whether an individual joined the Red Guard army, which is an individual-level measure.

The China Health and Longitudinal Retirement Study (CHARLS hereinafter) life history survey data set is ideal for this research. The CHARLS is nationally representative and provides complete records of work history and corresponding wage information. From this data set, we are able to construct lifetime panel data to

use in our subsequent analysis. The general results show that, within the CR cohort, the prefectures with more violence during the CR produced lower average lifetime income among males, while having joined the Red Guards increased the average lifetime income for males. The average lifetime income of females within the CR cohort was not significantly affected by either measure. In addition, the first-last wage differential was not affected by violence levels or having been a member of the Red Guards.

The rest of the chapter is organized as follows: Section 3.2 briefly introduces the background of the CR. Section 3.3 reviews the relevant literature. Section 3.4 introduces the data set and variables. Section 3.5 analyses the impact of the CR on lifetime income. Section 3.6 offers some concluding remarks and discussion.

3.2 The CR

The general historical and political explanations of the CR have been documented in several papers (see, e.g., Bonnin, 2006). The CR has affected many aspects of the country and its residents, such as cultural traditions, education system, personal beliefs, and long-term household welfare. In general, the CR can be divided into two subperiods: The first half of the CR consists of the Red Guard army movement and the massive chaos created by the Red Guard army members, and the second half consists of the "rustication program" or the sent-down movement.[1] Figure 3.1 depicts the overall timing of the event. We begin by introducing the second half (i.e. rustication) and then explain the core elements of the first half, which is the primary period we focus on in this chapter.

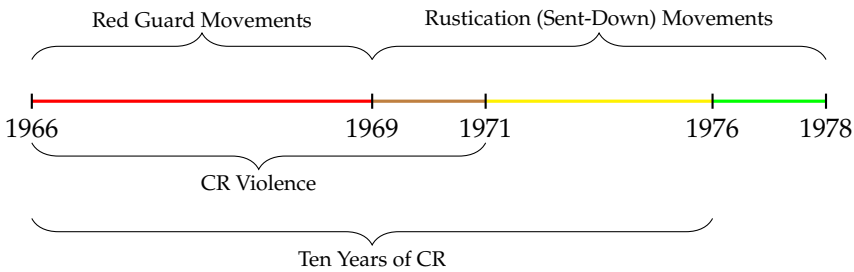


Figure 3.1. Major Events in the CR

[1] The rustication movement forced Chinese youths to the countryside to be reeducated by farmers.

3.2.1 Rustication and College Education during the CR

Most of the extant literature (see, e.g., Xie et al., 2008) studies the effect of the rustication program (or the sent-down movement) during the CR on individuals' current income. Identified as second half of the ten years of the CR, from 1969 to 1976, this period marks a revolution in education for the school-aged youth. Due to the massive interruption of schooling and the chaotic uprising among the young Red Guard members in the first half of the CR, a serious mismatch occurred between an increasingly limited labour force and the number of Red Guard members who experienced interrupted schooling in the second half of the CR. To address the potential massive unemployment, the youth were rusticated to the countryside or to factories. This rustication policy was formally terminated in 1978 even though the CR ended in 1976,[2] so 1977 is also an important year for significant rustication. During the period between 1978 and 1980, the majority of youths sent to work in the countrysides returned to urban areas when the policy ended. Some stayed in the countryside, however, for the rest of their lives.

Another stream of literature (see, e.g., Meng & Gregory, 2007) studies the potential income loss due to missed schooling years (especially for college-level schooling) during the CR. Before the CR, schools were open, and a college education was available for high school graduates who could pass the college entrance examinations. Throughout the CR, however, the college entrance examinations were suspended. Moreover, colleges did not admit new students between 1966 and 1970. Beginning in 1970, recruitment only took place via recommendations by local administrators based on the backgrounds of the candidates, such as whether candidate's family members were poor farmers (politically good) or landlords (politically bad). In total, approximately 940,000 candidates were recruited through this unique channel, and it is now widely believed that candidates in this small cohort were of the lowest quality in terms of their education. At the end of 1977, recruitment via recommendations was terminated, and the college entrance examinations were restored along with a minimum grade requirement and a relaxed age limit. Eligible individuals who could apply for colleges included high school graduates, returning rusticated youths, workers, farmers, demobilized soldiers, and cadres (those who have decent jobs) with at least primary diplomas. Age requirements differed across different groups. Individuals who were in middle or high school in 1966, 1967, and 1968 (called *laosanjie*), and who suffered the most severe interruptions of their schooling,

[2] Nevertheless, a small number of educated youths were still rusticated between 1978 and 1980.

were all eligible to apply. Individuals younger than 25 years of age were also eligible. In 1977, approximately 5,700,000 people applied to take the college entrance exam and fewer than 300,000 were recruited. In 1978, the numbers were 6,100,000 and 402,000, respectively.

3.2.2 The Red Guard Movement and Violence in the CR

In this chapter, we focus on the first half of the CR, and a similar strategy is employed by Bai (2014). While the most commonly recognized ten years of the CR occurred between 1966 and 1976, the CR *stricto sensu*, is defined with historical objectivity as the period from May 1966 to April 1969 (Bonnin, 2006). During this period, the youth, mostly from urban areas, were incited to "rebel" against the country's current leadership and join the Red Guard army. The majority of Red Guards came from the student population at that time. Nonetheless, some young factory workers were also involved. Many schools and factories were therefore closed. Consequently, students faced interrupted education in this period. Even for those students who remained in school, the curricula changed to emphasize the study of the "Little Red Book", a collection of famous quotations from Chairman Mao. At the end of 1968, all the students who were in middle or high school in 1966, 1967, and 1968 (*laosanjie*) were simply graduated, regardless of their grade. After the intense conflicts initiated by the Red Guard, several campaigns began to play a role, most of which were conducted by the military and newly established local governments. Figure 3.2 offers a detailed timeline of this period.

In Figure 3.2, the earliest events, such as "Wudou" (armed battles), "attack military/government", and "suppression", are often classified as collective conflicts and did not involve the military. Most of these events lasted throughout the first half of the CR. Beginning around 1969, events such as "Qingdui" (cleansing of class ranks), the "One Strike-Three Anti" campaign, and the "Anti-May 16" campaign[3] were initiated either by the military or the newly established local government (*geweihui*). In the literature, the first group of events are defined as *"collective conflicts"* and the second group of events are defined as *"campaigns"*. In the empirical research, collective conflicts are often used as the measure of severity of the CR because they occurred before various campaigns. The spatial distributions of conflicts, campaigns, and their associated number of deaths and injuries appear in Figures 3.3 and 3.4,

[3] The *Anti-May 16* event was created by a group of left-wing Red Guards in 1967, and the *Anti-May 16* campaign was initiated to punish these Red Guard members.

respectively.

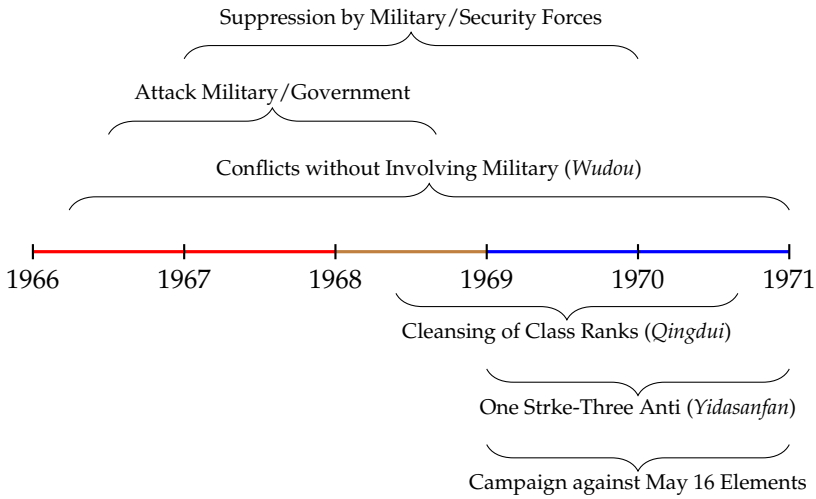
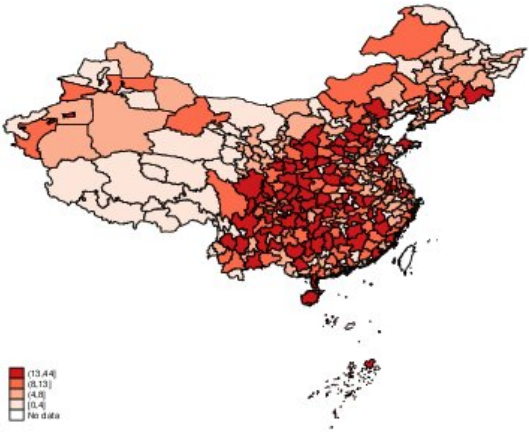


Figure 3.2. Six Named Collective Conflicts/Campaigns

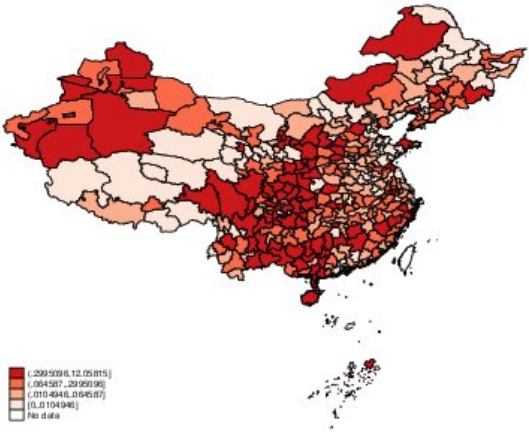
According to Walder (2014), approximately 1.1-1.6 million unnatural deaths occurred, and 20 to 30 million people were the victims of some of political persecution produced by the discord between the insurgents and the authorities during the CR. Human capital is not the only aspect that has been devastated; other critical aspects such as social capital and city infrastructures have also been damaged. Traditional wisdom argues that the rural areas were free from being affected, but this is perhaps not true. Walder & Su (2003) recently demonstrated that the rural areas were also extensively affected, though the effects were delayed compared with the urban areas. In the current empirical analysis, we only use the number of events and deaths from collective conflicts.

3.3 Relevant Literature

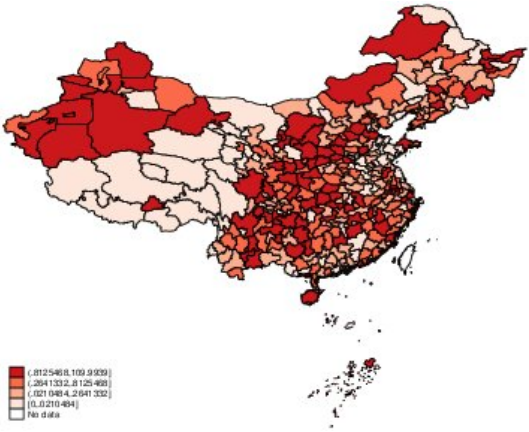
Several studies estimate the impacts of adverse political events on individuals' late-life well-being. For the purpose of this chapter, we examine literature that employs the CR as a natural experiment, with special attention given to research that uses income as the dependent variable. We summarize information on journals, data sets, methodologies, measures of exposure to the CR, and income measures in Table 3.1. Because income is closely related to educational choices, we also include the literature that uses education as the dependent variable into the table.



(a) Number of Conflicts* 1000 per Capita

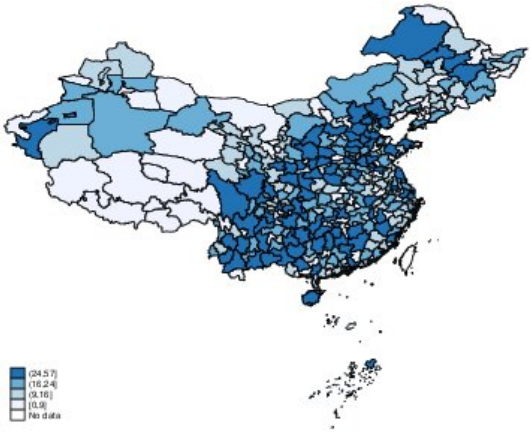


(b) Number of Deaths* 1000 per Capita

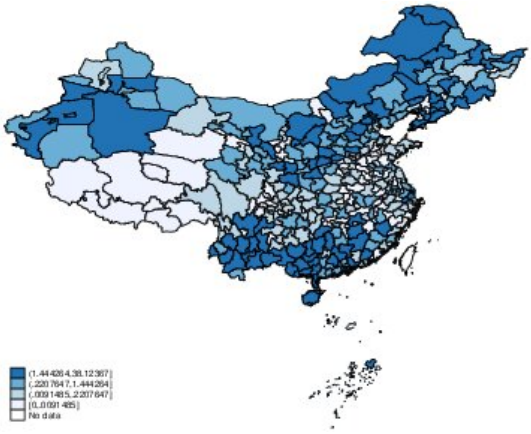


(c) Number of Injuries* 1000 per Capita

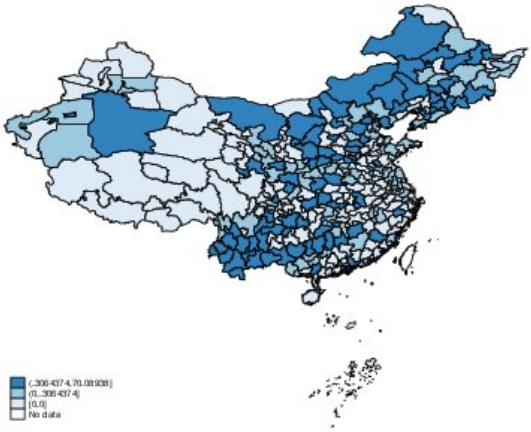
Figure 3.3. Collective Conflicts Violence during the CR 1966-1971



(a) Number of Campaigns* 1000 per Capita



(b) Number of Deaths* 1000 per Capita



(c) Number of Injuries* 1000 per Capita

Figure 3.4. Campaign Violence during the CR 1966-1971

Table 3.1. Literature Review of the CR Effects on Income and Education

Authors (Journals)	Data	Methods	Measures of CR Exposure	Dependent Variable
		Dependent Variable: Income		
Li et al., 2010 (JPE)	Twins Survey[1]	Twin FE	Sent-down to rural	Monthly earnings 2002
Zhou & Hou, 1999 (ASR)	Urban Residence[2]	OLS	Sent-down to rural	Annual income 1978, 1987, 1993
Zhou, 2013 (WP)	CGSS[3]2003	IV	Sent-down to rural	Monthly income 2003
Meng & Gregory, 2007 (IZA-WP)	UHIES[4] HIDS[5]	 LATE	Years of missed schooling CR Cohort 1947-1961	Monthly earnings 1992-2000 Years of schooling
Xie et al., 2008 (SSR)	SFLUC[6]	Sibling FE	Sent-down to rural	Annual income College education Years of schooling
		Dependent Variable: Education		
Deng & Treiman, 1997 (AJS)	1% Census[7]	Trends FE	Year-of-birth in CR	Educational attainment
Meng & Zhao, 2016 (IZA-WP)	CHIPS[8] CULS[9] UREES[10]	IV	Years of missed schooling Unnatural deaths and victims Regional growth compared to pre-CR	Years of schooling
Meng & Gregory, 2002 (EDCC)	UIDS[11] RFPS[12]	Probit	Years of Missed Schooling	Educational attainment
Bai (2014) (WP)	Gazetteers[13] Census 1982, 1990, 2000	IV	Unnatural Deaths and Victims	Regional development Education
Han et al., 2019 (JHC)	CHNS[14] UHS[15]	OLS,RD	CR cohorts 1947, 1948-50, 1951-55, 1956-57, 1958-61	Re-schooling
Fan, 2017 (ALCR)	CGSS[3]1990/2000	Semi-Para. Cox	CR cohorts 1947-1957	School returning timing

Note: 1. The Chinese Twin Survey was conducted by the Urban Survey Unit of the National Bureau of Statistics of China in five cities in 2002. 2. The sample of urban residents was collected in 1993 and 1994 in 20 cities from six provinces: Hebei, Heilongjiang, Gansu, Guangdong, Jiangsu, and Sichuan. 3. The China General Social Survey (CGSS) is the earliest nationally representative continuous survey project, launched by the Department of Sociology of the Renmin University of China and the Survey Research Center of Hong Kong University of Science and Technology in 2003. The first phase consists of 2003-2006 and 2008, and the second phase consists of 2010-2019. 4. The Urban Household Income and Expenditure Survey (UHIES) is conducted by the National Bureau of Statistics of China on an annual basis. The sample includes 15 provinces and lasts from 1992 to 2000. 5. The Household Income Distribution Surveys (HIDS) is conducted by the Institute of Economics at the Chinese Academy of Social Sciences. The surveys were conducted in 1995, 1999, and 2002 and cover ten, six and ten provinces, respectively. 6. The Study of Family Life in Urban China (SFLUC) was conducted in 1999 in three cities: Shanghai, Wuhan, and Xi'an. The authors collected the data by themselves. 7. The 1% sample of census is drawn from the 1982 population census of China. 8. The Chinese Household Income Project Survey (CHIPS) was conducted by China Institute for Income Distribution with the help of National Bureau of Statistics and now has five waves: 1988, 1995, 2002, 2007, and 2013. It covers all the provinces and both rural and urban households in China. 9. The China Urban labor Survey (CULS) was conducted by the Institute for Population and labor Economics at the Chinese Academy of Social Sciences with provincial and municipal branches of National Bureau of Statistics of China, which was conducted from November 2001 to January 2002. It covers five large cities: Fuzhou, Shanghai, Shenyang, Wuhan and Xi'an. 10. The Urban Residents Education and Employment Survey (UREES) was fielded in 2005 and consists of urban households in the following provinces: Beijing, Shanxi, Liaoning, Zhejiang, Anhui, Hubei, Guangdong, Sichuan, Guizhou, Shaanxi, and Gansu. 11. The Urban Income Distribution Survey (UIDS) was conducted nationwide by Institute of Economics at the Chinese Academy of Social Sciences in 1989. 12. The Residents and Floating Population Survey (RFPS) was fielded in Shanghai in 1995 by the Institute of Population Studies at the Shanghai Academy of Social Sciences. 13. The gazetteers recorded local histories, ranging broadly from geography, economics to politics in China; this data was collected by Walder (2014). 14. The China Health and Nutrition Survey (CHNS) was collected by the Carolina Population Center at University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention; the first round was fielded in nine provinces in 1989, and six additional panels were collected in 1991, 1993, 1997, 2000, 2004, and 2006. 15. The Urban Household Survey (UHS), which was designed to be nationally representative, was collected by National Bureau of Statistics of China from 1986 to 2006. It is time series of cross-sections.

The extant literature commonly employs cross-sectional income or wages as the dependent variable, but different research employs different measures of exposure to the CR. Therefore, the identification strategies and the targeted populations vary across studies. A typical example is given by Li et al. (2010), who employ a twins data set. They consider whether one of the identical (monozygotic) twins who was sent down to the rural area had a lower income than the other who was not sent down. Using within-twins variation, they find that one year spent in the countryside was associated with a 3.4% higher monthly wage in 2002. A similar data set is employed by Zhang et al. (2007). They study the effect of the CR on income by comparing returns to schooling between a CR cohort (born between 1947 and 1969) non-CR cohorts. Employing hourly wages as the measure of income, they find that the CR cohort does not have lower returns to schooling. Instead of using twins data, Zhou & Hou (1999) conduct a survey in urban areas of six provinces in China. They consider annual income across three years (1978, 1987, and 1993) as the dependent variable. They also restrict the sample to individuals who entered into the labour market between 1967 and 1968. They identify the impact of the CR by comparing those who were sent down to the rural areas with those stayed in urban areas. The results show that the CR negatively affected income, but the effects are insignificant. Zhou & Hou (1999) attribute the insignificant results to a policy issued in 1980s to facilitate the sent-down youth to labour market (e.g., government helps the sent-down to find jobs). Instead of using an urban survey, Zhou (2013) employs nationally representative survey data. Similar to Zhou & Hou (1999), Zhou (2013) compares the monthly income of the sent-down and the stayed-in-urban groups. The difference is that Zhou (2013) identifies the 1948-1963 cohort as the CR cohort and employs a nationally representative sample. The evidence shows that the income of males was positively affected by the sent-down experience while the income of females was not. Xie et al. (2008) disaggregate the effect of the sent-down experience into groups with different durations of stay. By investigating within-sibling variations in the sent-down experience, they find that the sent-down experience actually increased annual income if the sent-down duration was less than six months. If the duration of stay was longer than six months, the sent-down experience decreased income. Meng & Gregory (2007) calculate years of missed education for each cohort and estimate the earnings loss by multiplying the years of missed schooling and returns to schooling. They find that the CR cohort indeed has more years of missed schooling, but the calculated earnings loss based on estimated returns to schooling is not significantly different between the CR and the non-CR groups.

In summary, the existing research has investigated the impact of the CR on income through specific events that occurred during the CR. In addition, the measures of income are mainly from cross-sectional data sets. In this chapter, we estimate the general effect of the CR on lifetime income. Then we compare our results to the estimates in the literature.

3.4 Data and Descriptive Statistics

3.4.1 Data Source

Our primary data set is the life history survey of the CHARLS, conducted in 2014. The CHARLS is a HRS (Health and Retirement Study)-type ageing survey that has a common base with SHARE (Survey of Health, Ageing and Retirement in Europe) and ELSA (English Longitudinal Study of Ageing). It aims to collect a nationally representative sample of Chinese residents aged 45 years and more on a biannual basis. The baseline survey was conducted in 2011, and now it has three regular waves and one special survey on the life history of the respondents. The respondents in the previous two waves (i.e., 2011 and 2013) were asked to provide retrospective information on their family background, education, health, health care, wealth, and work from the childhood until the interview year. One advantage of this survey is the minimized error caused by recall problems of the elderly because it uses calendars that are anchored to some important events in history, such as the CR or certain major floods in local areas (Chen et al., 2017). In addition to the life history survey data, wage information in Wave 1 (2011), Wave 2 (2013), and wave 4 (2015) are also used for the sake of complete information and to double-check whether the wage information is correct.

We also employ another historical data source on the regional violence associated with the CR–namely, the China Political Events Dataset, 1966-1971, provided by Walder (2014). Walder (2014) constructs county-level CR violence data by collecting information in local annals (county or city gazetteers), including conflicts, injuries, and deaths. The data also provide information on the extent of potential levels of underreporting violence, such as the number of words used in the description of the CR history. We use the information about the number of conflicts, injuries, and victims from 1966 to 1971, the prominent period in the CR for the purposes of our research. Because the data set is collected from county-level gazetteers, the variables

are valid at the county-level. The life history survey of the CHARLS, however, provides only the prefecture-level of individuals' residence history. Therefore, we link the individuals to the prefecture-level cities.[4] This data set is also used by Meng & Zhao (2016) and Bai (2014).

3.4.2 Variable Construction

Lifetime Income

The core dependent variable in our analysis is lifetime income, which we calculate using self-reported annual income or monthly wages.[5] To capture both the levels and the curvature of lifetime income profiles, we construct two measures for lifetime income. The first is the average lifetime earnings per worked year; we describe the calculation procedures below, and they are similar to those used in Alessie et al. (2013). The respondents are asked to provide all the jobs they had that lasted at least six months, and each job is anchored by start and end dates. For each job episode, there are some missing values for either start or end dates, so we impute those missing values. Each job, depending on its duration, has different wage information. For the first job as well as other jobs that lasted more than five years but fewer than 20 years, the wages in the first and the last years are recorded. For the jobs that lasted more than 20 years, the wages in the first, the middle, and the last years are recorded. Together with the current wage information from waves 2011, 2013, and 2014, wage paths are obtained using linear interpolation by assuming constant wage growth rates. In the original data set, there are some missing values, which we impute using the strategy explained in appendix 3.A. We adjust the price levels to 2010 Chinese yuan using the urban-rural-specific province-level annual Consumer Price Index from 1960 to 2015, obtained from the National Bureau of Statistic in China. The constructed annual income is shown in Table 3.2. Table 3.2 reports four types of information: age, overall average annual income for all individuals, average annual income for males, and average annual income for females. The first row of the three-row block of each year for each variable provides the sample size, the second row provides the mean, and the third row provides the standard deviation. In

[4] Because the administrative divisions now are different from the 1960s and 1970s, resulting in some counties not belonging to certain prefectures consistently, we match the current prefectures (the division codes in CHARLS) to the multiple counties in the 1960s and 1970s (the division codes in Walder (2014)).

[5] In the data set, respondents are asked information about their lifetime job tenures lasting at least six months, such as the job category (e.g., own agricultural production and business activities, agricultural employment, nonagricultural employment, nonagricultural self-employment, unpaid household business help, and army), job start and end dates, hourly labour and wages.

general, the income paths follow the statistics from the National Bureau of Statistics of China closely.

Lifetime income should also include current pension benefits if the individuals are retired or future pension benefits if they are yet to retire. In China, however, the problem of calculating pension benefits arises because the majority of the population, who are from rural areas, are not covered by any work-related pension schemes. Even for those who began to participate the New Rural Pension Scheme initiated in 2009, the age of retirement is rather unclear. Thus, the calculation of lifetime income presents a challenge. To address this problem, we ignore the future pension plans and use the lifetime income individuals have already earned. Because some individuals in the sample are still working, we calculate the average lifetime income per worked year. Therefore, individuals who have retired are comparable with those who have not. The second measure–namely, first-last wage differentials, which is calculated by subtracting the log of first wage from the log of last wage of individuals–represents the proportional changes in the wages between a person's first and current jobs. This measure is constructed to capture the curvature of the lifetime earnings profiles. A higher wage differential demonstrates a relatively lower first wage compared with a higher last wage, indicating the potential high level of reinvestment of human capital during work after the CR. This reinvestment is mainly through on-the-job training.

Exposure to the CR

Ideally, the main explanatory variable should be the treatment that individuals who have experienced the CR and the corresponding controls are naturally the individuals who have not experienced the event. Because the CR affected almost the whole country, it is difficult to find the counterfactual groups who were not affected at all. Nevertheless, we can still employ three identification strategies by using three measures of exposure to the CR. The first identification strategy employs the measure of the CR cohort dummy. We can identify the impact of the CR by comparing the conditional income difference between the CR cohorts and the non-CR cohorts. Similar to traditional literature such as Meng & Gregory (2007), the dummy is constructed based on whether the individual was born between 1946 and 1961.[6] Because this cohort was most likely to experience the shutdown of

[6] Meng & Gregory (2007) define the CR cohort as the collection of 15 cohorts between 1947 and 1961. These 15 cohorts suffer from potential interrupted education. We consider one extra year ahead to capture

Table 3.2. Annual Income, 1981-2010: Full Sample and by Gender

Year	Age	All	Male	Female	Year	Age	All	Male	Female
1981 *Obs.*	14,478	14,303	7,211	7,092	1996 *Obs.*	17,628	17,293	8,580	8,713
Mean	27.72103	7,159.376	8,841.826	5,448.696	*Mean*	39.56422	6,586.335	8,856.648	4,350.677
S.D.	8.48599	22,558.05	28,543.57	13,861.07	*S.D.*	9.441301	19,703.2	25,820.58	10,195.64
1982	15,092	14,920	7,507	7,413	1997	17,583	17,262	8,549	8,713
	28.22316	7,899.564	9,788.934	5,986.237		40.49736	6,814.184	9,140.23	4,531.92
	8.68471	22,628.16	28,983.87	13,139.07		9.430995	21,921.41	28,930.58	10,971.06
1983	15,578	15,391	7,727	7,664	1998	17,521	17,199	8,516	8,683
	28.82411	8,783.144	11,055.31	6,492.299		41.43337	7,297.487	9,797.97	4,845.095
	8.834886	24,238.38	31,140.65	13,846.95		9.417657	24,954.45	33,073.55	12,189.17
1984	15,984	15,788	7,936	7,852	1999	17,451	17,125	8,474	8,651
	29.49468	9,668.497	12,230.24	7,079.348		42.37872	7,959.699	10,715.94	5,259.849
	8.955553	27,527.81	34,459.89	17,613.56		9.406149	28,773.22	38,348.52	13,553.65
1985	16,370	16,170	8,120	8,050	2000	17,402	17,077	8,441	8,636
	30.1796	9,703.309	12,377.51	7,005.853		43.33416	8,465.887	11,412.42	5,585.888
	9.076376	27,935.4	35,500.18	16,791.14		9.391411	31,430.95	42,132.19	14,205.51
1986	16,703	16,498	8,260	8,238	2001	17,307	16,954	8,372	8,582
	30.90056	10,029.09	12,909.4	7,141.084		44.26336	8,980.588	12,168.75	5,870.439
	9.189459	27,830.15	35,450.66	16,565.94		9.376955	34,884.57	47,096.13	14,861.75
1987	16,967	16,747	8,392	8,355	2002	17,248	16,907	8,346	8,561
	31.63547	10,215.61	13,312.87	7,104.639		45.22756	9,643.165	13,140.56	6,233.606
	9.266397	35,073.3	46,625.87	16,215.82		9.374357	38,741.2	52,544.78	15,782.26
1988	17,196	16,974	8,482	8,492	2003	17,173	16,831	8,298	8,533
	32.41382	9,097.036	11,933.41	6,264.001		46.179	10,064.98	13,750.65	6,480.809
	9.342662	28,321.08	37,824.17	12,580.57		9.369128	41,696.79	56,906.95	15,946.72
1989	17,370	17,118	8,548	8,570	2004	17,078	16,761	8,250	8,511
	33.22861	8,311.83	10,951.18	5,679.258		47.12993	10,374.7	14,278.45	6,590.662
	9.387296	24,073.18	31,886.18	11,384.95		9.372928	44,773.13	61,529.44	15,784.97
1990	17,499	17,259	8,612	8,647	2005	16,960	16,650	8,200	8,450
	34.09618	8,778.265	11,611.97	5,956.029		48.08296	10,727.54	14,833.1	6,743.444
	9.429264	24,238.64	31,357.42	13,320.73		9.37589	46,820.19	64,632.03	15,288.1
1991	17,580	17,301	8,623	8,678	2006	16,835	16,546	8,133	8,413
	34.98999	9,043.582	11,954.6	6,151.016		49.02709	10,914.78	14,851.18	7,109.4
	9.453399	23,787.73	30,079.42	14,572.74		9.387844	25,024.05	27,683.21	21,478.39
1992	17,629	17,345	8,634	8,711	2007	16,724	16,464	8,072	8,392
	35.91355	9,071.547	12,035.14	6,134.155		49.98344	10,968.52	15,099.59	6,994.973
	9.462351	23,783.31	30,200.48	14,318.32		9.398785	22,851.17	28,018.82	15,400.57
1993	17,654	17,346	8,637	8,709	2008	16,632	16,388	8,023	8,365
	36.81511	8,435.555	11,246.16	5,648.185		50.94835	11,000.91	15,248.77	6,926.721
	9.466224	21,731.4	27,589.89	13,043.89		9.410438	24,244.13	30,389.82	15,228.4
1994	17,665	17,347	8,627	8,720	2009	16,523	16,290	7,969	8,321
	37.73122	7,315.056	9,780.657	4,875.75		51.92592	11,552.73	16,018.49	7,275.878
	9.461006	19,524.77	25,052.49	11,203.21		9.428646	27,794.27	35,639.95	16,080.74
1995	17,643	17,316	8,596	8,720	2010	16,423	16,236	7,930	8,306
	38.65063	6,701.698	8,993.371	4,442.613		52.89728	11,644.8	16,204.2	7,291.799
	9.446832	19,358.27	25,179.4	10,437.04		9.444695	31,047.61	40,702.23	16,245.45

Note: In this table, we show some descriptive statistics of annual income between 1981 and 2010 from the life history survey of the CHARLS. In each three-row cube, we exhibit the number of observations, sample means, and standard deviation of each variable. By comparing the table with the statistics released by the National Bureau of Statistics of China, we find that the data are externally valid. The table also exhibits a substantial gender gap in income.

schools, the altering of curriculum, joining of the Red Guard army, and the sent-down movements, this assignment of the treatment is the simplest and, of course, the most imprecise one. The second peril is that it depends highly on the functional approximations of the cohort trends. Therefore, the results might be sensitive to the specifications with different polynomial orders of the year-of-birth term. Third, we cannot evaluate the impact of the CR on individuals within the CR cohort, because there are a lot of heterogeneities within the CR cohort. So we will not show this result in the main text, but it is available on request.

Therefore, we mainly focus on the second and the third groups of measures of exposure. The second group of measures is the regional violence during the CR. Not every region was affected by the CR in the same way (see, Figure 3.3), so the cross-regional variation can be employed to identify the impact of the CR. Thus, we introduce the prefecture-level violence caused by the CR as another treatment to help explain the variations in lifetime income. Specifically, we employ the number of collective conflicts per thousand population and the number of unnatural deaths due to the CR per thousand population. We make use of the conflicts and deaths from the following types: (1) armed battles between insurgents factions (*Wudou*), (2) attacks on military forces by insurgents, (3) attacks on government offices by insurgents, (4) suppression of insurgents by military/security forces, and (5) other conflicts involving insurgents. Because the violence counts account for all the events between 1966 and 1971, they are time invariant. Descriptive statistics related to the number of conflicts and deaths appear in Table 3.3. In addition to the collective conflict variables, we provide the variables from campaign events for comparison. Finally, we summarize the regional control variables in the data: the population sizes of the "cadre"[7] and urban residents in each prefecture at the start of the CR.[8]

We merge the violence data with the respondents in CHARLS through the residence information between 1966 and 1971. The final measure of exposure to the CR, especially the CR cohort, is the Red Guard army dummy. This represents an individual who had (or had not) joined the Red Guard army. It is a dummy for the "injurer" group during the CR because the Red Guard army was one the most important instigators of chaos during the CR.

potential early schooling. The results are robust with respect to these minor changes in the time window.

[7] "Cadre" refers to those who have management positions or political power.

[8] In the original data set, there are 354 prefectures.

Table 3.3. Descriptive Statistics of Prefectures

Variables	(1) N	(2) Mean	(3) S.D.	(4) Min	(5) Max	(6) Quartile 1	(7) Median	(8) Quartile 3
Cadre population	345	15,797	17,809	0	150,551	5,759	11,899	19,317
Urban population	345	331,617	527,448	0	5.270e+06	90,958	192,492	360,330
Conflicts	345	9.441	7.700	0	44	4	8	13
Campaigns	345	17.88	11.90	0	57	9	16	24
Deaths (per thousand)	345	2.061	4.455	0	39.47	0.0754	0.532	1.984
Deaths due to conflicts (per thousand)	345	0.404	1.119	0	12.06	0.0102	0.0645	0.294
Deaths due to campaign (per thousand)	345	1.657	4.152	0	38.12	0.00936	0.222	1.446

Note: 1. Variables cadre population and urban population are the counts of people in positions with power and with urban *hukou* status, respectively. 2. Conflicts measure how many collective battles or confrontations occurred without involving the military or social security forces. 3. Campagins measure the number of conflicts that were driven by the government or military forces.

Control Variables

We also construct some demographic controls, such as age (or year of birth), gender, first *Hukou* status, and whether the respondents have ever married. Moreover, we employ several control variables related to family backgrounds. First, we construct the parental education background. Intergenerational mobility has been revealed to have a significant impact on the offspring's income through both education and innate ability. Therefore, controlling for parental education can mitigate the potential bias due to omitted innate ability. Second, we consider the family's relative financial situations compared with the average community/village and whether the food supply was insufficient when the respondents were 17 years of age. This information serves as a proxy for childhood conditions. Third, we employ three variables related to family economic situations during the Great Chinese Famine (1958-1961): whether family members were starving, whether any members starved to death, and whether the family fled during the famine when the individual was between 0 and 5 years of age. These dummies are important because among the CR cohort members, some could have experienced the Great Chinese Famine between 1959 and 1961 when they were between 0 and 5 years of age. As a result, their abilities could have been adversely affected. Fourth, as an important background control, we consider whether the respondents received monetary or suggestive help from family members when they were young adults. Receiving help, either monetary or nonmonetary, can enhance the earnings capabilities of those who were self-employed.

3.4.3 Sample Selection

In total, 20,654 respondents have reported their information, but the valid sample size drops to 20,622 after we exclude missing values in the variables of year of birth (20,623 were not missing) and gender (20,652 non-missing). In addition, we drop two individuals due to wrong values in year of birth (143- and 146-year-old individuals). Next, we drop 696 individuals who never worked. Among this group, 197 individuals report no work history, and 499 individuals report that they had worked but the data reveals no job tenures. We also drop individuals with missing values for current *Hukou* and marriage status.[9] At this stage, we have 19,384 total respondents. We then construct a lifetime panel data and fill in the eduction and work histories of each individual from their year of birth to the year of the interview. The total number of individual-year observations is 1,158,838. Among the whole sample, 349,762 individual-year observations do not have any work information, and of these, 10,833 observations are gap years for all the individuals and the rest are the individual-year observations during nonworking ages (either before or after the working years). Among the 10,833 gap-year observations, 8,923 observations represent unemployment status, and 1,910 observations have no information but we also treat these observations as unemployed.[10] The total number of observations in working years is 819,909, and the number in working years with actual working information is 809,076. Furthermore we find that not all individuals who report working years also report complete wage information. We exclude all individuals who reported no wage information. Finally, we are left with 739,298 individual-year observations during working years without unemployment or breaks in careers, and the overall sample size is 1,069,884. This amounts to 18,198 individuals. After dropping the observations with missing administrative division codes, we finally obtain 17,614 individuals. Within the CR cohort, the sample size is 8,808.

3.4.4 Descriptive Statistics

The descriptive statistics of all the variables with final a sample size 17,614 appear in the Table 3.4. In our subsequent analysis, we focus more on the CR cohort sample;

[9] The majority of individuals who reported no work experience also have missing values in these important variables. After we exclude the observations with missing values, 49 individuals reported no work experience.

[10] We do not impute the earnings of those with an unemployment status with 80% of the adjacent wages as Alessie et al. (2013) do, because unlike developed economies, which pay approximately 80% of the wages of the previous job if individuals become unemployed, individuals in developing economies, such as China, receive nearly no public transfers when they are unemployed.

these descriptive statistics appear in Table 3.5.

To generate an impression of the overall wage patterns in the final sample, we estimate and plot the age-earnings profiles by gender, education, and CR/non-CR groups. We obtain age-earnings profiles by regressing the log-annual income on the fourth order polynomial in age for different groups with individual fixed effects. Panel A of Figure 3.5 shows the age-earnings of males by education levels. Panel B shows the equivalent for females. In general, the patterns follow the intuition that less educated individuals have a flatter age-earnings profile (see, e.g., Alessie et al., 2013). In Panel C, we estimate the profiles by gender and CR/overall cohorts, and we find that males dominate females in terms of income and that the CR cohorts are also inferior to the overall cohorts.

3.5 Empirical Models

We want to estimate the impact of the CR on individuals' lifetime income profiles. Specifically, we consider two variables to describe the lifetime earnings profiles: average lifetime income per worked year and the first-last wage differential (log last-log first). Typically it is common to use two types of variations to identify the effects of exposure to CR: the cross-cohort variations and the regional variations within the CR cohort. In the current context, however, the variations across CR and non-CR cohorts are not plausible. Because the pre-CR cohort (born before 1946) was also exposed to the CR (post-schooling period) and the great famine (*in utero* or childhood period), and the post-CR cohorts (born after 1962) experienced the opening and reform policy and other influential policies during the critical periods of their lifetime.

To avoid these potential confounding factors, we only use the sample of the CR cohort members (1946-1961). That said, we only use the cross-prefecture variations to identify the effects. As a robustness check, in the online appendix (available on request), we also make use of the cross-cohort variations only. The results show that the CR members have a lower lifetime income than the non-CR cohort members. Although we do not incorporate this variation in our model, the results from the robustness check reveal that the effect we obtain is actually a lower bound. If we found nontrivial negative effects, the true effects would be larger by including other non-CR cohort members.

In the following models, we consider two measures of exposure to the CR as our

Table 3.4. Descriptive Statistics of All Respondents (Cross-Sectional)

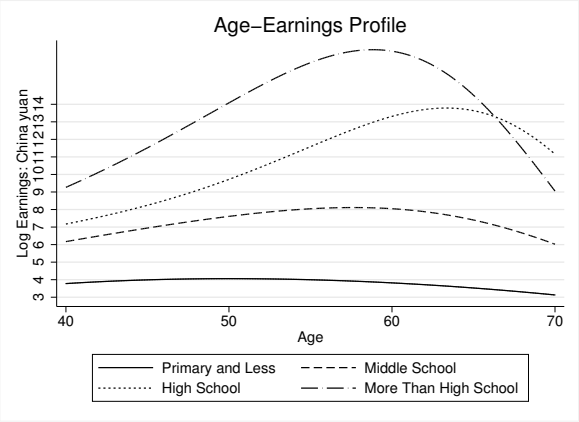
Variables	(1) N	(2) Mean	(3) S.D.	(4) Min	(5) Max	(6) Quartile 1	(7) Median	(8) Quartile 3
				Core Variable				
Lifetime income[1]	17,614	335,620	754,169	147.6	5.068e+07	77,366	175,015	373,442
Average lifetime income (per worked year)	17,614	8,968	21,683	3.433	1.810e+06	1,840	4,303	9,979
Current (last) income	17,614	13,027	55,582	1.234	4.767e+06	1,324	4,406	14,904
First income	17,614	5,869	30,433	0.576	1.837e+06	427.7	1,322	3,722
First-last differential	17,614	7,160	56,021	-1.829e+06	4.766e+06	-377.3	1,840	9,966
log last-*log* first	17,614	1.148	1.964	-7.160	9.352	-0.327	1.133	2.519
Red Guard	17,614	0.193	0.395	0	1	0	0	0
				Demographic Variables				
Age	17,614	58.33	9.449	43	94	50	58	65
Female	17,614	0.501	0.500	0	1	0	1	1
First Hukou is agricultural	17,614	0.897	0.304	0	1	1	1	1
Have ever married	17,614	0.990	0.0981	0	1	1	1	1
				Childhood and Family Background				
Illiterate parent	17,614	0.818	0.385	0	1	1	1	1
Better financial situations[2]	17,614	0.605	0.489	0	1	0	1	1
Family starving to death[3]	17,614	0.102	0.303	0	1	0	0	0
Family starving during famine[4]	17,614	0.763	0.426	0	1	1	1	1
Family fleeing during famine	17,614	0.0807	0.272	0	1	0	0	0
Insufficient food before age 17	17,614	0.673	0.469	0	1	0	1	1
Have received monetary help[5]	17,614	0.0773	0.267	0	1	0	0	0
Have received idea help	17,614	0.137	0.344	0	1	0	0	0

Note: 1. The lifetime income is constructed based on each job tenure throughout the person's history, and the current income is the latest annual income if not yet retired or the last annual income before retirement if already retired. 2. The respondent is asked how, before age 17, his/her family background was compared with the average family in the same community/village. 3. The variable is obtained by asking respondents who were between the ages of 0 and 5 years, between 1958 and 1961 (The Great Chinese Famine), if any family member starved to death. 4. The respondent was asked that if any family experienced starvation. 5. This variable is obtained by asking whether the respondent received financial support for business when (s)he was young.

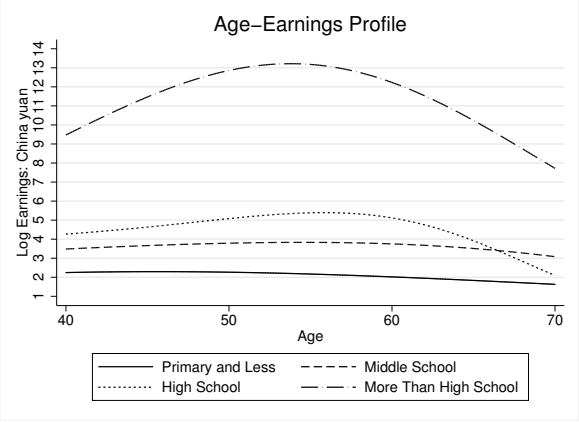
Table 3.5. Descriptive Statistics of the CR Cohort Members (Cross-Sectional)

Variables	(1) N	(2) Mean	(3) S.D.	(4) Min	(5) Max	(6) Quartile 1	(7) Median	(8) Quartile 3
				Core Variable				
Lifetime income[1]	8,808	320,951	695,493	147.6	2.254e+07	78,107	166,914	346,934
Average lifetime income (per worked year)	8,808	7,510	16,288	3.433	505,290	1,725	3,790	8,115
Current (last) income	8,808	11,577	58,551	1.748	4.767e+06	1,322	4,171	11,908
First income	8,808	3,744	20,598	0.576	1.314e+06	327.6	890.4	2,451
First-last differential	8,808	7,833	59,663	-678,473	4.766e+06	-5.164	2,259	9,217
log last-*log* first	8,808	1.417	1.963	-7.160	9.352	-0.00563	1.465	2.760
Red Guard	8,808	0.299	0.458	0	1	0	0	1
				Demographic Variables				
Age	8,808	60.55	4.241	53	68	57	60	64
Female	8,808	0.497	0.500	0	1	0	0	1
First Hukou is agricultural	8,808	0.889	0.314	0	1	1	1	1
Have ever married	8,808	0.989	0.103	0	1	1	1	1
				Childhood and Family Background				
Illiterate parent	8,808	0.869	0.338	0	1	1	1	1
Better financial situations[2]	8,808	0.594	0.491	0	1	0	1	1
Family starving to death[3]	8,808	0.00840	0.0913	0	1	0	0	0
Family starving during famine[4]	8,808	0.101	0.301	0	1	0	0	0
Family fleeing during famine	8,808	0.0110	0.104	0	1	0	0	0
Insufficient food before age 17	8,808	0.0903	0.287	0	1	0	0	0
Have received monetary help[5]	8,808	0.0662	0.249	0	1	0	0	0
Have received idea help	8,808	0.119	0.323	0	1	0	0	0

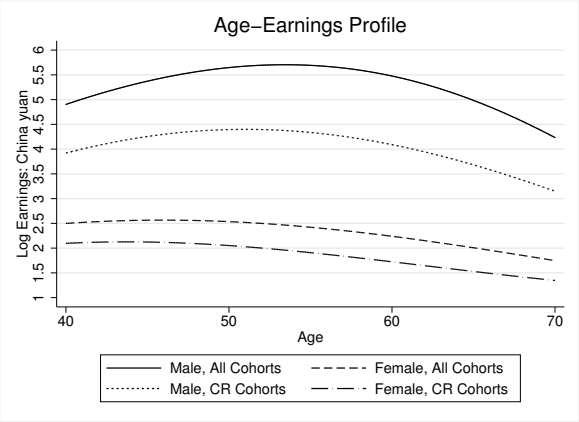
Note: 1. The lifetime income is constructed based on each job tenure throughout the person's history, and the current income is the latest annual income if not yet retired or the last annual income before retirement if already retired. 2. The respondent is asked how, before age 17, his/her family background was compared with the average family in the same community/village. 3. The variable is obtained by asking respondents who were between the ages of 0 and 5 years, between 1958 and 1961 (The Great Chinese Famine), if any family member starved to death. 4. The respondent was asked that if any family experienced starvation. 5. This variable is obtained by asking whether the respondent received financial support for business when (s)he was young.



(a) Age-Earnings Profile of Males by Educational Level



(b) Age-Earnings Profile of Females by Educational Level



(c) Age-Earnings Profile by Gender and Cohort

Figure 3.5. Age-Earnings Profile by Characteristics

main explanatory variables: (1) the prefecture-level violence between 1966 and 1971 that an individual experienced and (2) whether the individual joined the Red Guard army. The first measure is a proxy for the general effects of the CR on the population at large, while the second assesses a specific group of people who actively joined the CR. In the next subsection, we provide more detailed specifications.

3.5.1 Regional Violence

We first use prefecture-level violence to identify the impact of the CR. Each individual is merged to a prefecture in which he or she lived between 1966 and 1971. The estimated model is as follows:

$$y_{igc} = \alpha + \beta_1 Violence_g + \mathbf{X}_i\gamma + \mathbf{Z}_g\eta + \lambda_c + \varepsilon_{igc}, \tag{3.1}$$

where the violence variable, $Violence_g$, includes collective conflicts and the associated deaths between 1966 and 1971; thus, it is time-invariant. The collective conflicts are defined in section 2. All the violence measures are transformed into logarithm forms.[11] The parameter β_1 is the targeted parameter. Covariates $\mathbf{X}_i$ contain the important family background and childhood conditions. The regional controls, $\mathbf{Z}_g$, include the population sizes of urban and cadre people in the prefecture at the start of the CR. In addition, to control for common trends, we also consider the year-of-birth fixed effects λ_c. It is possible that individuals within each prefecture correlate with each other. Therefore, we cluster the idiosyncratic errors, ε_{igc}, at both the prefecture and the year-of-birth levels. The detailed estimation of two-way clustering of standard errors is from Cameron & Miller (2015).

Notably, a threat on using historical archives deserves more discussion. Because the CR is a political event, there is a possibility of underreporting of the counts of victims or other adverse outcomes. Walder (2014) argues, however, that the administration rules of local annals mitigate this problem because administrators of annals are required to write down the number zero if no violence is observed to have occurred. In addition, they must also record a reason for reporting a zero. Nevertheless, the underreporting problem is still possible. That is, if the levels of resources and attention allocated to writing the annals are low in some areas (e.g., in many rural areas), then the underreporting problem is more likely to occur. To examine whether the estimated effects are subject to potential underreporting

[11] The transformation is log(1+violence) to avoid infinite negative values.

bias, we also conduct a robustness check by including a proxy for underreporting behaviour in our regression model.

The proxy for the measure of potential underreporting is the lengths of the annals from which the measure of CR violence is constructed. Specifically, the proxy is the word counts in *dashiji*, the general chronology of events for the period from June 1966 to December 1971. The rationale to use the word counts is that if a county gave limited attention to writing the annals, then the underreporting behaviours can be simultaneously reflected in the lengths of the annals. All the violence counts are rescaled to 1,000,000 words. The robustness check results appear in the online appendix which is available on request.

3.5.2 Red Guard Army

The last identification strategy is also employed to find the within-CR cohort effects of the CR. The corresponding measure is the dummy indicating whether an individual joined the Red Guard army or not. The model to be estimated is as follows:

$$y_{igc} = \alpha + \beta_1 RedGuard_i + \mathbf{X}_i\gamma + \mathbf{Z}_g\eta + \lambda_c + \theta_g + \varepsilon_{igc}, \tag{3.2}$$

where, *ceteris parabus*, $RedGuard_i$ is the Red Guard dummy, which is measured at the individual level. The existing literature (see, e.g., Li et al., 2010) that employs the sent-down movement reveals a positive value of β_1. According to the history files (e.g., Bonnin, 2006), those who were sent down between 1969 and 1978 were also more likely to have joined the Red Guard army, and not *vice versa*.[12]

3.6 Results

In this section, we provide the estimation results for the models described in section 3.5. We first demonstrate the results using regional violence as the measure of exposure, and then we show the gender-specific estimates. Finally, we show the results with the Red Guard army dummy.

[12] In CHARLS, we identify individuals who were sent-down to the countryside; fewer than 300 respondents were identified. Therefore, most of the Red Guard army members in our sample stayed where they lived.

3.6.1 Regional Violence

The main results for model (1) appear in Table 3.6. Columns (1-4) show the estimates using the average lifetime income per worked year as the dependent variable, and Columns (5) and (6) show the results using the differential of the logarithm of first-last wages. We consider two measures for the severity of the CR violence. The first measure is the number of deaths that occurred during collective conflicts, and the second measure is the number of collective conflicts. Columns (1-2) and (5-6) are the estimates using the log of number of deaths, and Columns (3-4) and (7-8) are the estimates with the log of the number of conflicts.

For each pair of dependent variable measure and core explanatory variable measure, we first conduct the univariate regression (columns with odd numbers) and then include the full set of control variables and fixed effects (columns with even numbers). By comparing the two estimates in each pair, we can observe how sensitive the results are and whether selection effects dominate the true effects or not.

According to Columns (1-4), exposure to more severe violence decreases the average lifetime income by more than 10 percentage points without control variables and fixed effects. The effect is reduced to around 7 percentage points if we include the full set of controls and fixed effects. The estimate for the log of deaths in Column (2) also becomes insignificant, but the magnitude is very close to the log of number of conflicts in Column (4).

In terms of wage differentials, as shown in Columns (5-8), exposure to violence seems not to play a role. In addition to the insignificance of all estimates, the signs of the estimates between two measures of exposure are also divergent.

3.6.2 Heterogeneous Effects by Gender

Currently, Chinese households exhibit a large gender gap in income (see Table 3.2). Before the CR, the monetary system was run by the central state, and the gender gap problem was not yet reflected in terms of real income. It is therefore important to know whether the CR was the starting point of the gender gap. In other words, it is important to know whether males and females were affected by the CR to the same extent. In this subsection, we discuss the gender-specific effects by interacting the gender dummy with the violence terms, as shown in Table 3.7. The upper panel

contains the results using average lifetime income as the dependent variable, and the lower panel uses the wage differential. For each panel, Columns (1-2) and (5-6) show the estimates with the male sample, while other columns show the female counterpart.

Table 3.6. Lifetime Income on Regional Violence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Ln Avg. Lifetime Income				First-Last Wage Diff.			
log deaths	-0.102**	-0.0735			-0.0638	-0.0575		
	(0.0482)	(0.0450)			(0.102)	(0.0988)		
Log conflicts			-0.111***	-0.0720**			0.0879	0.0989
			(0.0384)	(0.0359)			(0.0603)	(0.0618)
Female		-0.696***		-0.695***		-0.847***		-0.849***
		(0.0461)		(0.0459)		(0.0583)		(0.0591)
Agricultural Hukou at birth		-0.466***		-0.457***		-0.00466		-0.0292
		(0.0722)		(0.0722)		(0.0904)		(0.0863)
Illiterate parent		-0.112**		-0.111**		0.0599		0.0628
		(0.0557)		(0.0558)		(0.0647)		(0.0642)
Constant	8.235***	8.531***	8.472***	8.672***	1.436***	1.288***	1.204***	1.048***
	(0.0618)	(0.0849)	(0.0956)	(0.115)	(0.0572)	(0.146)	(0.141)	(0.200)
Observations	8,808	8,808	8,808	8,808	8,808	8,808	8,808	8,808
Adjusted R-square	0.002	0.159	0.004	0.160	0.000	0.055	0.001	0.056
Year-of-birth FE	NO	YES	NO	YES	NO	YES	NO	YES
Regional controls	NO	YES	NO	YES	NO	YES	NO	YES
Childhood controls	NO	YES	NO	YES	NO	YES	NO	YES
Background controls	NO	YES	NO	YES	NO	YES	NO	YES

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. 2. Columns (1-4) show the results using the log of lifetime income per worked year as the dependent variable, and Columns (5-8) show the ones with first-last differential of log wage. 3. All standard errors are clustered at both the year-of-birth and prefecture levels. 4. The childhood control variables include whether the respondents' family experienced starvation or fled during the great famine when they were young. Background controls include the family financial situation and whether the respondents have received help when they were young.

In general, we find that the males are affected negatively in terms of average lifetime income while females are not. With regard to the wage differential, neither group seems to be affected. Similar to the results with the pooled sample, a one percentage point increase in deaths due to the CR reduces the males' lifetime income by 8.9 percentage points (see Column (2)). The effect is 7.77 percentage points (see Column (6)) when using the number of conflicts as the measure of severity. The results for the females become insignificant when we include the controls and year-of-birth fixed effects.

Table 3.7. Lifetime Income on Regional Violence: by Gender

Variables	(1) Male	(2) Male	(3) Female	(4) Female	(5) Male	(6) Male	(7) Female	(8) Female
				Avg. Lifetime Income				
log deaths	-0.110**	-0.0891**	-0.0911	-0.0546				
	(0.0437)	(0.0404)	(0.0586)	(0.0559)				
log conflicts					-0.100**	-0.0777*	-0.116***	-0.0607
					(0.0443)	(0.0441)	(0.0426)	(0.0409)
Agricultural Hukou at birth		-0.364***		-0.580***		-0.358***		-0.570***
		(0.0849)		(0.0745)		(0.0853)		(0.0742)
Illiterate parent		-0.139*		-0.0867		-0.138*		-0.0876
		(0.0760)		(0.0663)		(0.0765)		(0.0666)
Constant	8.589***	8.444***	7.877***	7.936***	8.797***	8.595***	8.129***	8.055***
	(0.0745)	(0.106)	(0.0592)	(0.112)	(0.115)	(0.144)	(0.109)	(0.134)
Observations	4,432	4,432	4,376	4,376	4,432	4,432	4,376	4,376
Adjusted R-squared	0.002	0.100	0.001	0.075	0.003	0.100	0.004	0.076
				First-Last Wage Diff.				
log deaths	-0.0893	-0.0865	-0.0358	-0.0247				
	(0.110)	(0.106)	(0.102)	(0.103)				
log conflicts					0.114	0.117	0.0681	0.0867
					(0.0748)	(0.0731)	(0.0711)	(0.0737)
Agricultural Hukou at birth		0.0471		-0.0702		0.0254		-0.0959
		(0.113)		(0.113)		(0.110)		(0.109)
Illiterate parent		0.0487		0.0572		0.0482		0.0619
		(0.0918)		(0.0976)		(0.0908)		(0.0972)
Constant	1.863***	1.168***	1.002***	0.600***	1.561***	0.874***	0.826***	0.400*
	(0.0679)	(0.187)	(0.0707)	(0.181)	(0.165)	(0.244)	(0.174)	(0.237)
Observations	4,432	4,432	4,376	4,376	4,432	4,432	4,376	4,376
Adjusted R-squared	0.000	0.015	-0.000	0.009	0.001	0.016	0.000	0.010
Year of Birth FE	No	Yes	No	Yes	No	Yes	No	Yes
Regional Controls	No	Yes	No	Yes	No	Yes	No	Yes
Childhood Controls	No	Yes	No	Yes	No	Yes	No	Yes
Background Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. 2. All standard errors are clustered at both the year-of-birth and prefecture levels. 3. The upper panel shows the results using log of lifetime income per worked year as the dependent variable, and the lower panel shows the results with first-last differential of log wage. 4. The childhood control variables include whether the respondents' family experienced starvation or fled during the great famine when they were young. Background controls include the family financial situation and whether the respondents received help when they were young.

3.6.3 Red Guard Army

Finally, we show the estimates of equation (2) in Table 3.8–namely, the estimates using the Red Guard army dummy as the core explanatory variable. Columns (1-4) are the estimates taking the log of average lifetime income per worked year as the dependent variable on the male and female samples, respectively. Columns (5-8) are the counterpart estimates using the first-last wage differential as the dependent

variable. Columns (1-2) and (5-6) show the results for the male sample, while the rest show the results for the female sample. After we add controls and fixed effects, the

Table 3.8. Lifetime Income on Red Guard Army

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Ln Avg Lifetime Income				First-Last Wage Diff.			
	Male	Male	Female	Female	Male	Male	Female	Female
Red Guard	0.0734	0.166***	0.114**	0.0818*	-0.0366	0.0349	0.0378	0.0305
	(0.0517)	(0.0366)	(0.0483)	(0.0468)	(0.0629)	(0.0651)	(0.0666)	(0.0753)
Agricultural Hukou at birth		-0.226***		-0.483***		0.0703		-0.202*
		(0.0737)		(0.0909)		(0.116)		(0.118)
Illiterate parent		-0.126*		-0.0768		-0.0352		0.0384
		(0.0706)		(0.0580)		(0.102)		(0.0927)
Constant	8.528***	8.672***	7.822***	8.532***	1.850***	1.893***	0.982***	0.973
	(0.0844)	(0.305)	(0.0641)	(0.377)	(0.0777)	(0.553)	(0.0672)	(0.619)
Observations	4,437	4,432	4,376	4,376	4,437	4,432	4,376	4,376
Adjusted R-square	0.001	0.207	0.001	0.174	-0.000	0.113	-0.000	0.121
Year-of-birth FE	No	Yes	No	Yes	No	Yes	No	Yes
Prefecture FE	No	Yes	No	Yes	No	Yes	No	Yes
Childhood controls	No	Yes	No	Yes	No	Yes	No	Yes
Background controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: 1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. 2. Columns (1-2) and (5-6) report on the male sample only, and Columns (3-4) and (7-8) report on the female sample only. 3. All standard errors are clustered at both the year-of-birth and prefecture levels. 4. The childhood control variables include whether the respondents' family experienced starvation or fled during the great famine when they were young. Background controls include the family financial situation and whether the respondents received help when they were young.

effect of having been a Red Guard army member becomes higher and significant for the males (see Columns (1) and (2)), while the opposite is for females (see Columns (3) and (4)). On average, the experience of being a Red Guard army member for males increases average lifetime income by about 16.6 percentage points, which is substantial. The effect is approximately 8.18 percentage points for females. The comparison between the univariate model and the full model also implies potential selection effects. Because the control variables we include are mainly related to family background during childhood (e.g., whether family members experienced starvation, whether the family's financial situation is better than that of neighbours), it might suggest that those with a good family background were more likely to join the Red Guard army. From the measures of the CR, we find that the regional violence measure is more representative for victims during the CR. On the contrary, members of the Red Guard army members were mostly the "suppliers" of the violence during the CR. Therefore, we assume that the Red Guard army represents a self-selected, more active, and younger population with better family backgrounds.

These individuals were more likely to self-motivate and seize (political/leadership) opportunities. Another possible explanation is that they have built connections with each other and that those connections helped them gain political power after the CR, which subsequently helped them earn more because being a public servant is often a well-paid occupation compared with others.

We have also estimated a separate specification, which allows for interaction between the CR violence measures and the Red Guard Army dummy. The interaction term is insignificant, regardless of whichever CR measure we use.

3.7 Conclusion

In this chapter, we study the long-term effects of the CR on individual lifetime income using the life history survey of the CHARLS data set. The extant literature reveals mixed evidence on the impact of the CR on income because different research sheds light on different aspects of the CR. Yet the effects of the violence during the CR on lifetime income need further investigation using alternative measures of exposure to the CR. For the measure of income, most research employs cross-sectional wages as the dependent variable, which has two potential drawbacks. First, cross-sectional data cannot smooth out the temporary fluctuations. To overcome this problem, we construct the average lifetime earnings per worked year as the first dependent variable. Second, cross-sectional data sets cannot capture the curvature of the individuals' lifetime income profile. The curvature is important because for those who have similar average lifetime income, it could distinguish those who have always earned a lot from those who started with low income but ended with high income. To account for this, we introduce the first-last wage differential, that is, the difference between an individual's first and last wages before retirement (or the latest wage for those who are still working).

We employ two identification strategies through two measures of exposure to the CR. First, we use regional violence during the CR between 1966 and 1971, and second we use whether individuals joined the Red Guard army as the measure of exposure. With the regional violence measure, we use the number of conflicts and the number of deaths at the prefecture level as a proxy for CR-induced violence. The results show that having been exposed to more violence decreases males' lifetime income by approximately 7%. Finally, we estimate the impact of having joined the Red Guard army on lifetime income. The results show that males in the CR cohort

who joined the Red Guard army have higher average lifetime income than those who did not join, and the effect is around 16.6 percentage points. The estimates with the first-last wage differential are insignificant for all gender groups. In future work, it will also be important to identify potential mechanisms of the long term effects and shed more light on the selection of being a Red Guard army member.

In the literature, several potential mechanisms can explain the results in this chapter. The first mechanism is an interrupted education (see, e.g., Meng & Gregory, 2002). During the CR, many students were not able to receive regular training, which translated to lower lifetime income after the CR. The second is social trust (see, e.g., Bai & Wu, 2020). The CR undermined mutual trust among individuals, which harmed various economic cooperation after the CR. The third is the disadvantaged local economies and governance. Red Guard Army members were more likely to become regional governors because of their active involvement. Compared with governors from other cohorts, those Red Guards are mostly inferior in terms of their governing abilities.

3.A Missing (Wrong) Values and Imputation Strategies

In the original data set, there are nontrivial number of missing values in either start or end dates of each job tenure; thus, we impute them using the following strategy. If the start dates are missing, we impute them using the end date of the previous job tenure. If the end dates are the missing dates, we impute them using the start date of the following job tenure. For those who were retired, we impute the missing end dates of the final job using the retirement date; otherwise, we impute them using the current date for the latest job. The detailed imputation numbers appear in Table 3.A.1.

We construct the wage paths as follows: first, we assign the individuals' wage of the start year, the end year, and the middle year (if the job lasted at least 20 years) to corresponding years of each individual-year observation. Second, we add or update the wage information from the other three waves at the years of interviews if available.[13] Then, we interpolate the wage paths using linear interpolation except during the unemployment and other career breaks.

Because there are still missing values, we further impute these by using the age-max-min imputation. Specifically, we record the minimum, mean, and maximum of the wages, and then we impute the missing values using minimum if the age was under 30 years of age, using the mean if the age was between 31 and 50 years of age, and using the maximum if the age was above 50 years of age. By using this method, we assume that during each employment tenure, except the linear wage growths, wages increase as work experience increases. Using this procedure, we further imputes 127,292 observations for the wage paths with all job categories.

[13] In wave 1, we drop those observations with extreme values in annual/monthly wages which is for example larger than 1,000,000.

Table 3.A.1. Missing Patterns on Job Variables

	Job Tenures			Job Category
i-th Job	Starting Date	Ending Date	Difference	
1	831	3901 (932) [2138]	3070[4]{0}	280
2	4269 (1462)[1]	11506 (1358)[2] [7343][3]	7237{-2}	3917
3	12951 (2843)	15819 (303) [5410]	2868{-2}[5]	12892
4	17372 (1861)	18556 (118) [2928]	1184{-1}	17360
5	19177 (741)	19677 (46) [1195]	500{0}	19183
6	19939 (309)	20190 (27) [533]	251{0}	19938
7	20285 (122)	20401 (8) [230]	116{0}	20286
8	20447 (54)	20515 (4) [118]	68{0}	20448
9	20540 (29)	20569 (5) [53]	29{0}	20540
10	20576 (12)	20589 (0) [25]	13{0}	20577
11	20598 (9)	20605 (0) [16]	7{0}	20599
12	20609 (4)	20611 (0) [6]	2{0}	20609
13	20611 (0)	20616 (1) [4]	5{0}	20611
14	20616 (1)	20620 (0) [5]	4{0}	20616
15	20621 (1)	20621 (0) [1]	0{0}	20621
16	20621 (0)	20621 (0) [0]	0{0}	20621
17	20621 (0)	20621 (0) [0]	0{0}	20621
18	20621 (0)	20621 (0) [0]	0{0}	20621

Note: 1. The table shows the numbers of missing values in the starting and ending dates of each job tenure and the job category variable. The total number of observations is 20,622, and numbers without any parentheses represent the number of missing values in each variable. In the brackets () of the start dates are the missing start dates imputed using the end date of the previous job tenure. 2. In the parentheses () of the end dates are the missing end dates imputed using the start date of the following job tenure. 3. In the square brackets [] are the missing end dates imputed using retirement date or the current date for the latest job, depending on retirement status. 4. The column "Difference" represents the difference of the missing values between the start and the end dates (positive number means the number of missing ending dates is larger, and *vice versa*). 5. The numbers in braces {} represent the difference between the missing start and end dates after imputation (a negative number indicates that the number of missing start dates is larger, and *vice versa*). After imputation, the job tenure information is of good quality.

Chapter 4

Exposure *in utero* to Adverse Events and Health Late-in-life*

4.1 Introduction

A growing body of literature has shown that exposure to adverse conditions around birth can have long-term negative consequences on health. The long shadow of early-life circumstances can be explained by both biological and social mechanisms. According to the "Fetal Origin Hypothesis" by Barker (see, e.g., Barker et al., 1993; Barker, 1995; Almond & Currie, 2011), malnutrition during pregnancy might cause disproportional fetal growth and program later coronary heart disease and type 2 diabetes (see, e.g., Lumey, Stein & Susser, 2011; Portrait et al., 2011; Xu et al., 2017; Kim et al., 2017; Hu et al., 2017; Dinkelman, 2017; Doblhammer, van den Berg & Lumey, 2013; Van den Berg et al., 2015). Moreover, prenatal stress might also be responsible for developing diseases such as depression and cognitive deficiency in the longer run (see, e.g., Akbulut-Yuksel, 2017; Gade & Wenger, 2011; Strauss et al., 2011; Bramsen et al., 2007; Grimard & Laszlo, 2014; Teerawichitchainan & Korinek, 2012; Islam et al., 2017). The pathway framework (Kuh & Shlomo, 2004), on the other hand, rests less on the biological imprinting and states that the impact of early life conditions on health later in life depends also on the interaction of the individual with the environment. The idea is that adverse conditions around birth might set in motion lifetime trajectories of health related disadvantages. For example, children

*This chapter is based on Wang et al. (2020).

born under adverse circumstances might be exposed to poor diets, smoking, worse educational and job opportunities, and all these factors in turn will have a negative impact on health.

In this chapter we estimate the effect of *in utero* exposure to the Chinese Great Famine (1959-1961) and the Cultural Revolution (mainly the Red Guard Army movement 1966-1971) on a range of physical and mental health outcomes in later life. While the famine was responsible for severe malnutrition for pregnant mothers living in rural areas, the Cultural Revolution caused a sharp increase in violence and stress in urban areas. Therefore, it is interesting to look at both events.

We draw data from the China Health and Retirement Longitudinal Study (CHARLS) survey, a nationally representative survey of the Chinese population aged 45 or more. The 2011 and 2015 waves of the survey include data from blood biomarkers, which allow us to construct an 8-year risk score for diabetes and a 10-year risk score for cardiovascular disease (CVD). Indeed some *in utero* effects might only manifest themselves later in life but early signals of disease might be detected at younger ages using blood-based biomarkers. For mental health, the data include standard survey scales to measure the presence of depressive symptoms and assess cognitive abilities.

Our results support the idea that adverse events around birth potentially have a long-lasting negative effect on both physical and mental health. We also find that there are gender differences in this effect: exposure *in utero* to the famine heightens the risk of type 2 diabetes for women, while it increases depressive symptoms for men. On the other hand, the Cultural Revolution mainly harmed cognitive abilities of men.

The rest of the chapter is organized as follows: section 4.2 summarizes the epidemiological and biomedical evidence. Section 4.3 briefly explains the background of the Chinese Great Famine and the Cultural Revolution. Section 4.4 introduces the data the construction details of the main variables. Section 4.5 describes the empirical model and section 4.6 presents the results. Finally, section 4.7 concludes the chapter.

4.2 Biological Mechanisms

In this section we summarize several commonly accepted biological mechanisms through which the *in utero* exposure to adverse events can be translated into late life diseases. Based on the literature, there are two important mechanisms: malnutrition and prenatal stress. Biomedical studies show that exposure to prenatal stress harms individuals' cortex development and it further induces depression and cognition deficiency, whereas exposure to malnutrition is mostly associated with later life cardiovascular disease or diabetes.

4.2.1 Stress and Cognitive Deficiency

In recent studies, prenatal stress is found to have a significant impact on the babies *in utero*. The hypothalamic-pituitary-adrenal (HPA) axis plays a crucial role in understanding the mechanism,[1] especially during early fetal development because plasticity is very high (see, Braithwaite et al., 2014). Basically, the HPA axis is a neuroendocrine system which regulates the body's response to stress. Higher prenatal stress implies more responses or reactions from the HPA axis, which in turn affects the brain formation of babies.[2]

The detailed processes are as follows (see, e.g., Molenaar et al., 2019; Bock et al., 2015). When pregnant mothers experience unexpected adverse events or a stressful environment, the hypothalamus (the H of HPA) in the brain area release excessively a type of hormone called corticotropin-releasing hormone (CRH). The CRH then travels to the pituitary gland (the P of HPA) which responds by releasing more adrenocorticotropic hormone (ACTH). The ACTH then travels across human bodies and reaches the adrenal glands (the A of HPA) on the top of kidneys. The adrenal glands are stimulated by the increase of ACTH and release more cortisol into the bloodstream. Normally the HPA axis has a feedback system that balance the hormone levels. When the stress level becomes too high, however, the "negative feedback" hits the HPA axis and the sense receptor in hippocampus will shut down the stress response mechanism. Finally, the cortisol level becomes out of control.

The elevated cortisol level then affects the foetus. Normally the placenta acts as

[1] There are also other mediation pathways to explain the association, but the one with HPA axis is commonly accepted.

[2] Various biological literature tries to find biomarkers to quantify the activity of the HPA axis and the prenatal stress, such as hair cortisone levels (see, e.g., Molenaar et al., 2019) and dehydroepiandrosterone (DHEA) and dehydroepiandrosteronesulfate (DHEAS) (see, e.g., Schmelter et al., 2019).

a barrier between the concentrated cortisol level and the foetus. However, when the cortisol density around the placenta becomes too high, it will penetrate the placenta and the foetus becomes exposed to this excessive cortisol. Recent epigenetic literature offers the mechanisms after the foetus exposing to the excessive cortisol. It has been shown that the gene expression and the DNA-methylation processes of babies are altered to adapt to the changing intrauterine environment. The methylation process of the genes such as *Nr3c1*, *CRH*, *CRHBP*, *Crb*,and *Nr3c1* F1 promoter (see, e.g., Welberg & Seckl, 2001; Hamada & Matthews, 2019; McGowan & Matthews, 2017) would be altered. For instance, the *Nr3c1* gene which controls the glucocorticoids receptor production and controls the HPA axis reactivity of babies. The altered HPA axis activities then affects the shaping of cortex and hippo-campus which reflects cognitive abilities late-in-life.

4.2.2 Malnutrition and Type-2 Diabetes and CVD Risks

Babies who experience malnutrition *in utero* are shown to have higher risks of type-2 diabetes (see, e.g., Vaiserman & Lushchak, 2019) and higher risks of CVD (see, e.g., Alessie et al., 2019). The literature finds that maternal malnutrition creates intrauterine growth restriction, which then produces foetal adipose tissue and pancreatic β-cell dysfunction for the foetuses. To be more specific, the β-cell dysfunction is essential to explain the mechanisms behind the association. The β-cell is made from the stem cells when the pancreas of the embryo is developing, and it is the basis for the islets of Langerhans (the home of hormone). The decreased β-cell in reaction to the environmental cues will lead to irreversibly reduced insulin secretion. That is why the babies exposed to malnutrition will develop higher diabetes and CVD risks.

The detailed processes are as follows: when pregnant mothers and their embryos experience malnutrition, the "thrifty genotype" starts to play a role (see, Hales & Barker, 1992). Epigenetic evidence (mostly from animal models) shows that, with reduced nutrition supply especially amino acids, the DNA-methylation processes for genes such as *IGF2*, *GNASAS*, *IL10*, *LEP*, *ABCA1*, *INSIGF* and *MEG3* will be altered (see, e.g., Lumey, Terry et al., 2011; Heijmans et al., 2008; Tobi et al., 2009). For example, one of the most famous epigenetic evidence by Heijmans et al. (2008) shows that, when the expression of the gene insulin-like growth factor-2 (IGF2) is reduced, the β-cell transformation from the stem cells is suppressed. Then the body becomes thriftier in keeping glucose and releasing less insulin, and the risk of having type 2 diabetes increase once the nutrition supply after birth is improved.

The babies finally have lower insulin secretion and a higher ability to store fat (see, e.g., Gluckman & Hanson, 2004).

The risks of having CVD also increase because of two reasons: first, since type 2 diabetes can also cause (or caused by) CVD, the risks for having CVD also increase. Second, the lack of amino acids lead to the decrease of β-cells, while the lack of glucose and oxygen will directly lead to a reprogramming of the neuro-endocrine system (see, e.g., Almond & Mazumder, 2011), similar to the HPA axis reaction due to stress. Then the CVD risks will be developed along with the type-2 diabetes risk.

The mechanisms summarized above are the most commonly accepted results. Some other research also finds that malnutrition could affect cognition (see, e.g., De Groot et al., 2011; Doblhammer, Van den Berg & Fritze, 2013), *vice versa* for the literature studying prenatal stress (see, e.g., Akbulut-Yuksel, 2017). These studies are mostly from the epidemiology literature and the findings might suggest that cognition deficiency is the complication of diabetes and CVD, *vice versa*.

4.2.3 Role of Gender

In the literature, the evidence on the sex-specific health effects of prenatal exposure to adverse events is mixed (Alessie et al., 2019). For example, in terms of exposure to prenatal stress, some studies reveal that it is mostly male offspring's emotionality to be negatively affected, while others find the opposite effect, i.e. that female offspring's anxiety and depressive symptoms respond significantly to the adverse environment (see, e.g., Bock et al., 2015, for a review of both animal and human studies). A strand of research finds that the gender-specific effects in response to stress can be explained by the timing of stress exposure (see, e.g., Mueller & Bale, 2008). Male babies are more responsive to prenatal stress during early gestational periods while female babies are more responsive during late gestational periods.

In terms of exposure to malnutrition, there is also some evidence showing that males are affected differently than females. For example, Eriksson et al. (2010) find that boys grow faster than girls *in utero*, therefore boys need more nutrition from mothers. If mothers experience malnutrition during pregnancy, boys *in utero* will more likely be affected. Moreover, boys *in utero* tend to develop their brains first. If exposed to undernutrition, boys' visceral development will be sacrificed to sustain the nutrition supply for their brains. In the long run, boys' kidney development might be hurt, and some diseases such as hypertension and CVD will incur due to

the underdevelopment of their kidney functions (see, e.g., Barker et al., 2006).

In this chapter, we will also empirically investigate the gender-specific health effects of exposure to the CR and the famine to see whether males and females are affected differently.

4.3 Background

4.3.1 The Chinese Great Famine

The Chinese Great Famine (CGF thereafter) is so far one of the largest famines in human history in terms of both the severity and the size of the affected population. The famine occurred mainly between 1959 and 1961, when the grain production substantially dropped. Given the fact that grain was the major source of food, the decline in grain production caused 16.5 to 45 million deaths during the famine period (see, e.g., Meng et al., 2015). In the year 1958, some regions had already experienced some grain production drops, but it had not yet become a nationwide disaster. After the famine, grain production increased monotonically until it reached the normal level.

The famine hit primarily the rural areas. Indeed, during the famine period, China was running a centrally planned economy, and rural households were only able to eat in local communal kitchens while urban households were able to consume normal food products within quotas depending on the household size. Moreover, rural households were not allowed to store food privately. So when the famine arrived, rural households were immediately hit by the food shortage shock in the communal kitchens. What was worse, the grain produced in rural areas was over-procured by the central government so it exacerbated the starvation of the rural households.

The famine not only produced substantial deaths nationwide but it also had a long-term impact on the health outcomes of survivors. With little food available for rural pregnant mothers, the foetuses in the famine cohort were more likely to suffer from malnutrition than any other cohorts. In addition, mothers could have also experienced stress which could have been passed on to the babies.

In Table 4.1, we show the annual province-level deaths rates provided by Lin & Yang (2000). From the table we observe that during the famine years death rates

are higher than in other non-famine years. Based on this, we calculate the excessive death rates and employ them to measure the famine severity across provinces. We calculate the excessive death rates in 0.1% unit by subtracting the average annual deaths rates between 1956 and 1958 from those in the famine years (1959-1961). During non-famine years, the measure of severity equals zero.

Table 4.1. Provincial Excessive Death Rates (unit 0.1%) between 1955 and 1964

Province / year	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964
Anhui	11.80	14.25	9.10	12.36	16.72	68.58	8.11	8.51	7.92	8.59
Fujian	8.26	8.43	9.02	7.50	7.95	15.61	12.18	8.27	7.51	8.68
Gansu	11.98	10.78	11.33	21.11	17.38	41.32	11.48	8.25	10.38	15.55
Guangdong	10.70	11.19	8.42	9.15	11.76	15.09	10.67	9.32	11.78	8.32
Guangxi	14.80	12.48	12.42	11.98	17.32	29.20	20.37	10.15	10.34	10.55
Guizhou	16.23	13.01	12.35	15.26	20.28	52.33	23.27	11.64	17.14	20.66
Hebei	11.86	11.26	11.59	10.92	12.31	12.19	13.34	8.97	10.66	10.48
Heilongjiang	11.33	10.08	10.45	9.17	12.76	10.52	11.12	8.61	8.56	11.47
Henan	11.75	14.00	11.80	12.69	14.10	39.65	10.10	8.03	9.39	10.65
Hubei	11.60	10.81	9.61	9.60	14.49	21.22	9.08	8.77	9.83	10.94
Hunan	16.41	11.50	10.35	11.58	12.92	29.26	17.48	10.23	10.26	12.88
Inner Mongolia	11.40	7.90	10.50	7.90	11.00	9.40	8.80	9.00	8.50	11.80
Jiangsu	11.65	12.81	10.03	9.33	14.55	18.41	13.35	10.36	9.03	10.13
Jiangxi	16.23	12.49	11.47	11.33	13.01	16.05	11.54	11.00	9.76	10.87
Jilin	9.91	7.53	9.05	9.11	13.43	10.13	11.12	9.96	9.44	12.62
Liaoning	9.40	6.60	9.40	8.80	11.80	11.50	17.50	8.50	7.90	9.30
Ningxia	N.A.	N.A.	N.A.	14.10	15.81	13.88	10.71	8.49	10.22	13.44
Qinghai	13.76	9.34	10.40	12.64	16.29	40.72	11.68	5.35	8.36	15.53
Shaanxi	10.55	9.85	10.31	11.04	12.76	12.27	8.76	9.35	10.55	15.60
Shandong	13.73	12.16	12.05	12.77	18.14	23.51	18.48	12.35	11.78	12.06
Shanxi	12.93	11.60	12.68	11.73	12.84	14.21	12.20	11.34	11.44	13.98
Sichuan	13.26	11.79	11.82	17.37	19.22	47.78	28.01	14.61	12.82	13.87
Yunnan	13.76	15.21	16.29	21.62	17.96	26.26	11.84	10.86	14.14	15.23
Zhejiang	12.58	9.46	9.32	9.15	10.81	11.88	9.84	8.61	7.89	9.21

Note: Date source comes from the National Bureau of Statistics of China and is summarized by Lin & Yang (2000).

4.3.2 The Cultural Revolution

The Cultural Revolution (CR thereafter) was a major adverse political event which happened between 1966 and 1976 (see, e.g., Walder & Su, 2003) in China. It is often called the "lost ten years" in China. In this period, there were millions of victims

who suffered from various types of political movements and prosecutions, and the negative impacts persist throughout the victims' whole life course. A more detailed description of the CR can be found in Bonnin (2006). The event itself can be divided into two sub-periods: the Red Guard Army movement between 1966 and 1971, and the rustication programme between 1971 and 1976. The Red Guard Army movement was accompanied by massive conflicts and associated victims, while the rustication movement produced very few conflicts and victims. The extant studies either employ the Red Guard Army movement to study the effect of violence (see, e.g., Wang, 2019) or employ the Rustication movement to study the impact of interrupted education (see, e.g., Meng & Gregory, 2002).

In this chapter, we focus on the Red Guard Army movement during the CR because we would like to focus on the effect of conflicts. During this period, students were encouraged by Mao to rebel against the central/local governments. Students worshipped him, followed his extreme-left ideology, and created a lot of violence in this movement. In response, the governments also repressed those student activities and produced plenty of victims and deaths. Those massive conflicts, victims, injuries and deaths were documented by local annals and gazetteers (see, e.g., Walder, 2014).

Pregnant mothers who experienced or witnessed the violence were more likely to suffer from prenatal stress (perhaps also malnutrition). This stress can be passed on to the foetuses. Mothers in different regions and in different years suffered from different levels of violence, and we make use of this geographical and temporal variation to identify the health effects.

More specifically, we employ the number of collective conflicts as the measure of violence intensity during the CR (Red Guard Army movement 1966-1971) years. The measure equals 0 during the non-CR years. In a sensitivity analysis, we also use an alternative measure based on the number of deaths due to these collective conflicts . A similar strategy is also employed by Bai (2014) and Wang (2019). In Figure 3.3, we show the severity of total violence during the CR years by prefecture.

4.4 Data on Health Outcomes and Descriptive Statistics

We draw data from the China Health and Retirement Longitudinal Study (CHARLS) 2011 and 2015 surveys. The CHARLS survey is designed to be nationally represent-

ative, and it focuses only on the population aged 45 and above. It has one national study with three regular waves (wave 2011, 2013 and 2015) and a retrospective wave (wave 2014). In this chapter, we use the 2011 and 2015 waves of the national sample because they are the only waves which include a biomarker section with blood tests. Biomarkers are powerful preclinical or pre-morbid signals for diseases such as diabetes or CVD and allow us to construct risk scores.

As the CHARLS data set does not provide suitable measures for the severity of the CGF and the CR, we merge it with two additional data sources. We use the province-level excessive death rates from Lin & Yang (2000) to measure the famine severity, and we use prefecture-level conflicts and deaths counts from Walder (2014) to measure the CR severity. We merge each individual in CHARLS with these two additional data sets using information on the prefecture/province and the year at birth. In CHARLS we also have information on the *hukou* status at birth, a household registration system used in China which identifies a person as resident in a rural or urban area.

4.4.1 Sample Selection

In the original data sets, there are 17,708 observations in the 2011 wave and 21,095 observations in the 2015 wave. Only 11,847 respondents in the 2011 wave and 13,420 respondents in the 2015 wave have participated in the blood-based biomarker sections. When we merge the different sections of the questionnaire, we are left with 10,138 observations in wave 2011 and 13,280 observations in wave 2013. We then construct a cross-sectional data set by pooling the two waves, and we use the latest information for each individual in the pooled data set. In other words, if a respondent has been interviewed in both 2011 and 2015, we use the information in 2015, which results in a sample size of 16,674.

Notably, we have dropped some observations, because not all respondents surveyed in CHARLS have participated in the blood-based biomarker sections. The response rates of the blood-based biomarker section are around 65%. The CHARLS team has estimated a Logit regression explaining the participation of the blood sample (see Zhao et al. (2013) and Chen et al. (2019) for details). We have replicated their results and introduced two core explanatory variables into their model: the CR violence measure and the famine severity measure. The estimates for these two variables are insignificant, so the participation does not correlate with the core explanatory variables. Therefore, there is no evidence of attrition bias in this

analysis.

We restrict our sample to those who were born between 1950 and 1971, which leaves us with 10,959 observations. We choose this period for the following reasons: first, the period not only covers both the CGF (1959-1961) and the Red Guard Army movement of the CR (1966-1971) but also includes a relatively peaceful time (1950-1959, 1962-1965). Second, we do not want to include war periods such as the Second World War (1938-1945) and the civil war (1945-1949).

After merging with the CR violence data and the CGF excessive death rates data, the sample size reduces to 9,794. We also drop observations with missing values in *hukou* status (144 observations), cognitive tests (595 respondents), depression symptoms (687) and the blood-based biomarkers (483).

In the final sample, there are 7,885 respondents distributed in 254 prefectures in 29 provinces.[3] In this sample, there are 4,341 female and 3,544 male respondents. Of those 7,885 respondents, 7,232 individuals have rural *hukou* at birth and 653 individuals have non-rural *hukou* at birth. There are 3,668 individuals who were born between 1950 and 1958, 799 individuals born between 1959 and 1961, 1,758 individuals born between 1961 and 1965, and 1,660 individuals born between 1966 and 1971.

4.4.2 Health Outcomes

Diabetes and CVD

Different from the extant studies which employ mostly self-reported health measures, we use biomarker data to construct risk scores for developing diabetes and CVD. For type-2 diabetes we consider the 8-year risk score, which clinically can be served as early monitor for diabetes prevalence after 8 years. We use various sources of information to construct the risk score from the Framingham Heart Study (FHS, see, Wilson et al., 2007): age, gender, BMI, blood pressure, blood High Density Level Cholesterol (HDL Cholesterol), Triglycerides, and fasting blood glucose. The detailed calculations can be found in appendix 4.A. The risk score works appropriately for undiagnosed patients but works inappropriately for those who have been diagnosed with diabetes and are taking medications. Therefore, we assign a risk score of 1 to those who were diagnosed with diabetes but have normal blood sugar

[3] Provinces of Xizang, Chongqing (province-level municipality) and Hainan are not included in CHARLS.

levels because they are taking medications, which we can identify from the CHARLS data. The distribution of the risk score can be found in Table 4.2. In our sample, on average, the risk of getting diabetes in 8 years is 32.5%, and 16.2% of the respondents have already been diagnosed with diabetes. The figures are comparable with other data sources or public reports. For example, Xu et al. (2013) report that nearly half of the Chinese adult population are prediabetic, and the 2015 Report on Nutrition and Chronic Disease in Chinese Residents reveals that the prevalence rate of diabetes among Chinese adults aged 40 is 9.7 percent (see, Burns & Liu, 2017, chapter 6). We also disaggregate the numbers by cohort, gender and *hukou* status, and the statistics can be found in Table 4.3. From Panel A of Table 4.3, we find that normally the younger cohorts have lower diabetes risks except for the famine cohort (1959-1961). The eight-year diabetes risk for the famine cohort is 35.7%, which is higher than that of both the pre- and post-famine cohorts (34.1% and 31.9%),[4] indicating that the famine might play a role in increasing the risk of getting diabetes. Panel B of the table shows that females have lower diabetes risk but higher prevalence rates.

For CVD, we follow D'agostino et al. (2008) and construct a 10-year risk score from the FHS. Other studies employ the Systematic COronary Risk Evaluation (hereinafter SCORE) (see, e.g., Conroy et al., 2003; Alessie et al., 2019) but we focus on the FHS because Selvarajah et al. (2014) shows that it performs better than SCORE for the Asian population. We refer the reader to appendix 4.B for details on the construction of the risk score. As for diabetes, we assign value one to the CVD risk score for respondents with a history of CVD events such as heart failure and stroke. The descriptive statistics can be found in Table 4.2. On average, 8.27% of the sample has already experienced a CVD event (stroke and heart failures), and the average CVD risk in the sample is 15.4%. The figure is close to external sources. For example, the reported prevalence rate of CVD in China is about 20% (one in five adults in China has a CVD).[5] In Table 4.3 we also offer disaggregated statistics of CVD risks and prevalences by cohort and *hukou* status. As expected, younger cohorts have on average lower CVD risk and prevalence compared to older cohorts. Urban-born residents suffer from higher risks than rural-born residents and the risk is higher for males than females.

[4] The difference between famine and post-famine cohorts are statistically significant.

[5] The statistics are provided by World Heart Federation: https://www.world-heart-federation.org.

Table 4.2. Descriptive Statistics CHARLS 2011/15: Born 1950-1971

VARIABLES	(1) Mean	(2) S.D.	(3) Min	(4) Max	(5) Quartile 1	(6) Median	(7) Quartile 3
			Sample size 7,885				
Female (0-1)	0.551	0.497	0	1	0	1	1
Rural Hukou (0-1)	0.917	0.276	0	1	1	1	1
Cognition & Depression							
Episodic Memory (Word Recall) (0-10)	3.699	1.765	0	10	2.500	4	5
Mental Intactness (TICS) (0-10)	6.815	2.754	0	10	5	7	9
Graphical Cognition (0-1)	0.694	0.461	0	1	0	1	1
CESD-10 (0-30)	7.792	6.295	0	30	3	6	11
CVD & Diabetes							
Diabetes	0.162	0.368	0	1	0	0	0
8-year Diabetes Risk	0.325	0.341	3.51e-05	1	0.0478	0.166	0.551
CVD	0.0827	0.275	0	1	0	0	0
10-year CVD Risk	0.154	0.266	0.00220	1	0.0249	0.0572	0.128

Note: Sample size 7,885 include all the respondents in wave 2011 and 2015, and we keep the latest information if the repondents appear twice. Rural *hukou* equals one if the repondents have the rural *hukou* at birth, zero otherwise. Episodic memory is constructed by asking the respondents recall the ten words listed on cards, ranging from 0 (no words recalled) to 10 (all words recalled). Mental intactness is measured through questions which ask the respondent to subtract 7 from 100 and keep subtracting 7 from the result for a maximum of five times and to name the correct date (day of the week, month, day, year, and season). The maximum score is 10. Graphical cognitive ability is measured by a dummy which is equal to 1 if the respondent is able to draw a picture showed by the interviewer. CESD-10 measures depressive symptoms, ranging from 0 (no depression at all) to 30 (the most serious depression). Diabetes is a dummy which equals one if the respondents have been diagnosed with diabetes and/or taking medications to control the blood sugar level. The 8-year risk score for diabetes is constructed by following Wilson et al. (2007). CVD is a dummy for actual cardiovascular disease and equals one if the respondents have been diagnosed with stroke or heart failures. The Cardiovascular Disease 10-year risk measure is from Framingham Heart Study(D'agostino et al. (2008)).

Cognition and Depression

We employ various measures to describe cognitive abilities and depressive symptoms of older Chinese individuals. The first measure of cognition, *episodic memory*, uses the word recalling questions in CHARLS (see, e.g., Lei et al., 2012). Respondents were requested to read ten Chinese words, and then they were asked to recall as many words as possible at two different moments in time. The first is just after reading all the words, and the second is after answering several other questions. We measure the numbers of correct words by taking the average of the two counts. The second measure of cognition is *mental intactness*, which is generated from ten

Table 4.3. Descriptive Statistics of Health Outcomes by Groups in CHARLS 2011 and 2015

Panel A: By Cohort

	pre-famine 1950-1959 (N=3,668)		famine 1959-1961 (N=799)				pre-CR 1962-1965 (N=1,758)				CR cohort 1966-1971 (N=1,660)			
Health Variables	Mean	SD	Mean	SD	Diff.[2]	*p*-value[3]	Mean	SD	Diff.	*p*-value	Mean	SD	Diff.	*p*-value
Female	0.524	0.499	0.546	0.498	-0.022	0.2587	0.572	0.495	-0.026	0.2193	0.589	0.492	-0.017	0.3143
Rural	0.910	0.286	0.900	0.300	0.010	0.3748	0.922	0.269	-0.022*	0.0647	0.937	0.243	-0.015*	0.0878
Cognition & Depression														
Episodic Memory (Word Recall)	3.334	1.743	3.683	1.684	-0.349***	0.0001	4.011	1.715	-0.328***	0.0001	4.183	1.728	-0.172***	0.0035
Mental Intactness (TICS)	6.417	2.880	6.780	2.662	-0.363***	0.0011	7.011	2.592	-0.231**	0.0384	7.502	2.514	-0.491***	0.0001
Graphical Cognition	0.622	0.485	0.701	0.458	-0.079***	0.0001	0.763	0.425	-0.062***	0.0009	0.779	0.415	-0.016	0.2659
CESD-10	8.282	6.470	7.458	6.073	0.824***	0.0015	7.659	6.190	-0.201	0.4440	7.013	6.019	0.646***	0.0020
CVD & Diabetes														
Diabetes	0.187	0.390	0.184	0.388	0.000	0.8437	0.141	0.348	0.043***	0.0053	0.117	0.321	0.024**	0.0365
8-year Diabetes Risk	0.341	0.346	0.357	0.357	-0.016	0.2390	0.319	0.332	0.038***	0.0089	0.282	0.324	0.037***	0.0010
CVD	0.110	0.313	0.100	0.300	0.010	0.4098	0.0666	0.249	0.033***	0.0033	0.0301	0.171	0.037***	0.0001
10-year CVD Risk	0.207	0.294	0.170	0.286	0.037***	0.0012	0.119	0.241	-0.05***	0.0033	0.0674	0.169	0.0516***	0.0001

Note: 1. In this table we disaggregate the statistics from Table 4.2 by cohort, gender and *hukou* status. Panel A exhibits the statistics by cohort, and Panel B exhibits the ones by gender and *hukou*. Total sample size: 7,885. 2. We compare the mean differences between two consecutive cohorts. 3. The *p*-values are calculated based on a two-tailed test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4.3 Continued

Panel B: By Hukou and Gender

	Male (N=3,544)		Female (N=4,341)				Rural (N=7,232)		Non-Rural (N=653)			
Health Variables	Mean	SD	Mean	SD	Diff.(Male-Female)	*p*-value	Mean	SD	Mean	SD	Diff(Rur-NonRur).	*p*-value
Female	-	-	-	-	-	-	0.550	0.497	0.551	0.498	-0.001	0.9607
Rural	0.917	0.275	0.917	0.276	0.000	1.0000	-	-	-	-	-	-
Cognition & Depression												
Episodic Memory (Word Recall)	3.735	1.694	3.670	1.821	0.065	0.1038	3.608	1.750	4.715	1.611	-1.107***	0.0001
Mental Intactness (TICS)	7.477	2.417	6.274	2.891	1.203***	0.0001	6.706	2.777	8.015	2.149	-1.309***	0.0001
Graphical Cognition	0.793	0.405	0.613	0.487	0.180***	0.0001	0.680	0.466	0.850	0.357	-0.170***	0.0001
CESD-10	6.620	5.750	8.750	6.554	-2.130***	0.0001	7.924	6.336	6.332	5.618	1.5920***	0.0001
CVD & Diabetes												
Diabetes	0.166	0.373	0.158	0.364	0.008	0.3371	0.157	0.364	0.213	0.410	-0.056***	0.0002
8-year Diabetes Risk	0.321	0.347	0.328	0.335	-0.007	0.3638	0.319	0.339	0.395	0.356	-0.076***	0.0001
CVD	0.0643	0.245	0.0977	0.297	-0.033***	0.0001	0.0784	0.269	0.130	0.337	-0.052***	0.0001
10-year CVD Risk	0.182	0.233	0.132	0.288	0.050***	0.0001	0.150	0.260	0.206	0.319	-0.056***	0.0001

Note: 1. In this table we disaggregate the statistics from Table 4.2 by cohort, gender and *hukou* status. Panel A exhibits the statistics by cohort, and Panel B exhibits the ones by gender and *hukou*. Total sample size: 7,885. 2. We compare the mean differences between two consecutive cohorts. 3. The *p*-values are calculated based on a two-tailed test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

questions of the Telephone Interview of Cognitive Status (TICS hereinafter), (see, e.g., Lei et al., 2012; Huang & Zhou, 2013). This measure is constructed based on whether the respondent is able to subtract 7 from 100 and keep subtracting 7 from the result for a maximum of five times, and to name the correct date (day of the week, month, day, year, and season). The maximum score is ten. The third measure considers graphical cognitive ability, which is a dummy indicating whether the respondent is able to draw a picture showed by the interviewer.

For depressive symptoms, we employ the well-known CESD-10 measure (see, e.g., Björgvinsson et al., 2013). Respondents were asked to report how often they had experienced the following situations: being bothered by small things, having difficulty in keeping their mind on what they were doing, feeling depressed, being tired of doing things, feeling hopeful, feeling fearful, feeling restless, feeling happy, feeling lonely, and feeling hard to get going. The answers were recorded on a four-point scale from 0 to 3, corresponding to rarely, some days, occasionally, and most of the time for the negative questions, and reversed for the positive ones. The total score ranges between 0 and 30, where a cutoff score of 10 or higher indicates clinically relevant depression.

In Panel A of Table 4.3, we observe that older cohorts have on average lower cognitive abilities (episodic memory, mental intactness and graphical cognition), whereas depressive symptoms are more severe among younger respondents. Panel B shows that females have lower cognitive performance than males, and they have higher levels of depressive symptoms. Rural-born residents have more depressive symptoms and less cognitive abilities than their urban counterparts.

4.5 Empirical Framework

The model to be estimated is of the following form:

$$y_{igpt} = \alpha + \beta_1 \cdot EDR_{pt} \cdot Rural_{igp} + \beta_2 \cdot Vio_{gt} \cdot Urban_{igp} + \mathbf{X}_{igp}\gamma + \lambda_t + \theta_g + \epsilon_{igpt}, \tag{4.1}$$

where y_{igpt} is one of the health outcomes we have mentioned in section 4.4. The subscript i refers to the individual, g denotes prefecture (254), p means province (29), and t stands for year of birth. The dummy $Rural_{igp}$ equals one if the individual was born in a rural area (having a rural *hukou* at birth) and zero if the individual was

born in an urban area (having an urban *hukou* at birth). The opposite holds for the dummy $Urban_{igp}$. As we only consider their *hukou* status at birth, these dummies are time-invariant. The variable excess death rate EDR_{pt} is the excess mortality rate in province p in the year in which the respondent was born t for the CGF years (1959-1961). It is computed by subtracting the average death rate between 1956 and 1958 from the annual death rate during the famine years. During the non-famine years, EDR_{pt} equals zero. The variable violence Vio_{gt} is measured at the level of the prefecture g in year t during the CR and is the number of conflicts per thousand individuals during the CR years. It is equal to zero in the non-CR years. Since the famine mainly hit rural areas and the CR mainly affected urban areas, we interact EDR_{pt} with $Rural_{igp}$ and interact Vio_{gt} with $Urban_{igp}$.

The covariate vector $\mathbf{X}_{igp}$ includes variables like gender and *hukou* status. The cross-cohort correlations are captured by the cohort fixed effect λ_t, and we also introduce the prefecture fixed effects θ_g to account for unobserved cross-regional heterogeneity. The rest of the unobserved shocks are included in the idiosyncratic error term ϵ_{igpt}. To account for the correlation within each prefecture and each cohort, the standard errors are clustered at both the prefecture and the year of birth level.

4.6 Results

4.6.1 Main Results

We estimate model (1) for the full sample, which consists of all individuals who were born between 1950 and 1971, and the results can be found in Table 4.4. Column (1) exhibits the regression estimates using the predicted 8-year type-2 diabetes risk as the dependent variable. Column (2) shows the estimates using the predicted 10-year CVD risk. Column (3) displays the estimates using depression as the dependent variable which is measured by the CESD-10 scale. Columns (4-6) are the estimates for three cognition measures. The results show that mothers' exposure to a 1% (the unit of the excess deaths rates in the raw data is 0.1%) more severe famine in rural areas during pregnancy would increase the risk of type-2 diabetes of the offspring by 0.03 percentage points (Column (1)) on average, while the effect on CVD risk is insignificant. These results suggest that exposure to famine might have triggered the "thrifty genotype" mechanism which induces insulin secretion during adult

Table 4.4. Main Results: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression	Cognition		
VARIABLES	Predicted 8-year Risk	Predicted 10-year Risk	CESD-10	Graphical Cognition	Mental Intactness (TICS)	Episodic Memory (Word Recall)
	%	%	Score 0-30(Worst)	0-1(Best)	0-10(Best)	0-10(Best)
EDR*Rural	0.00354**	0.000783	0.0597**	-0.000878	-0.0112	-0.00961
	(0.00170)	(0.00122)	(0.0291)	(0.00220)	(0.0123)	(0.00805)
Conflicts*Urban	-0.0108	-0.0358	-0.0410	-0.0644	0.369	0.0276
	(0.0372)	(0.0253)	(0.677)	(0.0437)	(0.255)	(0.210)
Rural Hukou	-0.0325	-0.0185	1.292***	-0.138***	-0.815***	-0.862***
	(0.0247)	(0.0157)	(0.353)	(0.0269)	(0.129)	(0.122)
Female	0.00872	-0.0443***	2.204***	-0.193***	-1.274***	-0.134**
	(0.0120)	(0.00781)	(0.158)	(0.0162)	(0.111)	(0.0541)
Constant	0.484**	0.216	0.808	1.053***	10.45***	5.689***
	(0.193)	(0.145)	(3.487)	(0.252)	(1.468)	(0.964)
Observations	7,885	7,885	7,885	7,885	7,885	7,885
R-squared	0.090	0.149	0.127	0.152	0.191	0.153
Interview Year Dummy	YES	YES	YES	YES	YES	YES
Year of Birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note:
1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses.
2. In Columns (1-2) we examine the impacts of violence (measured by log of number of collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk.
3. Column (3) shows the results using CESD-10 as the measure of depression, and Columns (4-6) use three measures of cognition: i) graphical cognition, ii) mental intactness, and iii) memory ability using word recall questions (see the chapter for details).
4. *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956-1958 from the annual death rates in the three famine years. The variable *Conflicts* measures the level of violence in each of the years between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

life. The magnitude of the effect is non-negligible because it is around 10.9% of the sample mean and around 10.4% of a standard deviation. This increase in the risk of developing diabetes might generate substantial healthcare costs and welfare losses for the society. A 1% increase in famine severity leads to a 0.6 points increase in the depression scale, while there is no significant effect on cognitive abilities. Exposure to conflicts *in utero* has no significant long-run effects on any of the health outcomes.

4.6.2 Heterogeneous Effects by Gender

Some studies, such as Eriksson et al. (2010), suggest that boys might be affected differently than girls (see section 4.2.3). To account for potential heterogeneous

effects between males and females, we interact the famine severity and regional conflicts with gender. The results can be found in Table 4.5. The results show that the

Table 4.5. Gender Specific: CHARLS 2011 and 2015

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression	Cognition		
VARIABLES	Predicted 8-year Risk	Predicted 10-year Risk	CESD-10	Graphical Cognition	Mental Intactness (TICS)	Episodic Memory (Word Recall)
	%	%	Score 0-30(Worst)	0-1(Best)	0-10(Best)	0-10(Best)
Male*EDR*Rural	0.000457	0.00138	0.0803**	0.00162	-0.0138	-0.00825
	(0.00222)	(0.00193)	(0.0402)	(0.00303)	(0.0172)	(0.0115)
Female*EDR*Rural	0.00566***	0.000376	0.0454	-0.00258	-0.00937	-0.0105
	(0.00210)	(0.00170)	(0.0348)	(0.00260)	(0.0147)	(0.00962)
Male*Conflicts*Urban	-0.00619	-0.0622	0.869	-0.167**	-0.421	-0.444
	(0.0531)	(0.0391)	(0.937)	(0.0764)	(0.395)	(0.283)
Female*Conflicts*Urban	-0.0138	-0.0189	-0.621	0.00115	0.871***	0.329
	(0.0438)	(0.0316)	(0.799)	(0.0547)	(0.319)	(0.272)
Rural Hukou	-0.0324	-0.0188	1.302***	-0.139***	-0.823***	-0.867***
	(0.0248)	(0.0158)	(0.356)	(0.0271)	(0.129)	(0.121)
Female	0.00638	-0.0442***	2.231***	-0.192***	-1.286***	-0.139**
	(0.0118)	(0.00810)	(0.163)	(0.0161)	(0.112)	(0.0542)
Constant	0.486**	0.216	0.783	1.052***	10.46***	5.692***
	(0.193)	(0.145)	(3.487)	(0.251)	(1.468)	(0.963)
Observations	7,885	7,885	7,885	7,885	7,885	7,885
R-squared	0.091	0.149	0.127	0.153	0.192	0.153
Interview Year Dummy	YES	YES	YES	YES	YES	YES
Year of Birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note:
1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses.
2. This table show the gender-specific estimates with the full sample.
3. In Columns (1-2) we examine the impacts of violence (measured by log of number of collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk.
3. Column (3) shows the results using CESD-10 as the measure of depression, and Columns (4-6) use three measures of cognition: i) graphical cognition, ii) mental intactness, and iii) memory ability using word recall questions (see the chapter for details).
4. *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956-1958 from the annual death rates in the three famine years. The variable *Conflicts* measures the level of violence in each of the years between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

effect of the famine on diabetes risk matters only for females, and a 1% increase in excessive death rate would imply a 0.06 percentage points higher risk of getting type-2 diabetes. The effect is around 17.5% of the sample mean and 16.7% of a standard deviation, which is statistically and economically relevant. The effects for males are statistically insignificant. The results might also be driven by selective mortality, namely males were more likely to die earlier. In terms of the impact on CVD risk, we also find insignificant results. The pattern for depression is different: a 1% increase

in famine-induced deaths increases males' depression score by 0.8, suggesting that boys' emotionality is more affected by the adverse prenatal environment, especially prenatal stress.

Two mechanisms could explain the results that significant effect appears only on females. The first is that, according to section 4.2.3, boys grow faster than girls do *in utero*. Therefore, boys need more nutrition during pregnancy than girls. When their mothers experience malnutrition during pregnancy, boys kidney functions will be underdeveloped, which triggers the onsite of various metabolic diseases (Eriksson et al. (2010)). The second is selective mortality. Since boys are more vulnerable to adverse events, they are more likely to die in utero (Di Renzo et al. (2007); Ingemarsson (2003)). Therefore, those males who survived in the sample are biologically stronger than their female peers.

4.6.3 Robustness Checks

We conduct two robustness checks to ensure that the effects are not spurious. First, since the number of conflicts might not capture the full damage during the CR, we use an alternative violence measure: the (log) number of deaths due to the collective conflicts. Second, since the results might be sensitive to the choice of the control group, in a sensitivity analysis we restrict the control group to those who were born after 1955 (the original sample include all respondents who were born after 1949).

In Table 4.6 we report the results when we replace the measure of CR violence with the log of number of deaths. Using an alternative measure does not substantially alter the effect of the famine on diabetes and CVD risk scores. The estimated effects of violence on cognition are less robust, although both results in Columns (4-6) of Table 4.4 and 4.5 show that cognition is negatively affected by violence exposure.

The results for the second robustness check are exhibited in Table 4.7. In this table, we restrict the sample to those who born after 1955. From the table we can find that the estimates do not qualitatively change from the ones in Table 4.5. For instance, the estimated coefficient in Column (1) for the female-EDR-rural interaction term changes from 0.00566 to 0.00558. The effect of the male-EDR-rural interaction term in Column (3) does not change at all. The coefficient on the male-conflict-urban interaction term changes from -0.167 to -0.128. Overall, the results are robust to changes in the size of the control group. To further examine how control group composition affects the sensitivity of our results, we estimate a specification

Table 4.6. Gender Specific: CHARLS 2011 and 2015 Deaths

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression		Cognition	
VARIABLES	Predicted 8-year Risk	Predicted 10-year Risk	CESD-10	Graphical Cognition	Mental Intactness (TICS)	Episodic Memory (Word Recall)
	%	%	Score 0-30(Worst)	0-1(Best)	0-10(Best)	0-10(Best)
Male*EDR*Rural	0.000443	0.00133	0.0807**	0.00163	-0.0135	-0.00790
	(0.00222)	(0.00194)	(0.0402)	(0.00303)	(0.0172)	(0.0115)
Female*EDR*Rural	0.00566***	0.000327	0.0458	-0.00262	-0.00934	-0.0104
	(0.00209)	(0.00169)	(0.0348)	(0.00260)	(0.0147)	(0.00962)
Male*Deaths*Urban	-0.891	-0.691	15.00	-2.562	-7.533	-7.455*
	(0.883)	(0.667)	(15.98)	(1.753)	(6.729)	(4.416)
Female*Deaths*Urban	0.390	0.880	-30.04*	-0.318	16.24**	-1.765
	(1.106)	(0.737)	(16.78)	(1.210)	(7.067)	(5.082)
Female	0.00601	-0.0442***	2.230***	-0.191***	-1.282***	-0.134**
	(0.0119)	(0.00808)	(0.163)	(0.0162)	(0.112)	(0.0544)
Rural Hukou	-0.0321	-0.0142	1.266***	-0.136***	-0.837***	-0.881***
	(0.0247)	(0.0157)	(0.352)	(0.0269)	(0.127)	(0.115)
Constant	0.486**	0.216	0.790	1.051***	10.45***	5.686***
	(0.193)	(0.145)	(3.486)	(0.251)	(1.468)	(0.963)
Observations	7,885	7,885	7,885	7,885	7,885	7,885
R-squared	0.091	0.149	0.128	0.153	0.192	0.153
Interview Year Dummy	YES	YES	YES	YES	YES	YES
Year of Birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note:
1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses.
2. This table shows the results using the log of number of deaths due to Cultural Revolution as the measure of violence.
3. In Columns (1-2) we examine the impacts of violence (measured by log of number of deaths due to collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk.
4. Column (3) shows the results using CESD-10 as the measure of depression, and Columns (4-6) use three measures of cognition: i) graphical cognition, ii) mental intactness, and iii) memory ability using word recall questions (see the chapter for details).
5. *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956-1958 from the annual death rates in the three famine years. The variable *Deaths* measures the number of deaths due to CR violence in each year between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

with only the Chinese Great Famine indicator on the subsample 1950-1965 and a specification with only the CR violence measure on the subsample 1960-1975. The coefficient estimates from these two specifications are very similar to the results in Table 4.5. The effects of exposure to famine on diabetes risk for rural females are always around 0.005, and the effects of exposure to the CR on the graphical cognitive ability for urban males are always significantly negative. Therefore, the results are quite robust.

Table 4.7. Robustness Check: Born after 1955

	(1)	(2)	(3)	(4)	(5)	(6)
	Diabetes	CVD	Depression		Cognition	
VARIABLES	Predicted 8-year Risk	Predicted 10-year Risk	CESD-10	Graphical Cognition	Mental Intactness (TICS)	Episodic Memory (Word Recall)
	%	%	Score 0-30(Worst)	0-1(Best)	0-10(Best)	0-10(Best)
Male*EDR*Rural	-0.000245	0.000773	0.0803**	0.000921	-0.0182	-0.00691
	(0.00223)	(0.00202)	(0.0401)	(0.00293)	(0.0168)	(0.0115)
Female*EDR*Rural	0.00558**	-0.00135	0.0429	-0.00414	-0.0204	-0.0136
	(0.00218)	(0.00147)	(0.0348)	(0.00271)	(0.0144)	(0.00976)
Male*Conflicts*Urban	-0.00514	-0.0385	0.815	-0.128*	-0.212	-0.367
	(0.0549)	(0.0365)	(0.932)	(0.0750)	(0.387)	(0.285)
Female*Conflicts*Urban	-0.0137	-0.00579	-0.705	0.0311	0.942***	0.435*
	(0.0449)	(0.0296)	(0.831)	(0.0541)	(0.314)	(0.253)
Female	-0.0133	-0.0270***	2.221***	-0.166***	-1.118***	-0.0744
	(0.0149)	(0.00696)	(0.170)	(0.0170)	(0.105)	(0.0604)
Rural Hukou	-0.0246	-0.00868	1.362***	-0.102***	-0.675***	-0.724***
	(0.0258)	(0.0173)	(0.425)	(0.0266)	(0.154)	(0.154)
Constant	0.290	0.119	-0.510	1.042***	10.23***	5.723***
	(0.233)	(0.164)	(4.183)	(0.299)	(1.736)	(1.173)
Observations	5,355	5,355	5,355	5,355	5,355	5,355
R-squared	0.100	0.144	0.136	0.134	0.188	0.143
Interview Year Dummy	YES	YES	YES	YES	YES	YES
Year of Birth FE	YES	YES	YES	YES	YES	YES
Prefecture FE	YES	YES	YES	YES	YES	YES

Note:
1. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the prefecture and year of birth levels and reported in parentheses.
2. This table shows the gender-specific effects using the subsample of people who were born after 1955.
3. In Columns (1-2) we examine the impact of violence (measured by log of number of collective conflicts) and famine (measured by excessive death rates) on the predicted 8-year diabetes risk and 10-year CVD risk.
4. Column (3) shows the results using CESD-10 as the measure of depression, and Columns (4-6) use three measures of cognition: i) graphical cognition, ii) mental intactness, and iii) memory ability using word recall questions (see the chapter for details).
5. *EDR* is the province-level excessive death rates in 1959, 1960 and 1961, which is calculated by subtracting the average death rates between year 1956-1958 from the annual death rates in the three famine years. The variable *Conflicts* measure the level of violence in each of the years between 1966 and 1971. It is measured at the prefecture level per thousand population. The variable *Rural* is the indicator for rural *hukou* at birth.

4.7 Conclusion

We estimate the long-run impact of exposure *in utero* to the Chinese Great Famine and the Cultural Revolution on physical and mental health later in life. We merge data on the regional violence during the Cultural Revolution and the excessive death rates during the Chinese Great Famine with data from the China Health and Retirement Longitudinal Study (CHARLS) survey. While the Chinese Great Famine was responsible for severe malnutrition and a stressful environment for pregnant mothers in rural areas, the Cultural Revolution mainly affected pregnant mothers in

urban areas.

The results show that *in utero* exposure to the famine increases the risk of getting type-2 diabetes later in life, and that the effect is more pronounced for females. A 1% increase in excessive death rates increases diabetes risk by around 0.06 percentage points for girls who were exposed to famine. Boys exposed to the famine are shown to have a higher tendency to develop depressive symptoms later in life, while exposure to the Cultural Revolution has a negative effect on their cognitive abilities. As in our data we only observe those who survived into adulthood, our results are likely to underestimate the negative effect of adverse events at birth on health later in life.

Overall, our results suggest that early life intervention is crucial. Future research should focus on disentangling the mechanisms through which adverse conditions around birth translate into poor health outcomes later in life, which is crucial to identify potential targets for prevention and intervention strategies.

4.A Eight-year Diabetes Risk Calculation

For the eight-year diabetes risk calculation, we use the prediction model in Table 4 of Wilson et al. (2007) which is reproduced in Table 4.A.1 . We adopt the prediction model using continuous explanatory variables. The general prediction model for the 8-year risk $\hat{p}_{Diabetes}$ is

$$\hat{p}_{Diabetes} = \frac{1}{1 + \exp(-\sum_{i=1}^{m} \hat{\gamma}_i X_i)}, \tag{4.A.1}$$

where the m risk factors include age, gender, parental history of diabetes mellitus, BMI, systolic blood pressure, HDL cholesterol level per mg/dL, triglyceride level per mg/dL, waist circumference in cm, and fasting glucose per ng/dL. Notably, because in the data set we do not have information on parental history of diabetes, we use the default value 0.17 suggested by Wilson et al. (2007) for all respondents. Similar to the calculation of CVD, we assign value 1 to the risk score if the respondents have been diagnosed with diabetes or are now taking the medications to control the blood sugar levels.

Table 4.A.1. Diabetes Risk Factors, Table 4 of Wilson et al. (2007)

Variable	OR (95% CI)	p-value
Age, y	0.99 (0.97-1.01)	.42
Male	0.65 (0.41-1.02)	.06
Parental history of diabetes mellitus	1.55 (1.01-2.38)	.04
BMI	1.04 (0.97-1.11)	.24
Systolic blood pressure, mm Hg	1.01 (1.00-1.02)	.11
HDL-C level per mg/dL	0.96 (0.95-0.98)	$<$.001
Triglyceride level per mg/dL	1.00 (1.00-1.00)	.16
Waist circumference, cm	1.05 (0.97-1.12)	.22
Fasting glucose level, mg/dL	1.15 (1.12-1.17)	$<$.001

4.B Ten-year CVD Risk Calculation

The procedures for calculating the 10-year CVD risk can be found in the appendix of D'agostino et al. (2008). The study estimates a simple CVD risk prediction model and provides the regression coefficients of the model using data from the Framingham Heart Study. We reproduce the estimates in Table 4.B.1. The general formula for calculating the CVD risk $\hat{p}_{CVD}$ is

$$\hat{p}_{CVD} = 1 - S_0(10)^{\exp(\sum_{i=1}^{k} \hat{\beta}_i X_i - \sum_{i=1}^{k} \hat{\beta}_i \bar{X}_i)}, \tag{4.B.1}$$

where $S_0(10)$ is the baseline survival rate in 10 years. In this chapter, we use the same baseline gender-specific survival rates as D'agostino et al. (2008), which is 0.95012 for females and 0.88936 for males. The covariate vector X_i includes the log of the i_{th} risk factor where i ranges from 1 to k. Vector $\bar{X}_i$ stands for the sample means of the covariates. The risk factors include age, total cholesterol, high density level cholesterol, a dummy for whether being undertreated for hypertension, a dummy for whether smoking or not, and a dummy for whether being diagnosed with diabetes or not. The regression coefficients $\hat{\beta}$ are provided in the Table 2 of D'agostino et al. (2008). For those who were diagnosed with CVD, we assign value 1 to the risk score.

Table 4.B.1. CVD Risk Factors, Table 2 of D'agostino et al. (2008)

Variable	$\hat{\beta}$	p-value
Women ($S_0(10) = 0.95012$)		
Log of age	2.32888	$<.0001$
Log of total cholesterol	1.20904	$<.0001$
Log of HDL cholesterol	-0.70833	$<.0001$
Log of SBP if not treated	2.76157	$<.0001$
Log of SBP if treated	2.82263	$<.0001$
Smoking	0.52873	$<.0001$
Diabetes	0.69154	$<.0001$
Men ($S_0(10) = 0.88936$)		
Log of age	3.06117	$<.0001$
Log of total cholesterol	1.12370	$<.0001$
Log of HDL cholesterol	-0.93263	$<.0001$
Log of SBP if not treated	1.93303	$<.0001$
Log of SBP if treated	1.99881	$<.0001$
Smoking	0.65451	$<.0001$
Diabetes	0.57367	$<.0001$

References

Akbulut-Yuksel, M. (2017). War during childhood: The long run effects of warfare on health. *Journal of Health Economics*, *53*, 117–130.

Alessie, R., Angelini, V. & Pasini, G. (2014). Is it true love? Altruism versus exchange in time and money transfers. *De Economist*, *162*(2), 193–213.

Alessie, R., Angelini, V., van den Berg, G. J., Mierau, J. O. & Viluma, L. (2019). Economic conditions at birth and cardiovascular disease risk in adulthood: Evidence from post-1950 cohorts. *Social Science & Medicine*, *224*, 77–84.

Alessie, R., Angelini, V. & van Santen, P. (2013). Pension wealth and household savings in Europe: Evidence from SHARELIFE. *European Economic Review*, *63*, 308–328.

Almond, D. & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, *25*(3), 153–72.

Almond, D., Currie, J. & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360–1446.

Almond, D. & Mazumder, B. (2011). Health capital and the prenatal environment: The effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics*, *3*(4), 56–85.

Bai, L. (2014). Economic legacies of the Cultural Revolution. *Working Paper, Department of Economics, University of Edinbourgh*.

Bai, L. & Wu, L. (2020). Political movement and trust formation: Evidence from the Cultural Revolution (1966–76). *European Economic Review*, *122*, 103331.

Barham, V., Boadway, R., Marchand, M. & Pestieau, P. (1995). Education and the poverty trap. *European Economic Review*, *39*(7), 1257–1275.

Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ*, *311*(6998), 171–174.

Barker, D. J., Bagby, S. P. & Hanson, M. A. (2006). Mechanisms of disease: In utero programming in the pathogenesis of hypertension. *Nature Clinical Practice Nephrology*, *2*(12), 700–707.

Barker, D. J., Godfrey, K. M., Gluckman, P. D., Harding, J. E., Owens, J. A. & Robinson, J. S. (1993). Fetal nutrition and cardiovascular disease in adult life. *Lancet*, *341*(8850), 938–941.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, *82*(6), 1095–1117.

Becker, G. S. (2009). *A treatise on the family*. Harvard University Press.

Björgvinsson, T., Kertz, S. J., Bigda-Peyton, J. S., McCoy, K. L. & Aderka, I. M. (2013). Psychometric properties of the CES-D-10 in a psychiatric sample. *Assessment*, *20*(4), 429–436.

Bock, J., Wainstock, T., Braun, K. & Segal, M. (2015). Stress in utero: Prenatal programming of brain plasticity and cognition. *Biological Psychiatry*, *78*(5), 315–326.

Bonnin, M. (2006). The "lost generation": Its definition and its role in today's Chinese elite politics. *Social Research*, *73*(1), 245–274.

Bonsang, E. (2009). Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, *28*(1), 143–154.

Braithwaite, E. C., Murphy, S. E. & Ramchandani, P. G. (2014). Prenatal risk factors for depression: A critical review of the evidence and potential mechanisms. *Journal of Developmental Origins of Health and Disease*, *5*(5), 339–350.

Bramsen, I., Deeg, D. J., van der Ploeg, E. & Fransman, S. (2007). Wartime stressors and mental health symptoms as predictors of late-life mortality in World War II survivors. *Journal of Affective Disorders*, *103*(1), 121 - 129.

Brown, M., Karl Scholz, J. & Seshadri, A. (2011). A new test of borrowing constraints for education. *Review of Economic Studies*, *79*(2), 511–538.

Burns, L. R. & Liu, G. G. (2017). *China's healthcare system and reform*. Cambridge University Press.

Cai, F., Giles, J. & Meng, X. (2006). How well do children insure parents against low retirement income? An analysis using survey data from urban China. *Journal of Public Economics*, *90*(12), 2229–2255.

Cameron, A. C. & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317–372.

Checchi, D., Fiorio, C. V. & Leonardi, M. (2014). Parents' risk aversion and children's educational attainment. *Labour Economics*, *30*, 164–175.

Chen, X., Crimmins, E., Hu, P., Kim, J. K., Meng, Q., Strauss, J., ... Zhao, Y. (2019). Venous blood-based biomarkers in the China Health and Retirement Longitudinal Study: Rationale, design, and results from the 2015 wave. *American Journal of Epidemiology*, *188*(11), 1871–1877.

Chen, X., Smith, J., Strauss, J., Wang, Y. & Zhao, Y. (2017). China Health and Retirement Longitudinal Study (CHARLS). *Encyclopedia of Geropsychology*, 463–469.

Cogan, J. F. (1981). Fixed costs and labor supply. *Econometrica*, *49*(4), 945–963.

Conroy, R., Pyörälä, K., Fitzgerald, A. e., Sans, S., Menotti, A., De Backer, G., ... Wilhelmsen, L. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *European Heart Journal*, *24*(11), 987–1003.

D'agostino, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M. & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care. *Circulation*, *117*(6), 743–753.

De Groot, R. H., Stein, A. D., Jolles, J., van Boxtel, M. P., Blauw, G.-J., Van de Bor, M. & Lumey, L. H. (2011). Prenatal famine exposure and cognition at age 59 years. *International Journal of Epidemiology*, *40*(2), 327–337.

Deng, Z. & Treiman, D. J. (1997). The impact of the Cultural Revolution on trends in educational attainment in the People's Republic of China. *American Journal of Sociology*, *103*(2), 391–428.

Dinkelman, T. (2017). Long-run health repercussions of drought shocks: Evidence from South African homelands. *Economic Journal*, *127*(604), 1906–1939.

Di Renzo, G. C., Rosati, A., Sarti, R. D., Cruciani, L. & Cutuli, A. M. (2007). Does fetal sex affect pregnancy outcome? *Gender Medicine*, *4*(1), 19–30.

Doblhammer, G., Van den Berg, G. J. & Fritze, T. (2013). Economic conditions at the time of birth and cognitive abilities late in life: Evidence from ten European countries. *PloS One*, *8*(9), e74915.

Doblhammer, G., van den Berg, G. J. & Lumey, L. H. (2013). A re-analysis of the long-term effects on life expectancy of the Great Finnish Famine of 1866–68. *Population Studies*, *67*(3), 309–322.

Duan, N. (1983). Smearing estimate: A nonparametric retransformation method. *Journal of the American Statistical Association*, *78*(383), 605–610.

Eriksson, J. G., Kajantie, E., Osmond, C., Thornburg, K. & Barker, D. J. (2010). Boys live dangerously in the womb. *American Journal of Human Biology*, *22*(3), 330–335.

Fan, W. (2017). Education delayed but not denied: The Chinese Cultural Revolution cohort returning to school. *Advances in Life Course Research*, *33*, 53–65.

Gade, D. M. & Wenger, J. B. (2011). Combat exposure and mental health: The long-term effects among US Vietnam and Gulf war veterans. *Health Economics*, *20*(4), 401-416.

Gluckman, P. D. & Hanson, M. A. (2004). Living with the past: Evolution, development, and patterns of disease. *Science*, *305*(5691), 1733–1736.

Grimard, F. & Laszlo, S. (2014). Long-term effects of civil conflict on women's health outcomes in peru. *World Development*, *54*, 139 - 155.

Guiso, L., Jappelli, T. & Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *The American Economic Review*, *86*(1), 158.

Haider, S. & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, *96*(4), 1308–1320.

Hales, C. N. & Barker, D. J. (1992). Type 2 (non-insulin-dependent) diabetes mellitus: The thrifty phenotype hypothesis. *Diabetologia*, *35*(7), 595–601.

Ham, J. C. & Song, H. (2014). The determinants of bargaining power in an empirical model of transfers between adult children, parents, and in-laws for South Korea. *Journal of Development Economics*, *109*, 73–86.

Hamada, H. & Matthews, S. G. (2019). Prenatal programming of stress responsiveness and behaviours: Progress and perspectives. *Journal of Neuroendocrinology*, *31*(3), 1-13.

Han, J., Suen, W. & Zhang, J. (2019). Picking up the losses: The impact of the Cultural Revolution on human capital reinvestment in urban China. *Journal of Human Capital*, *13*(1), 56–94.

Hausman, J. A. (1980). The effect of wages, taxes, and fixed costs on women's labor force participation. *Journal of Public Economics*, *14*(2), 161–194.

Heckman, J. J. (2005). China's human capital investment. *China Economic Review*, *16*(1), 50–70.

Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., ... Lumey, L. H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences*, *105*(44), 17046–17049.

Hu, X. F., Liu, G. G. & Fan, M. (2017). Long-term effects of famine on chronic diseases: Evidence from China's great leap forward famine. *Health Economics*, *26*(7), 922–936.

Huang, W. & Zhou, Y. (2013). Effects of education on cognition at older ages: Evidence from China's great famine. *Social Science & Medicine*, *98*, 54–62.

Ingemarsson, I. (2003). Gender aspects of preterm birth. *BJOG: an international journal of obstetrics and gynaecology*, *110*(Suppl 20), 34–38.

Islam, A., Raschky, P. & Smyth, R. (2017). The long-term health effects of mass political violence: Evidence from China's Cultural Revolution. *Social Indicators Research*, *132*(1), 257–272.

Jappelli, T., Pischke, J.-S. & Souleles, N. S. (1998). Testing for liquidity constraints in Euler equations with complementary data sources. *Review of Economics and statistics*, *80*(2), 251–262.

Kalwij, A., Pasini, G. & Wu, M. (2014). Home care for the elderly: the role of relatives, friends and neighbors. *Review of Economics of the Household*, *12*(2), 379–404.

Keane, M. P. & Wolpin, K. I. (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, *42*(4), 1051–1103.

Kim, S., Fleisher, B. & Sun, J. Y. (2017). The long-term health effects of fetal malnutrition: Evidence from the 1959–1961 China great leap forward famine. *Health Economics*, *26*(10), 1264–1277.

Kuh, D. & Shlomo, Y. B. (2004). *A life course approach to chronic disease epidemiology* (No. 2). Oxford University Press.

Lei, X., Hu, Y., McArdle, J. J., Smith, J. P. & Zhao, Y. (2012). Gender differences in cognition among older adults in China. *Journal of Human Resources*, *47*(4), 951–971.

Li, H., Rosenzweig, M. & Zhang, J. (2010). Altruism, favoritism, and guilt in the allocation of family resources: Sophie's choice in mao's mass send-down movement. *Journal of Political Economy*, *118*(1), 1–38.

Lin, J. Y. & Yang, D. T. (2000). Food availability, entitlements and the Chinese famine of 1959–61. *Economic Journal*, *110*(460), 136–158.

Lochner, L. J. & Monge-Naranjo, A. (2011). The nature of credit constraints and human capital. *American Economic Review*, *101*(6), 2487–2529.

Lumey, L. H., Stein, A. D. & Susser, E. (2011). Prenatal famine and adult health. *Annual Review of Public Health*, *32*, 237–262.

Lumey, L. H., Terry, M. B., Delgado-Cruzata, L., Liao, Y., Wang, Q., Susser, E., ... Santella, R. M. (2011). Adult global DNA methylation in relation to pre-natal nutrition. *International Journal of Epidemiology*, *41*(1), 116–123.

McGowan, P. O. & Matthews, S. G. (2017). Prenatal stress, glucocorticoids, and developmental programming of the stress response. *Endocrinology*, *159*(1), 69–82.

Meng, L. (2013). Evaluating China's poverty alleviation program: A regression discontinuity approach. *Journal of Public Economics*, *101*, 1–11.

Meng, X. & Gregory, R. (2007). Exploring the impact of interrupted education on earnings: The educational cost of the Chinese Cultural Revolution. *Working Paper*.

Meng, X. & Gregory, R. G. (2002). The impact of interrupted education on subsequent educational attainment: A cost of the Chinese Cultural Revolution. *Economic Development and Cultural Change*, *50*(4), 935–959.

Meng, X., Qian, N. & Yared, P. (2015). The institutional causes of China's great famine, 1959–1961. *Review of Economic Studies*, *82*(4), 1568–1611.

Meng, X. & Zhao, G. (2016). The long shadow of the Chinese Cultural Revolution: The intergenerational transmission of education. *IZA discussion paper series, IZA DP No. 10460*.

Mincer, J. A. (1974). Age and experience profiles of earnings. In *Schooling, experience, and earnings* (pp. 64–82). Cambridge, MA: National Bureau of Economic Research.

Molenaar, N. M., Tiemeier, H., van Rossum, E. F., Hillegers, M., Bockting, C., Hoogendijk, W., … El Marroun, H. (2019). Prenatal maternal psychopathology and stress and offspring HPA axis function at 6 years. *Psychoneuroendocrinology*, *99*, 120–127.

Mueller, B. R. & Bale, T. L. (2008). Sex-specific programming of offspring emotionality after stress early in pregnancy. *Journal of Neuroscience*, *28*(36), 9055–9065.

Oliveira, J. (2016). The value of children: Inter-generational support, fertility, and human capital. *Journal of Development Economics*, *120*, 1–16.

Portrait, F., Teeuwiszen, E. & Deeg, D. (2011). Early life undernutrition and chronic diseases at older ages: The effects of the Dutch famine on cardiovascular diseases and diabetes. *Social Science & Medicine*, *73*(5), 711–718.

Psacharopoulos, G. (1985). Returns to education: A further international update and implications. *Journal of Human Resources*, *20*(4), 583–604.

Raut, L. K. & Tran, L. H. (2005). Parental human capital investment and old-age transfers from children: Is it a loan contract or reciprocity for Indonesian families? *Journal of Development Economics*, *77*(2), 389–414.

Schmelter, V., Antonaci, L., Fröhlich, S., Herebian, D., Verde, P., Heil, M. & Hoehn, T. (2019). Quantification of fetal steroids in nails of neonates to quantify prenatal stress and growth restriction. *Biological Psychology*, *140*, 81–85.

Schultz, T. W. (1961). Investment in human capital. *American Economic Review*, *51*(1), 1–17.

Selvarajah, S., Kaur, G., Haniff, J., Cheong, K. C., Hiong, T. G., van der Graaf, Y. & Bots, M. L. (2014). Comparison of the Framingham Risk Score, SCORE and WHO/ISH cardiovascular risk prediction models in an Asian population. *International Journal of Cardiology*, *176*(1), 211–218.

Strauss, K., Dapp, U., Anders, J., von Renteln-Kruse, W. & Schmidt, S. (2011). Range and specificity of war-related trauma to posttraumatic stress; depression and general health perception: Displaced former World War II children in late life. *Journal of Affective Disorders*, *128*(3), 267 - 276.

Teerawichitchainan, B. & Korinek, K. (2012). The long-term impact of war on health and wellbeing in northern Vietnam: Some glimpses from a recent survey. *Social Science & Medicine*, *74*(12), 1995 - 2004.

Tobi, E. W., Lumey, L. H., Talens, R. P., Kremer, D., Putter, H., Stein, A. D., ... Heijmans, B. T. (2009). DNA methylation differences after exposure to prenatal famine are common and timing-and sex-specific. *Human Molecular Genetics*, *18*(21), 4046–4053.

Vaiserman, A. & Lushchak, O. (2019). Prenatal malnutrition-induced epigenetic dysregulation as a risk factor for type 2 diabetes. *International Journal of Genomics*, *2019*, 1–11.

Van den Berg, G. J., Pinger, P. R. & Schoch, J. (2015). Instrumental variable estimation of the causal effect of hunger early in life on health later in life. *Economic Journal*, *126*(591), 465–506.

Walder, A. G. (2014). Rebellion and repression in China, 1966–1971. *Social Science History*, *38*(3-4), 513–539.

Walder, A. G. & Su, Y. (2003). The Cultural Revolution in the countryside: Scope, timing and human impact. *China Quarterly*, *173*, 74–99.

Wang, J. (2019). Impoverishing or improving? The impact of the Chinese Cultural Revolution on lifetime income. *Working Paper*.

Wang, J., Alessie, R. & Angelini, V. (2017). Why do parents underinvest in their children's education? Evidence from China. *Working Paper*.

Wang, J., Alessie, R. & Angelini, V. (2020). Exposure in utero to adverse events and health late-in-life: Evidence from China. *Working Paper*.

Welberg, L. A. & Seckl, J. R. (2001). Prenatal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology*, *13*(2), 113–128.

Wilson, P. W., Meigs, J. B., Sullivan, L., Fox, C. S., Nathan, D. M. & D'Agostino, R. B. (2007). Prediction of incident diabetes mellitus in middle-aged adults: The Framingham offspring study. *Archives of Internal Medicine*, *167*(10), 1068–1074.

Xie, Y., Jiang, Y. & Greenman, E. (2008). Did send-down experience benefit youth? A reevaluation of the social consequences of forced urban-rural migration during China's Cultural Revolution. *Social Science Research*, *37*(2), 686–700.

Xu, H., Zhang, Z., Li, L. & Liu, J. (2017). Early life exposure to China's 1959–61 famine and midlife cognition. *International Journal of Epidemiology*, *47*(1), 109–120.

Xu, Y., Wang, L., He, J., Bi, Y., Li, M., Wang, T., ... the 2010 China Noncommunicable Disease Surveillance Group (2013). Prevalence and control of diabetes in Chinese adults. *JAMA*, *310*(9), 948–959.

Zeldes, S. P. (1989). Consumption and liquidity constraints: An empirical investigation. *Journal of Political Economy*, *97*(2), 305–346.

Zhang, J., Liu, P.-W. & Yung, L. (2007). The Cultural Revolution and returns to schooling in China: Estimates based on twins. *Journal of Development Economics*, *84*(2), 631–639.

Zhang, X. & Kanbur, R. (2005). Spatial inequality in education and health care in China. *China Economic Review*, *16*(2), 189–204.

Zhao, Y., Hu, Y., Smith, J. P., Strauss, J. & Yang, G. (2012). Cohort profile: The China health and Retirement Longitudinal Study (CHARLS). *International Journal of Epidemiology*, *43*(1), 61–68.

Zhao, Y., Strauss, J., Yang, G., Giles, J., Hu, P., Hu, Y., ... Wang, Y. (2013). China Health and Retirement Longitudinal Study–2011–2012 national baseline users guide. *Beijing: National School of Development, Peking University*, 1–56.

Zhou, W. (2013). How does a traumatic experience during youth affect life later? The long-term impact of the send-down program during the Chinese Cultural Revolution. *Working Paper*.

Zhou, X. & Hou, L. (1999). Children of the Cultural Revolution: The state and the life course in the People's Republic of China. *American Sociological Review*, *64*(1), 12–36.

Samenvatting (Summary in Dutch)

Als leefomstandigheden en de technologie in de zorg verbeteren, stijgt de individuele levensverwachting. Dit patroon doet zich niet alleen voor in ontwikkelde landen, maar ook in ontwikkelingslanden als China. Volgens het rapport *World Population Prospect 2017* van de Verenigde Naties is de gemiddelde levensverwachting bij geboorte in China gestegen van 69,7 in de periode 1990-1995 naar 76,5 in de periode van 2015-2020, en zal zij voor de periode 2045-2050 verder stijgen naar 81,1. De volksgezondheid draait echter niet alleen om kwantiteit, maar ook om kwaliteit van leven. Mensen leven tegenwoordig langer, maar willen dat wel in goede fysieke en financile gezondheid doen. Uit de literatuur blijkt steeds meer dat omstandigheden in de vroege levensfase belangrijke determinanten zijn voor het welzijn op oudere leeftijd. Om te weten hoe we de financile, fysieke en geestelijke gezondheid van ouderen kunnen verbeteren, moeten we dus naar de levenscyclus als geheel kijken. In dit proefschrift bestudeer ik de omstandigheden in de vroege levensfase die van invloed zijn op het welzijn op oudere leeftijd. Daarbij richt ik me met name op twee factoren die mogelijk van invloed zijn op het welzijn van ouderen: steun van kinderen en tegenslagen op jonge leeftijd. China kent een minder royaal socialezekerheidsstelsel dan wij, en dus moeten arme huishoudens daar andere vormen van ondersteuning zoeken om op oudere leeftijd in hun levensonderhoud te kunnen voorzien. Een voor de hand liggende oplossing is dat ouders hulp krijgen van hun kinderen. De mate waarin kinderen hun ouders economisch ondersteunen, zou verder bepaald kunnen worden door het menselijk kapitaal dat ouders in hun kinderen genvesteerd hebben toen die de schoolgaande leeftijd hadden. Een tweede factor is de rol van tegenslagen op jonge leeftijd. In China heeft het huidige cohort

ouderen de Grote Chinese Hongersnood en/of de Culturele Revolutie misschien nog meegemaakt, wat mogelijk heeft geleid tot verlies van inkomsten tijdens hun leven en negatieve langetermijngevolgen voor de gezondheid. De belangrijkste gegevens zijn ontleend aan de *China Health and Retirement Longitudinal Study* (CHARLS, longitudinaal onderzoek naar gezondheid en pensionering in China). CHARLS is een nationaal representatief onderzoek gericht op mensen boven de 45 jaar en omvat gegevens over intergenerationele overdracht, het arbeidsverleden van individuele personen tijdens hun levens, en bloedgerelateerde biomarkergegevens.

Waarom investeren ouders te weinig in onderwijs voor hun kinderen?

In hoofdstuk 2 bestuderen we de steun die kinderen hun ouders op hun oude dag bieden. De steun van kinderen aan hun oude ouders wordt, met name in ontwikkelingslanden, gezien als een goede aanvulling op sociale verzekeringen van de overheid (zie, bijv., Cai et al., 2006; Oliveira, 2016). De mate waarin kinderen hun oude ouders ondersteunen, hangt af van de mate waarin de kinderen in hun eigen levensonderhoud kunnen voorzien, wat weer wordt bepaald door het menselijk kapitaal dat de ouders in het onderwijs van hun kinderen hebben genvesteerd toen die de schoolgaande leeftijd hadden (zie, bijv., Alessie et al., 2014; Cai et al., 2006; Raut & Tran, 2005). Het verband tussen menselijk kapitaal dat door ouders in kinderen is genvesteerd en de steun die kinderen hun ouders bieden, wordt mogelijk ook benvloed door bepaalde sociaaleconomische en omgevingsfactoren toen ouders beslissingen moesten nemen over investeren in het menselijk kapitaal van hun kinderen. Twee belangrijke omstandigheden zijn liquiditeitsbeperking en vaste kosten voor onderwijs (zie, bijv., Keane & Wolpin, 2001; Lochner & Monge-Naranjo, 2011). Zo kan het zijn dat ouders, hoewel ze bereid zijn te investeren, toch onvoldoende investeren omdat ze financieel te zwak zijn om geld van banken of buren te lenen, of omdat de instapkosten om in het menselijk kapitaal van hun kinderen te investeren, te hoog voor hen is. Daardoor wordt er te weinig genvesteerd in het menselijk kapitaal van hun kinderen. Het is van groot belang deze twee omstandigheden te betrekken bij het verhaal van intergenerationele overdracht en de voorspellingen te toetsen aan de hand van Chinese gegevens, omdat we dan kunnen achterhalen welke omstandigheden ouders beletten te investeren in het onderwijs van hun kinderen. De overheid kan dan gericht ingrijpen. De eerste

onderzoeksvraag is daarom:

In hoeverre worden de mate waarin ouders investeren in het menselijk kapitaal van hun kinderen en de steun van kinderen aan hun oude ouders benvloed door het bestaan van liquiditeitsbeperking en vaste kosten voor onderwijs?

Om deze onderzoeksvraag te bestuderen, wordt in dit hoofdstuk gebruikgemaakt van een kader voor intergenerationele overdracht op basis van wederzijds altrusme, zoals voorgesteld door Raut & Tran (2005). We breiden het model uit door de liquiditeitsbeperking van ouders en de vaste kosten voor onderwijs er ook in op te nemen, en we toetsen de voorspellingen op empirische wijze aan de hand van de onderzoeksgegevens van de *2013 China Health and Retirement Longitudinal Study* (CHARLS).

De resultaten tonen, in de eerste plaats, aan dat er, ook als er geen sprake is van niet-bindende liquiditeitsbeperking en vaste kosten, te weinig wordt genvesteerd in het menselijk kapitaal van kinderen wegens het onderhandelingsproces tussen ouders en kinderen, die allen niet volmaakt altrustisch zijn. In de tweede plaats worden investeringen in menselijk kapitaal nog verder beperkt door de bindende liquiditeitsbeperking van ouders. Ten slotte tonen we aan dat het bestaan van vaste kosten voor onderwijs de overdracht van kinderen naar hun oude ouders versterkt.

Effect van de Culturele Revolutie op het inkomen tijdens de levensloop

In hoofdstuk 3 kijk ik naar het effect van blootstelling, op jonge leeftijd (5-19 jaar), aan ongunstige politieke gebeurtenissen op het inkomen tijdens de levensloop, waarbij ik de Chinese Culturele Revolutie (1966-1976) als natuurlijk experiment gebruik. De Culturele Revolutie had in elke prefectuur een andere uitwerking, en ik kan de variaties in geweldpleging in de prefecturen dus gebruiken om het langetermijneffect van de Culturele Revolutie vast te stellen. Ook zijn er verschillende partijen bij de Culturele Revolutie betrokken: de algemene bevolking, de slachtoffers, en de mensen die geweld creerden door zich aan te sluiten bij de Rode Garde. Het is daarom ook van belang te onderzoeken of leden van de Rode Garde op vergelijkbare wijze worden getroffen als mensen die geen lid waren van de Rode Garde. Bestaande studies meten inkomen aan de hand van transversale inkomensgegevens (zie, bijv., Meng & Gregory, 2007), maar ik ga uit van inkomen tijdens de levensloop, waarbij

het menselijk kapitaal gedurende de levenscyclus wordt gemeten. In hoofdstuk 3 is mijn onderzoeksvraag:

Wat is het effect van blootstelling aan de Culturele Revolutie tijdens de schoolgaande leeftijd (5-19) op het individuele inkomen tijdens de levensloop? Zijn er wat dit betreft verschillen tussen leden van de Rode Garde en mensen die geen lid waren van de Rode Garde?

Om deze onderzoeksvraag te bestuderen, maak ik gebruik van gegevens uit de dataset van de CHARLS-enqułte over levensgeschiedenissen. De CHARLS-enqułte over levensgeschiedenissen verzamelt informatie over alle levensgebeurtenissen, zoals werk- en onderwijsperiodes, van respondenten die aanvankelijk in 2011 en 2013 zijn ondervraagd. Op basis van deze gegevens kunnen twee criteria worden opgesteld voor inkomen tijdens de levensloop: gemiddeld inkomen tijdens de levensloop per gewerkt jaar, en inkomensgroei tussen de eerste en de laatste baan in het beroepsleven. Vervolgens voeg ik de variaties, op prefectuurniveau, in geweldpleging tijdens de Culturele Revolutie (1966-1976) toe aan het individueel inkomen tijdens de levensloop.

Uit de resultaten blijkt dat het gemiddeld inkomen tijdens de levensloop per gewerkt jaar van mannen die hebben blootgestaan aan 1% meer geweld, 8 procentpunten lager is, maar dat dit effect bij vrouwen verwaarloosbaar is. Bovendien is niet iedereen in dezelfde mate getroffen. Mannen die actief lid zijn geweest van de Rode Garde, profiteerden in feite van de Culturele Revolutie omdat hun gemiddeld inkomen tijdens de levensloop per gewerkt jaar met 16% toenam. Ik constateer ook dat groei van het individueel inkomen tijdens de levenscyclus niet wordt benvloed door blootstelling aan de Culturele Revolutie.

Blootstelling in de baarmoeder aan ongunstige gebeurtenissen, en gezondheid op latere leeftijd

In dit hoofdstuk kijken we naar de effecten van blootstelling in de baarmoeder aan ongunstige gebeurtenissen op de gezondheid op latere leeftijd. In dit onderzoek maken we gebruik van twee natuurlijke experimenten: de Grote Chinese Hongersnood, die vooral het platteland trof, en de Culturele Revolutie, die vooral de stedelijke gebieden trof. Om de gezondheid te meten, maken we gebruik van bloedgerelateerde biomarkergegevens en stellen we risicoscores voor diabetes en

hart- en vaatziekten vast. Dat is belangrijk omdat we zo de onderliggende risico's voor het krijgen van diabetes en hart- en vaatziekten kunnen achterhalen, met name wanneer deze chronische ziekten zich nog niet hebben geopenbaard. Behalve voor de fysieke gezondheid stellen we ook criteria op voor de geestelijke gezondheid, zoals depressie en cognitie. Omdat uit de biomedische literatuur blijkt dat jongetjes op een andere manier weefsel en organen ontwikkelen in de baarmoeder dan meisjes (zie, bijv., Eriksson et al., 2010), doen we ook afzonderlijk onderzoek per sekse. De derde onderzoeksvraag is daarom:

Wat zijn de effecten van blootstelling in de baarmoeder aan de Grote Chinese Hongersnood en de Culturele Revolutie op de fysieke en geestelijke gezondheid op oudere leeftijd? Is het patroon van die effecten seksegebonden?

Om de dataset op te stellen, meten we de hoogte van de oversterftecijfers op provincieniveau, die worden vastgesteld door de gemiddelde sterftecijfers in de jaren waarin er geen hongersnood was, af te trekken van de sterftecijfers tijdens de jaren van de hongersnood, en maken we gebruik van dezelfde criteria voor het meten van geweldpleging op prefectuurniveau als in hoofdstuk 3. Vervolgens voegen we deze gegevens toe aan de individuele gezondheidsresultaten uit de CHARLS-enqułtes van 2011 en 2013, waarbij we gebruikmaken van geboorteplaats- en het geboortejaargegevens. Wat betreft gezondheidsresultaten stellen we een risicoscore over acht jaar op voor diabetes, aan de hand van het algoritme uit Wilson et al. (2007), een risicoscore over tien jaar voor hart- en vaatziekten, aan de hand van het algoritme uit D'agostino et al. (2008), en verschillende criteria op het gebied van cognitie en depressie.

Uit de resultaten blijkt dat meisjes die in de baarmoeder hebben blootgestaan aan een 1% hoger, door hongersnood veroorzaakt oversterftecijfer, 0,06 procentpunt meer kans hebben om diabetes type 2 te krijgen. Voor jongetjes zijn de effecten verwaarloosbaar. Jongetjes die tijdens de Culturele Revolutie hebben blootgestaan aan meer geweld, blijken bovendien minder goede cognitieve vaardigheden te hebben.

Beleidsadvies

Om de problemen door vergrijzing en de daarbij behorende stijgende zorgkosten het hoofd te bieden, moeten overheden proactief beleid opstellen en uitvoeren. Mijn proefschrift biedt enkele inzichten voor het opstellen van relevant beleid. De

belangrijkste boodschap is eenvoudig: ingrijpen in de vroege levensfase levert meer voordelen op de lange termijn op, zeker voor mensen die te maken hebben gehad met tegenslag.

Algemeen beleidsadvies

In hoofdstuk 2 van mijn proefschrift constateren we dat zowel bindende liquiditeitsbeperkingen als het bestaan van vaste onderwijskosten een verklaring kunnen vormen voor het feit dat ouders te weinig investeren in het onderwijs van hun kinderen. Het verlichten van de bindende liquiditeitsbeperking bij de ouders en het wegnemen van de vaste kosten kunnen dus bijdragen aan meer investeringen in het onderwijs van de kinderen. Wat betreft het verlichten van de liquiditeitsbeperking van huishoudens zouden de huidige, landelijk uitgerolde armoedebestrijdingsprogrammas (zie, bijv., Meng, 2013) de onderwijsresultaten van kinderen uit arme gezinnen kunnen verbeteren. Overheden kunnen echter nog meer doen om vaste onderwijskosten te verlagen. Die kosten kunnen materieel van aard zijn, bijvoorbeeld vervoerskosten en lesgeld, en kunnen dan verlicht worden door uitbreiding van studieleningen tegen lage rentes voor studenten of door andere beleidsinstrumenten. Sommige ontwikkelde landen hebben al bepaalde beleidsmaatregelen gemplementeerd. Zo krijgen Nederlandse studenten korting op het openbaar vervoer (studentenreisproduct).[1] Vaste kosten kunnen ook immaterieel van aard zijn, bijvoorbeeld ongelijke kansen voor studenten uit gezinnen met een laag inkomen wat betreft het volgen van hoogwaardig universitair onderwijs. Zo bieden kustprovincies studenten uit de buurt meer kansen om naar goede universiteiten te gaan, terwijl studenten uit provincies in het westen of midden van het land intensief moeten concurreren voor schaarse kansen (zie, bijv., Zhang & Kanbur, 2005). Als de investering in menselijk kapitaal stijgt naar een optimaal niveau, doen de resultaten in hoofdstuk 2 vermoeden dat de overdracht naar ouders op hoge leeftijd ook zal toenemen. Dit is belangrijk omdat ouderen in een land als China, waar het pensioenstelsel minder goed ontwikkeld is, in hoge mate afhankelijk zijn van financile ondersteuning van hun kinderen. In hoofdstuk 3 van mijn proefschrift constateer ik dat mannen die in de schoolgaande leeftijd hebben blootgestaan aan de Culturele Revolutie, hun gehele leven te maken hebben gehad met verlies van inkomsten. De resultaten impliceren dat het effect van de Culturele Revolutie op het inkomen de gehele levenscyclus doorwerkt. Met andere woorden: wanneer het proces van inves-

[1] Zie bijvoorbeeld: https://www.duo.nl/particulier/aanvragen-studentenreisproduct.jsp.

teren in menselijk kapitaal eenmaal is onderbroken, is het lastig de achterstand in het beroepsleven na de Culturele Revolutie weer in te halen. Als het onderwijs van kinderen dus wordt onderbroken, ofwel door binnenlandse conflicten ofwel door schoolsluitingen, is het belangrijk moeite en geld te steken in kinderen die het meest zouden profiteren van fysiek onderwijs. In hoofdstuk 3 wordt verder geconstateerd dat mensen die zich hebben aangesloten bij de Rode Garde en die meestal al een gunstige sociaaleconomische achtergrond hadden, een veel hoger inkomen tijdens de levensloop hebben. Dit resultaat wekt de suggestie dat gebeurtenissen als de Culturele Revolutie de ongelijkheid op lange termijn zou kunnen versterken.

In hoofdstuk 4 constateren we dat meisjes die in de baarmoeder hebben blootgestaan aan hongersnood, meer kans hebben om diabetes type 2 te krijgen, en dat mensen die blootgestaan hebben aan de Culturele Revolutie op latere leeftijd minder goede cognitieve vaardigheden hebben. De algemene conclusie is dat ongunstige omstandigheden rondom de geboorte aanzienlijke negatieve gevolgen kan hebben op de gezondheid op latere leeftijd. Dit doet vermoeden dat er al op zeer jonge leeftijd, zelfs al tijdens de zwangerschap, moet worden gestart met preventie. In sommige ontwikkelingslanden waar sprake is van binnenlandse conflicten of hongersnoden is het daarom belangrijk prioriteit te geven aan voedselvoorziening aan en het bieden van een veilig onderdak voor zwangere vrouwen.

Beleidsadvies in verband met COVID-19

De recente wereldwijde uitbraak van het coronavirus (COVID-19) heeft in veel landen ernstige bedreigingen op het gebied van de volksgezondheid en de economie veroorzaakt. In de meeste Europese en Oost-Aziatische landen zijn veel mensen overleden aan de gevolgen van deze pandemie. Na een beleid van lockdowns en afstand houden lijkt de situatie in deze landen nu onder controle.

Maar aan de andere kant van de wereld verslechtert de situatie juist. Miljoenen burgers in (ontwikkelings)landen in Afrika en het Midden-Oosten hebben volgens een recent rapport van het Wereldvoedselprogramma niet alleen te maken met de wereldwijde gezondheidscrisis maar ook met zware hongersnoden.[2] Naast het voeren van een beleid van lockdowns en afstand houden is het van groot belang dat overheden onmiddellijk ingrijpen om de voedselvoorziening, met name aan zwangere vrouwen, veilig te stellen. Dit omdat uit de resultaten in hoofdstuk 4 blijkt

[2]https://www.wfp.org/news/wfp-chief-warns-hunger-pandemic-covid-19-spreads-statement-un-security-council.

dat baby's meer kans hebben om op oudere leeftijd chronische ziekten als diabetes te ontwikkelen als hun moeders te kampen hebben met ernstige voedseltekorten.

De dreigende hongerpandemie zal niet alleen pasgeborenen treffen maar kan ook voor de huidige generaties kinderen in de schoolgaande leeftijd verlies van inkomen tijdens de levensloop veroorzaken. Dit rapport van het Wereldvoedselprogramma toont aan dat landen die meer risico lopen op grootschalige hongersnoden, meestal conflicthaarden zijn. Volgens de resultaten uit hoofdstuk 3 hebben mensen die op jonge leeftijd vaker hebben blootgestaan aan conflicten, meer kans om tijdens hun levensloop een lager inkomen te hebben. Overheden moeten bij de aanpak van voedseltekorten ook een veilige leef- en scholingsomgeving bieden voor jonge mensen in de leeftijd van 5 tot 19 jaar.