

2020

## Three essays on family economics

Jia Cao  
*Iowa State University*

Follow this and additional works at: <https://lib.dr.iastate.edu/etd>

---

### Recommended Citation

Cao, Jia, "Three essays on family economics" (2020). *Graduate Theses and Dissertations*. 17871.  
<https://lib.dr.iastate.edu/etd/17871>

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

**Three essays on family economics**

by

**Jia Cao**

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of  
**DOCTOR OF PHILOSOPHY**

Major: Economics

Program of Study Committee:  
Elizabeth Hoffman, Major Professor  
Joshua L. Rosenbloom  
Sunanda Roy  
John V. Winters  
Wendong Zhang

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2020

Copyright © Jia Cao, 2020. All rights reserved.

## DEDICATION

I would like to dedicate this dissertation to my major professor Elizabeth (Betsy) Hoffman without whose support I would not have been able to complete this work.

## TABLE OF CONTENTS

	Page
LIST OF TABLES . . . . .	v
LIST OF FIGURES . . . . .	vi
ACKNOWLEDGMENTS . . . . .	vii
ABSTRACT . . . . .	ix
CHAPTER 1. INTRODUCTION . . . . .	1
CHAPTER 2. SOCIAL MOBILITY ACROSS THREE GENERATIONS: EVIDENCE FROM CHINA . . . . .	3
2.1 Introduction . . . . .	3
2.2 Literature Review . . . . .	5
2.3 The Model . . . . .	7
2.3.1 The Case of AR(1) Process . . . . .	7
2.3.2 The Case of AR(2) Process . . . . .	9
2.4 The Data . . . . .	10
2.5 Empirical Analysis . . . . .	12
2.5.1 Basic Regression Equation . . . . .	12
2.5.2 Basic Regression Results . . . . .	12
2.5.3 Uncle and Aunt Effects . . . . .	13
2.5.4 Exogenous Shock . . . . .	14
2.5.5 Robustness . . . . .	16
2.6 Conclusions . . . . .	18
CHAPTER 3. MODELING PARENTAL LOVE: WARM GLOW OR PURE ALTRUISM? . . . . .	19
3.1 Introduction . . . . .	19
3.2 Literature Review . . . . .	20
3.3 The Model . . . . .	22
3.3.1 The Case of Warm Glow . . . . .	22
3.3.2 The Case of Pure Altruism . . . . .	29
3.4 An Example . . . . .	33
3.4.1 Case One: Unskilled Workers . . . . .	34
3.4.2 Case Two: Skilled Workers as Borrowers . . . . .	36
3.4.3 Case Three: Skilled Workers as Lenders . . . . .	38
3.5 Conclusions . . . . .	39

CHAPTER 4. FERTILITY, PURE ALTRUISM AND HYPERBOLIC DISCOUNTING . . .	41
4.1 Introduction . . . . .	41
4.2 The Model . . . . .	44
4.2.1 Model 1: A Naive Agent's Commitment Problem . . . . .	45
4.2.2 Model 2: A Sophisticated Agent's Time-Consistent Problem . . . . .	49
4.3 An Example . . . . .	54
4.3.1 Model 1: A Naive Agent's Commitment Problem . . . . .	55
4.3.2 Model 2: A Sophisticated Agent's Time-Consistent Problem . . . . .	55
4.3.3 Discussion . . . . .	56
4.4 Conclusions . . . . .	56
BIBLIOGRAPHY . . . . .	58
APPENDIX. PROOF OF CHAPTER 2 . . . . .	61
.1 The Proof of AR(1) Case . . . . .	61
.2 The Proof of AR(2) Case . . . . .	62

## LIST OF TABLES

		<b>Page</b>
Table 2.1	Sample Characteristics . . . . .	11
Table 2.2	Basic Regression Results . . . . .	12
Table 2.3	Regression Results with Uncle and Aunt Effects . . . . .	14
Table 2.4	Regression Results to Explore Same-Gender Preferences . . . . .	15
Table 2.5	Regression Results of Father's on Dummy Variable . . . . .	15
Table 2.6	IV Regression Results of Father's Schooling on Child's Schooling . . . . .	16
Table 2.7	Regression Results of Shock on Multigenerational Social Mobility . . . . .	16
Table 2.8	Regression Results of the Dad's Schooling on the Paternal Grandparents' Schooling . . . . .	17
Table 2.9	Regression Results of the Grandson's Schooling on the Dad's Schooling . . . . .	17
Table 2.10	Regression Results of the Grandson's Schooling on the Paternal Grandparents' Schooling . . . . .	17
Table 4.1	The Position of This Paper in Literature . . . . .	43

## LIST OF FIGURES

		<b>Page</b>
Figure 2.1	Education Comparison of A and B Families . . . . .	9
Figure 3.1	Warm Glow and Pure Altruism in Modeling Parental Love . . . . .	21
Figure 3.2	The Illustration of $\frac{\partial n_{wg}}{\partial \gamma} > 0$ . . . . .	27
Figure 3.3	The Illustration of $\frac{\partial n_{wg}}{\partial \epsilon} < 0$ . . . . .	28
Figure 3.4	The Dynamics of Bequests in GZ Model . . . . .	36
Figure 4.1	The Order of Decisions . . . . .	44

## ACKNOWLEDGMENTS

One afternoon in the summer of 2008, I came across a friend on campus when I was walking home after completing the last class of my high school, carrying several textbooks and study guides. The friend was in the same grade, so we walked together and talked about the imminent college life.

My friend asked, “What’s your plan for your college life?”

“I’m determined to think deeply on some issues.” I replied calmly.

“You are always thinking too much!” My friend responded sarcastically as if I were a nerd.

I was serious. Reading fired my enthusiasm for the world: *The Records about Insects* aroused my curiosity about nature; *Sophie’s World* provoked my interest in philosophy; *Magic Math* enabled me to appreciate the elegance of conic sections; *Mencius* brought me into ancient China. I was eager to know how the world works. During the next seven years obtaining my Bachelor’s Degree in Economics and Master’s Degree Finance, I spent a considerable amount of time auditing courses in philosophy, history, literature and arts. However, after auditing numerous courses and reading hundreds of books, I was still confused. What’s worse, my ambition shrank from understanding the world to exploring the laws of human society.

I decided to move on. So, I came to the United States to pursue my Doctoral Degree in economics. Obviously, my ambition shrank again from understanding human society to investigating how the economy works. During the process of taking courses, doing homework, preparing for qualifying exams and reading papers, it became evident that it is impossible to demystify the black box economy with a Doctoral Degree in Economics. Consequently, my ambition shrank from comprehending the economy to writing three essays on family economics.

It is discouraging to see how a huge ambition dwindled to a small goal in my twelve years of higher education. But that’s how we humans make progress in the pilgrimage to understanding the



world. At the beginning, we have a whole but blurred map of the world. We are attracted by its beauty and amazed by its grace. Then, we delve into a small part of the map we find interesting and discover more details with the magnifier we have been honing for years. Drawing a completely clear map of the world is an impossible task for one person; nevertheless, it is the most sacred duty for humans as a race.

I am greatly indebted to many people, who gave me unconditional support and invaluable advice in my journey to the small goal. First and foremost, I would like to express my deep gratitude to my major professor Elizabeth (Betsy) Hoffman. I enjoy talking with Betsy not because she is my major professor, but because she can grasp my untypical, and sometimes whimsical ideas due to her extraordinarily broad knowledge. Our conversations range from family dissolution to ancestor worship, from the U-shaped relationship between income and fertility to the time gap between giving nicknames and formal names for Chinese children in ancient times. Betsy holds two Doctoral Degrees, one in economics and the other in history. Disciplines cannot constrain her talents. During her splendid academic life, she served as Professor of Economics, History, Political Science, Psychology and Law in different universities. Betsy is the person of my acquaintance with the clearest map of the world. She is a shining example of what a scholar should be.

I would also like to thank my reference letter writers Minghao Li and Tao Xiong. Minghao is also my coauthor who teaches me how to do empirical research with great patience. Tao helped me a lot in the job market. Meanwhile, I enjoyed playing squash and tasting delicious cuisine with the two young scholars.

I would additionally like to thank my committee members for offering insightful comments. They are Joshua L. Rosenbloom, Sunanda Roy, John V. Winters, and Wendong Zhang. I also thank Otávio Bartalotti and Xiyang Liu for the help in my job market. I'm grateful to Joydeep Bhattacharya and Juan Carlos Cordoba for the help at the early stage.

At last but not least, I want to thank my parents for their unconditional love and support.

## ABSTRACT

This dissertation explores several issues on family economics. Chapter 2 uses a large and nationally representative sample to investigate the effects on children's education from their grandparents and extended families including uncles and aunts. Chapter 3 compares warm glow and pure altruism when they are both used to model parental love. The comparison focuses on their implications for fertility decisions and educational investments in the children. Chapter 4 studies the fertility and education investment choices of parents when they hold both pure altruism towards their children and hyperbolic discounting preferences of their own.

## CHAPTER 1. INTRODUCTION

This dissertation focuses on family economics. More specifically, I have explored how grandparents and the extended families influence the grandchildren's education, and how parents with different preferences make decisions about fertility and educational investments in their children. The behaviors of members within a family are so interconnected that a family is a better representative unit of a society than an individual. That is to say, aggregating the behaviors of families can provide us with a more accurate map of the economy than aggregating those of individuals.

Chapter 2 (joint with Minghao Li) uses empirical methods to explore the influence of grandparents and extended families on grandchildren's education. Our paper is the first to measure multi-generational social mobility with a nationally representative data set in China. First, we find that the education of grandfathers is positively correlated with the education of grandchildren. Second, we explore uncle and aunt effects, which are often neglected in previous research due to data limitations. Third, our paper shows that grandparent effects will decrease after the inclusion of uncle and aunt effects, suggesting the influence through uncles and aunts as an important channel for the grandparent effects. At last but not least, we claim that three-generational social mobility is not sensitive to the shock on the second generation's education.

Chapter 3 (joint with Elizabeth Hoffman) is purely theoretical. Warm glow and pure altruism are two widely used ways to model parental love in the literature. The two methods seem to be the same essentially since they share some important attributes. For example, parents will sacrifice part of their consumption to leave bequests to their children in both cases. Therefore, their differences have never been explored rigorously before. In this paper, we compare the two methods in the same backgrounds and explore their differences in the implications for fertility and bequests. We find that fertility and bequests have dramatically different solution forms in the two cases. More specifically, fertility has a much more complicated solution form than bequests under warm glow

preference. Fertility does not have a closed-form solution in this case. On the contrary, fertility is much simpler than bequests when parents show pure altruism towards their children, although both have closed-form solutions. Intuitions about the differences between the two similar modeling methods are discussed. Last, we use an example to show how the results will change after we replace warm glow in the paper with pure altruism.

Chapter 4 introduces behavioral economics into family economics. Exponential discounting is a widely accepted assumption in mainstream economics. Hyperbolic discounting, which presents different attributes from exponential discounting, receives evidence from the lab and field and therefore is applied to many branches of economics. In this paper, I investigate the effects of hyperbolic discounting on demographic economics where parental love is modeled by pure altruism. I find that hyperbolic discounting will reduce fertility and increase educational investments in both the commitment problem and time-consistent problem. The magnitude of the changes in fertility and educational investments in the time-consistent problem is bigger than that in the commitment problem. Intuitions about the differences are discussed.

From the discussion above, we can see that this dissertation delves into the interactions among members who belong to different generations. One important reason is that I have strong interests in the long-run inequality of education and fertility of a society.

## CHAPTER 2. SOCIAL MOBILITY ACROSS THREE GENERATIONS: EVIDENCE FROM CHINA

Jia Cao and Minghao Li

### 2.1 Introduction

The majority of the literature on social mobility focuses on two adjacent generations, namely, the correlation between parents and children [Becker and Tomes (1979); Becker et al. (2018)]. One implicit assumption of such research is that the dynamics of social mobility can be modeled by a first-order autoregressive [AR(1)] process. Within this framework, the influence of the first generation's social status (such as income and education) on their descendants will decline geometrically across the future generations. Consequently, grandparents will have no direct effects on their grandchildren; the only grandparent effect is indirect and mediated through the parents.

However, there are plausible reasons that the AR(1) assumption is not valid. First, genetic information from the grandparents may be repressed in parents and manifested again in children. Second, grandparents may also exert direct economical, educational, and cultural influence on the children. The effects are especially relevant for certain demographic groups in countries that have closely-tied extended families. If these direct grandparent effects do exist, the results of two-generation studies under AR(1) assumption cannot be extended to multiple generations.

Several papers provide empirical evidence showing the existence of higher-order autoregressive processes (i.e. direct grandparent effects). However, due to data availability, most of the research is on developed countries, especially Scandinavian countries [Lindahl et al. (2015); Møllegaard and Jæger (2015)]. Research on multigenerational social mobility in developing countries is limited.

The multigenerational social mobility in China is of special importance for this literature. First, China is the world's largest developing country and the evidence from China is indispensable in the

debate on grandparent effects. Second, China's historical and institutional backgrounds provide unique opportunities to understand the mechanisms of multigenerational social mobility. However, despite the importance of this topic, there is a dearth of multigenerational mobility research on China. To our best knowledge, the only available research is by Zeng and Xie (2014), which uses a relatively small rural sample with incomplete educational attainment information for the grandchildren generation.

This paper is the first to study social mobility across three generations in China using a large, nationally representative sample. Data used in this study is from multiple waves (2011, 2013, 2014, 2015) of the China Health and Retirement Survey. We combined all waves of the survey to create a comprehensive dataset with rich information on the three generations.

The explicit grandparent effects are more likely to happen in China than in Western countries because, traditionally, big families with multiple generations living together are considered more successful than small families with only two generations. Although small families are pretty common at present, it is still a usual practice for the grandparents to take care of the children when the parents are busy with work. Grandparents usually consider it as a kind of pleasure instead of a burden to take care of their grandchildren. Using nationally representative datasets, this paper measures three-generational social mobility directly in China for the first time.

We find strong and positive direct grandparent effects. Besides contributing to the long-debated grandparent effects, this study sheds light on possible mechanisms that such effects take place. We also find significantly positive uncle and aunt effects from both paternal and maternal sides. Previous studies often overlook uncle/aunt effects due to data limitations. The inclusion of uncle/aunt effects reduces grandparent effects, suggesting the influence through uncles and aunts as an important channel for the grandparent effects. Last but not least, we attempt to quantify the grandparent effects through parents' formal education. In China, the education of some cohorts in the parent generation was reduced by the Cultural Revolution. Exploiting this unique historical background, we find that parents' education has a causal and positive relationship with children's education. Surprisingly, three-generational social mobility is not sensitive to the shocks on parents' education.

Our paper can contribute to the debate about grandparent effects in mobility research: are grandparent effects on grandchildren entirely mediated through parents, or do grandparents have independent and direct effects through other channels? Previous literature has provided various methods, which can serve as the departure point of this study, to answer these questions empirically,

The paper is organized as follows. Section 2.2 provides a literature review on the empirical studies about multigenerational social mobility. Section 2.3 presents the theoretical framework of intergenerational mobility. Section 2.4 describes the data we use. Section 2.5 summarizes the empirical results. Section 2.6 concludes.

## 2.2 Literature Review

Most of the existing multigenerational mobility studies are conducted in developed countries due to data availability. Most of them find no significant grandparent effect when parents' effects are controlled. In an early study in the U.S., Hodge (1966) finds that besides the indirect effects through parents, the occupation of grandparents has no direct effect on the occupation of the children. This conclusion is echoed in later years. For example, using data from the Wisconsin Longitudinal Study, Warren and Hauser (1997) find that grandparents' social status has no statistically significant impact on children's social status once parents' social status is controlled. Using a sample of twins, Behrman and Taubman (1985) find that grandparents' schooling has no significant effect on children's schooling. Erola and Moisio (2007) construct 57,585 three-generation lineages in Finland from 1950-2000 and find that the grandchildren's social class is almost conditionally independent from the grandparents' social class after parents' social class is controlled. Lucas and Kerr (2013) and Peters (1992) also find no significant grandparent effect.

However, there are still studies finding statistically significant grandparent effects with the data sets of Western countries. Chan and Boliver (2013) use data from three British birth cohort studies and find a statistically significant grandparent effect on their grandchildren's relative mobility patterns, after parents' social class is controlled. Lindahl et al. (2015) find strong evidence that grandparents' education and income directly affect children's income in a Swedish four-generation

study. Interestingly, Chan and Boliver (2014) argue that the main conclusion of “almost conditional independence” from Erola and Moio (2007) is not supported by the results in the article. They demonstrate that the grandparent effect in social mobility in Finland is not only statistically significant, but is also of substantial importance.

Researchers also try to identify the possible mechanism of the grandparent effects. Møllegaard and Jæger (2015) analyze data from Denmark. They find that it is grandparents’ cultural capital, instead of economic and social capital that plays a positive role when grandchildren choose the academic track in upper secondary education. The results of Møllegaard and Jæger (2015) show that the possible mechanism for the grandparent effects on grandchildren’s education success is carried out through the transmission of non-economic resources. The results, as the authors suggest, may be valid only in wealthy societies such as Scandinavian countries.

For the case of China, most empirical studies on multigenerational mobility focus on specific groups. Mare and Song (2014) investigate two datasets: one is genealogical data from the Qing Dynasty Imperial Lineage which contains 12 generations of Qing emperors and their relatives from the 17th to the 20th centuries; the other is population registry data which contains 10 generations of male peasants in the northeastern province of Liaoning from the mid-18th to the early 20th centuries. The former dataset contains individuals at the top while the latter contains individuals at the bottom of the society. Despite the huge differences between the two datasets, both find that men’s social positions are affected not only by the positions of their fathers but also of their grandfathers and great-grandfathers.

Shiue (2016) uses data covering information on seven lineages of nearly 10,000 men to explore social mobility. The author finds that educational inequality is closely related to changes in mobility over time. As for the grandparent effects, this paper also finds that the lineal impact of grandfathers and older generations is overshadowed by non-lineal interactions coming from higher status men in the same generation as the father. The results are consistent with the historical phenomenon that extended family members usually lived together and had strong communal ties among each other



in history. Note that all of the individuals in Shiue's data lived in one county of Anhui province, located in south of China.

Zeng and Xie (2014) is the only multigenerational mobility study for modern China with a relatively large sample. They used data from the 2002 Chinese Household Income Project (CHIP) to study the effect of grandparents' education on children's education in rural China. Because the third generation in the data was still in school, the authors could only observe the final education outcome for those students who dropped out. They have to use logit models to estimate the probabilities of dropping out from schools. Consequently, the authors use only rural data, because dropout rates are very low in urban China. Nonetheless, they not only find a significant grandparent effect after controlling for parents' education and social status, but they also show that the effect exists only if the grandparents live with the children. That means the direct grandparent effect is less likely to work through genetic inheritance, and more likely to work through personal interactions. It also means the intensity of the interaction matters, since living non co-resident grandparents (who presumably also have a certain level of interaction with the children) have no effect on grandchildren's education.

## 2.3 The Model

### 2.3.1 The Case of AR(1) Process

The model here follows Solon (2014) with one revision. Family  $i$  contains one parent born at time  $t-1$  and one child born at time  $t$ . The parent's income  $y_{i,t-1}$  is used for her own consumption  $C_{i,t-1}$  and investment  $I_{i,t-1}$  in the child's education. The budget constraint is

$$y_{i,t-1} = C_{i,t-1} + I_{i,t-1} \quad (2.1)$$

The child's schooling  $S_{i,t}$  is a function of education investment  $I_{i,t-1}$  and endowment from the parent  $e_{i,t}$

$$S_{i,t} = \theta \log I_{i,t-1} + e_{i,t} \quad (2.2)$$

where  $\theta$  is assumed to be positive.

In Solon (2014), the equation above describes the formation of human capital. I change it into schooling because schooling can be measured more accurately than human capital. This is the only revision I made. The endowment follows an AR(1) process

$$e_{it} = \delta + \lambda e_{i,t-1} + v_{it} \quad (2.3)$$

where  $v_{it}$  is an error term which is not correlated with endowment and  $0 < \lambda < 1$ .

The child's income is a function of her schooling

$$\log y_{it} = u + p S_{it} \quad (2.4)$$

where  $p$  is assumed to be positive.

The parent's utility function is

$$U_i = (1 - \alpha) \log C_{i,t-1} + \alpha \log y_{it} \quad (2.5)$$

$\alpha$  is the altruism parameter because the parent cares about her child's welfare which is represented by a function of the child's income.

Solving the problem, we have

$$S_{it} = \Omega_1 + (\lambda + \theta p) S_{i,t-1} - \lambda \theta p S_{i,t-2} + v_{it} \quad (2.6)$$

where  $\Omega_1 = \delta + (1 - \lambda) \theta \left[ u + \frac{\alpha p}{1 - \alpha(1 - \theta p)} \right]$ . The proof will be given in the Appendix.

We can see that  $0 < \lambda \theta p < \lambda + \theta p$  from the assumptions of the three parameters. If the assumption of AR(1) process is true, the model predicts a negative coefficient of grandparental schooling and the magnitude of the coefficient is smaller than that of the parental schooling. The prediction of negative sign seems counter-intuitive at the first glance. It is easy to understand the prediction if we know that the influence of grandparental schooling is calculated after the influence of parental schooling is controlled. It is illustrated in figure 2.1. Suppose children A and B's parents have the same schooling. However, A's grandparent has more schooling than B's. More schooling implies higher income, and higher income implies more investment. We can infer that A's

parent probably has lower endowment (i.e. inherited socially productive traits) because with more investment in her education, she achieved the same schooling as B's parent. Since A only gets her endowment from her parent, the model will predict that she probably has a lower endowment than B, hence less schooling.

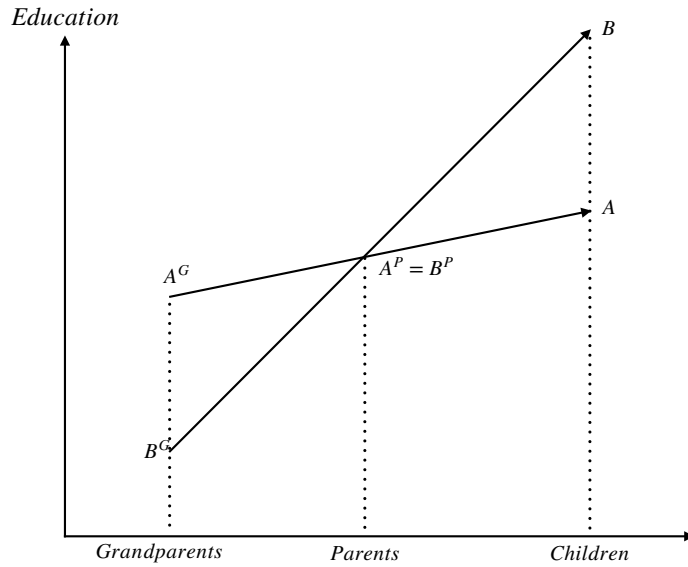


Figure 2.1 Education Comparison of A and B Families

### 2.3.2 The Case of AR(2) Process

We can model the possibility of grandparent effects with an AR(2) process of the endowments

$$e_{it} = \delta + \lambda_1 e_{i,t-1} + \lambda_2 e_{i,t-2} + v_{it} \quad (2.7)$$

where  $\lambda_2$  represents grandparent effects. When it equals zero, it degrades to an AR(1) process.

Solving the same problem as the case of AR(1) process, we have the following equation

$$S_{it} = \Omega_2 + (\lambda_1 + \theta p) S_{i,t-1} + (\lambda_2 - \lambda_1 \theta p) S_{i,t-2} - \lambda_2 \theta p S_{i,t-3} + v_{it} \quad (2.8)$$

where  $\Omega_2 = \delta + (1 - \lambda_1 - \lambda_2)\theta \left[ u + \log \frac{\alpha p}{1 - \alpha(1 - \theta p)} \right]$ . The proof will be given in the Appendix.

The sign of the coefficient of grandparental schooling is uncertain. However, if empirical evidence shows a positive coefficient of grandparental schooling, we can conclude that AR(2) process is a better assumption than AR(1) process about the dynamics of the endowments since  $\lambda_2 - \lambda_1\theta p > 0 \Rightarrow \lambda_2 > 0$ .

## 2.4 The Data

Data used in this study are from the China Health and Retirement Longitudinal Study (CHARLS). It is a nationally panel survey targeting the middle-aged and senior populations carried out from 2011. The following three waves of data were collected in 2013, 2014 and 2015 respectively. In this paper, we combined all waves of the survey to create a comprehensive dataset with rich information on the three generations. In the survey, a household with at least one member 45 years old or above is randomly selected, and this member becomes the main respondent. Information is collected on main respondents and their spouses, together with the parents on both sides and all children of the couple regardless of where they live. Information on other family members, such as the grandchildren of the main respondents, are available if they live together with the main respondents. In the first wave of data collected in 2011 and 2012, 17708 individuals who are from 10257 households and 150 counties successfully responded to the survey. This random sample is large enough to represent the whole aged population. This dataset contains detailed educational attainment information for three generations regardless of whether they live together and the fourth generation if they live in the same household. Using the information on the first three generations, nationally representative three-generation mobility can be measured for the first time for China.

Income and occupation are also used in the literature to represent social status in the research on social mobility. It may not be a problem in developed countries where income increases slowly and occupation seldom changes. However, it can be a serious problem in developing countries. China has experienced rapid economic growth for the previous four decades, when both individuals' incomes and occupations changed frequently. Another issue on income is that Chinese like to make money

in their leisure time. The informal income is quite unpredictable and suffers huge measurement errors. Education level is the variable which can be measured much more accurately than income level and occupation in China. Shiue (2016) also finds that educational inequality is closely related to changes in mobility over time. So, we choose educational attainment to represent social status for our research.

Table 2.1 Sample Characteristics

Variables	Observations	Mean	Standard errors	Min	Max
Child's Schooling	20,237	8.58	4.33	0	22
Child's Age	20,237	34.18	9.32	10	72
Child's Gender(Boy=0, Girl=1)	20,237	0.48	0.50	0	1
Area Type(Rural=0, Urban=1)	20,237	0.33	0.47	0	1
Father's Age	20,237	63.20	10.12	35	94
Mother's Age	20,237	60.93	9.77	32	93
Father's Schooling	20,237	7.06	3.55	0	22
Mother's Schooling	20,237	4.57	3.97	0	22
Paternal Grandfather's Schooling	20,237	2.27	3.30	0	16
Paternal Grandmother's Schooling	20,237	0.50	1.75	0	22
Maternal Grandfather's Schooling	20,237	2.32	3.38	0	16
Maternal Grandmother's Schooling	20,237	0.53	1.81	0	22

As can be seen from the Table 2.1, the average age of the third generation is 34, which means that most of them have finished school. We dropped the small sample when grandchildren are still in school. So, we can use schooling directly instead of estimating the probabilities of dropouts like Zeng and Xie (2014). The observations of paternal grandparents are less than maternal grandparents. The possible reason is that there are more mother respondents than father respondents and that mothers are not able to answer the schooling of their parents-in-law if their husbands are not alive. Our sample is also larger and more representative. It contains more than 20,000 observations from the whole country. Zeng and Xie (2014) 's sample contains only rural data and has only 833 households with the information of three generations.

## 2.5 Empirical Analysis

### 2.5.1 Basic Regression Equation

As illustrated in Section 3, our basic regression equation is

$$S_c = \alpha + \beta_1 S_f + \beta_2 S_m + \beta_3 S_{pf} + \beta_4 S_{pm} + \beta_5 S_{mf} + \beta_6 S_{mm} + \gamma X + \epsilon \quad (2.9)$$

$S_c$  is the schooling of the child.  $S_f$  and  $S_m$  are the schooling of the child's father and mother.  $S_{pf}$  and  $S_{pm}$  are the schooling of the child's paternal grandfather and grandmother.  $S_{mf}$  and  $S_{mm}$  are the schooling of the child's maternal grandfather and grandmother.  $X$  contains control variables which are the area type, province, parents' ages, child's age and gender.

### 2.5.2 Basic Regression Results

The regression results are shown in table 2.2. In order to save space, province effects are not shown in the following regression results.

Table 2.2 Basic Regression Results

	Coefficient	Standard Errors	p-value
Constant	7.243***	0.238	0.000
Age	-0.066***	0.006	0.000
Girl	-0.670***	0.052	0.000
Urban	1.452***	0.058	0.000
Father's Age	-0.001	0.007	0.869
Mother's Age	0.001	0.008	0.882
Father' Schooling	0.285***	0.008	0.000
Mother' Schooling	0.203***	0.008	0.000
Paternal Grandfather's Schooling	0.064***	0.009	0.000
Paternal Grandmother's Schooling	0.025	0.016	0.127
Maternal Grandfather's Schooling	0.065***	0.008	0.000
Maternal Grandmother's Schooling	0.026	0.016	0.105
Observations		20,237	
Adjusted R-squared		0.268	

We can see that both paternal and maternal grandparents' schooling have positive effects on children's schooling. Paternal grandfather effects and maternal grandfather effects are significant

and have the same magnitude. Grandmother effects are not significant. The possible reason is that only a small portion of grandmothers received very limited formal education. The negative sign of the girl dummy variable shows that girls receive less schooling than boys. The positive sign of the urban dummy variable indicates that children in urban areas stay longer in school than children from rural areas. The negative sign of age tells us that educational levels increases with time.

### 2.5.3 Uncle and Aunt Effects

The One Child Policy in China was introduced in 1979, long after the majority of fathers and mothers in our dataset were born. In fact, fathers have an average of 3.71 siblings, while mothers have 3.86 siblings. So, we also explore uncle and aunt effects, which were often neglected in previous research due to data limitations.

Four models are used to explore uncle and aunt effects. In model 1, we use the highest education of parents' siblings. We can see that uncle effects are stronger than aunt effects. On the paternal side, uncle effects are 34% higher than aunt effects. On the maternal side, the gap increases to 49%. If we ignore the gender and estimate the effects of the highest education of parents' siblings, the magnitudes are similar on both sides. The results are shown in Table 2.3. In models 3 and 4, we use the average education of parents' siblings to test uncle and aunt effects. The results follow similar patterns.

The regression results also show that grandfather effects decline after the inclusion of uncle and aunt effects. The results suggest that the influence through uncles and aunts is an important channel for the grandparent effects. Grandmother effects are still not significant.

Why are uncle effects so much stronger than aunt effects on both paternal and maternal sides? Do both uncles and aunts show same-gender preferences: do uncles prefer nephews and aunts prefer nieces? In order to solve these puzzles, we will do regressions for male and female grandchildren separately.

Table 2.3 Regression Results with Uncle and Aunt Effects

	Basic	Model 1	Model 2	Model 3	Model 4
The Highest Schooling of Father's Brothers		0.087***			
The Highest Schooling of Father's Sisters		0.065***			
The Highest Schooling of Mother's Brothers		0.106***			
The Highest Schooling of Mother's Sisters		0.071***			
The Highest Schooling of Father's Siblings			0.104***		
The Highest Schooling of Mother's Siblings			0.127***		
The Average Schooling of Father's Brothers				0.097***	
The Average Schooling of Father's Sisters				0.072***	
The Average Schooling of Mother's Brothers				0.120***	
The Average Schooling of Mother's Sisters				0.074***	
The Average Schooling of Father's Siblings					0.169***
The Average Schooling of Mother's Siblings					0.194***
Father's Schooling	0.284***	0.231***	0.238***	0.231***	0.233***
Mother's Schooling	0.185***	0.131***	0.146***	0.127***	0.124***
Paternal Grandfather's Schooling	0.085***	0.056***	0.062***	0.053***	0.051***
Paternal Grandmother's Schooling	0.016	-0.007	0.001	-0.017	-0.020
Maternal Grandfather's Schooling	0.073***	0.038***	0.045***	0.035***	0.036***
Maternal Grandmother's Schooling	-0.001	-0.015	-0.010	-0.023	-0.029
Observations	9,263	9,263	9,263	9,263	9,263
Adjusted R-squared	0.266	0.294	0.290	0.293	0.291

Whether we use the highest or the average schooling, we find strong evidence for same-gender preferences from uncles towards their nephews. However, the regression results above do not show that aunts have same-gender preferences towards their nieces.

#### 2.5.4 Exogenous Shock

During the Cultural Revolution (1966-1976), some schools were closed for more than one month. The Cultural Revolution affect a cohort of parents when they were in school. We use this event as an exogenous shock on the parents' education. We focus on the father's education at present and regress it on the dummy variable whether the father's school was closed during the Cultural Revolution. Control variables include father's age, area type, province and paternal grandparents'



Table 2.4 Regression Results to Explore Same-Gender Preferences

	Male Grandchildren	Female Grandchildren
The Highest Schooling of Father's Brothers	0.084***	0.089***
The Highest Schooling of Father's Sisters	0.039***	0.090***
The Highest Schooling of Mother's Brothers	0.103***	0.110***
The Highest Schooling of Mother's Sisters	0.067***	0.075***
The Average Schooling of Father's Brothers	0.105***	0.090***
The Average Schooling of Father's Sisters	0.039**	0.105***
The Average Schooling of Mother's Brothers	0.111***	0.131***
The Average Schooling of Mother's Sisters	0.071***	0.075***

education. The regression results are shown as following. We can see that the exogenous shock has a negative and significant effect on father's education.

Table 2.5 Regression Results of Father's on Dummy Variable

	Coefficient	Standard Errors	p-value
Constant	14.291***	1.295	0.000
Age	-0.061***	0.012	0.000
Urban	0.526***	0.082	0.000
Grandfather's Schooling	0.065***	0.012	0.000
Grandmother's Schooling	0.076***	0.023	0.001
School Close	-0.474***	0.083	0.000
Observations		4,679	
Adjusted R-squared		0.088	

Now we use the exogenous shock as an instrument variable for father's education and estimate the causal effect of father's education on children's education. The regression results are given in Table 2.6.

At last, we use the exogenous shock to explore the three-generational social mobility. The regression equation can be expressed as

$$S_c = \alpha + \beta_1 Close + \beta_2 S_{pg} + \beta_3 Close * S_{pg} + \gamma X + \epsilon$$

where *Close* is the dummy variable indicating father's school was closed.  $S_{pg}$  is the sum and paternal grandfather and mother's schooling. We add them up because both will influence father's schooling. Control variables include child's age and gender, mother's age and schooling, area type,

Table 2.6 IV Regression Results of Father's Schooling on Child's Schooling

	Coefficient	Standard Errors	p-value
Constant	1.032	3.989	0.796
Child's Age	-0.089***	0.017	0.000
Girl	-0.392***	0.136	0.004
Urban	1.304***	0.184	0.000
Father's Schooling	1.378***	0.308	0.000
Mother's Schooling	0.010	0.050	0.847
Father's Age	0.041	0.037	0.260
Mother's Age	-0.043	0.027	0.111
Observations		4,434	

province and maternal grandparents' schooling. Control variables don't include father's age and schooling in that they are highly correlated with the dummy variable. Regression results are given in Table 2.7. The results show that the interactive term is not significant.

Table 2.7 Regression Results of Shock on Multigenerational Social Mobility

	Coefficient	Standard Errors	p-value
Constant	14.511***	1.786	0.000
Child's Age	-0.112***	0.015	0.000
Girl	-0.402***	0.117	0.001
Urban	1.664***	0.135	0.000
Mother's Schooling	0.188***	0.016	0.000
Mother's Age	0.006	0.017	0.720
Maternal Grandparents' Schooling	0.057***	0.013	0.000
Close	-0.705***	0.156	0.000
Paternal Grandparents' Schooling	0.031***	0.050	0.538
Close*Paternal Grandparents' Schooling	0.029	0.029	0.311
Observations		4,434	

### 2.5.5 Robustness

The method requires three regressions. First, we regress the dad's schooling on the paternal grandparents' schooling and get a coefficient  $\beta_1$ . Then, we regress the grandson's schooling on the dad's schooling and get another coefficient  $\beta_2$ . At last, we regress the grandson's schooling on the paternal grandparents' schooling and get coefficient  $\beta_3$ . If grandparent effects do not exist,

$\beta_3$  should be close to the product of  $\beta_1$  and  $\beta_2$ . We constrain our sample with positive paternal grandparents' schooling and grandsons who have finished their educations. Our regression results shows that

$$\beta_3 = 0.153 \gg 0.151 * 0.374 = \beta_1\beta_2$$

which suggests that a (paternal) grandparent effect exists.

Table 2.8 Regression Results of the Dad's Schooling on the Paternal Grandparents' Schooling

	Coefficient	Standard Errors	p-value
Constant	12.97***	0.136	0.000
Dad's Age	-0.107***	0.002	0.000
Urban	1.183***	0.045	0.000
Paternal Grandparents' Schooling	0.151***	0.005	0.000
Observations		23,917	
Adjusted R-squared		0.179	

Table 2.9 Regression Results of the Grandson's Schooling on the Dad's Schooling

	Coefficient	Standard Errors	p-value
Constant	8.402***	0.157	0.000
Urban	2.515***	0.072	0.000
Grandson's Age	-0.110***	0.003	0.000
Dad's Schooling	0.374***	0.009	0.000
Observations		13,151	
Adjusted R-squared		0.293	

Table 2.10 Regression Results of the Grandson's Schooling on the Paternal Grandparents' Schooling

	Coefficient	Standard Errors	p-value
Constant	11.533***	0.154	0.000
Child's Age	-0.139***	0.004	0.000
Urban	2.816***	0.083	0.000
Paternal Grandparents' Schooling	0.153***	0.009	0.000
Observations		10,757	
Adjusted R-squared		0.212	

## 2.6 Conclusions

Social mobility is a topic which has been discussed heatedly, both in academia and in public, for a long time. Most of the previous research on social mobility focuses on two generations due to data availability. Multigenerational social mobility has gained increasing attention during the last decade. However, most datasets are limited to developed countries, especially Scandinavian countries. Research on multigenerational social mobility in developing countries is rare. Using years of schooling as the main social status indicator, we use a large and nationally representative dataset to estimate three-generational social mobility in China for the first time.

We have several findings. The first finding is that independent grandparent effects are positive with statistical significance. The magnitude of grandparent effects are about 20%-25% of parent effects. This finding contributes to the current literature on multigenerational social mobility with a better dataset. The second finding is that paternal grandparent effects are stronger than maternal grandparent effects. We also show that uncle and aunt effects are positive and grandparent effects decline dramatically after the inclusion of uncle and aunt effects. Last, we use uncle and aunt effects to explore the topic of same-gender preferences and find that both paternal and maternal aunts show stronger preferences for same-gender nieces than nephews.

The positive grandparent effects imply that second-order autoregressive processes can model the dynamics of endowments better than first-order autoregressive processes. But is the second-order autoregressive process the true model? It depends on the existence of great-grandparent effects which needs an analysis of four-generation social mobility. This will be future work.

## CHAPTER 3. MODELING PARENTAL LOVE: WARM GLOW OR PURE ALTRUISM?

Jia Cao and Elizabeth Hoffman

### 3.1 Introduction

Self-interest is the main driving force of human behavior, but it is not the only one. Altruism is widely observed, and thus commonly modeled in many fields including economics [Becker (1976); Simon (1993); Fehr and Schmidt (2006)]. Parental love is probably the kind of altruism to which people pay the most attention. Two kinds of methods are usually used to model parental love in the literature. One kind is called pure altruism. A behavior is motivated by pure altruism if its only motive is to improve the welfare of others. In the literature of modeling parental love, pure altruism implies that parents care about the lifetime utility of their children [Becker (1974); Becker (1976); Becker and Barro (1986)]. The implication of pure altruism in Becker's model is that the higher the children's lifetime utility is, the higher their parents' utility will be. The alternative method used to model parental love is warm glow. An agent with warm glow preferences receives utility from the behavior of giving, instead of the welfare impact on receivers. Although giving will probably improve the welfare of others, the giver does not care about the consequences of her giving. The mixture of pure altruism and warm glow is called impure altruism which is originally proposed to explain charity and public donations [Andreoni (1989, 1990)].

The mechanisms of warm glow and pure altruism in modeling parental love are illustrated in Figure 3.1. In the case of warm glow, parents give bequests to their children. The more bequests the parents leave, the happier they will feel. In the case of pure altruism, parents also give bequests to their children, and the purpose of leaving bequests is to improve the children's welfare. The higher the children's lifetime welfare is, the happier the parents will feel. The same thing between

warm glow and pure altruism in modeling parental love is that parents will leave bequests to their children in both cases. The only difference is that the parents with warm glow preferences care about how many bequests to leave, while the parents with pure altruism care about the positive effect of the bequests on their children's welfare.

The difference between warm glow and pure altruism seems not essential since the parents have the same behavior, leaving bequests to their children, in both cases. Also, warm glow is much easier to handle than pure altruism since the children's utility will not appear in warm glow models. Consequently, warm glow is usually considered as a proxy of pure altruism to model parental love. The purpose of this paper, however, is to explore whether the difference between warm glow and pure altruism in modeling parental love is essential or not. The only difference between the parents with warm glow and pure altruism is in motivation, that is to say, parents love their children from the perspective of their children or themselves. The difference in motivation seems to be subtle and is always ignored in the literature. We want to explore whether the differences are subtle enough to be ignored. As a result, we think it necessary to compare warm glow and pure altruism systematically and comprehensively in the literature of modeling parental love. As far as we know, our paper is the first to fill in this gap.

The paper is organized as follows. Section 3.2 shows more related literature. Section 3.3 presents the model in detail. We first consider the cases of warm glow and pure altruism in a general way, which means that the models will not contain specific forms of utility function and altruism parameters, then we consider the two cases with specific functions to solve the models completely. Section 3.4 uses an example to show the differences of the two kinds of modeling methods. Section 3.5 concludes.

### 3.2 Literature Review

The theory of warm glow is firstly proposed by Andreoni (1989) to explain charitable giving. Andreoni (1990) extends the warm glow theory to the literature on public donations. In the two seminal papers, impure altruism is created to mean the mixture of pure altruism and warm glow.

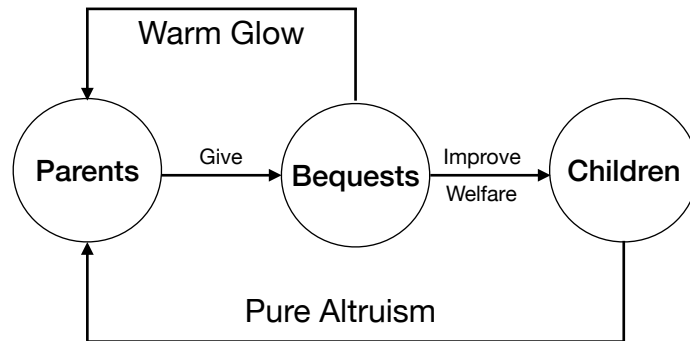


Figure 3.1 Warm Glow and Pure Altruism in Modeling Parental Love

The theory of altruism has a long history. In economics, pure altruism is introduced by Becker (1974) to model parental love.

The different implications of warm glow and pure altruism for charity giving and public goods donations are studied both theoretically and empirically [Andreoni (1989, 1990); Harbaugh et al. (2007); Ottoni-Wilhelm et al. (2017)]. If warm glow plays no role in donating to public goods, namely pure altruism is the only motivation to donate, the neutrality hypothesis which states that charity is subject to complete crowding out should be valid. However, the neutrality hypothesis is rejected empirically, and the crowd-out effect is found to decrease with output [Ottoni-Wilhelm et al. (2017)]. Crumpler and Grossman (2008) also uses an experimental method to show the existence of warm glow in charity giving. In other words, donations are probably motivated by both warm glow and pure altruism, although pure altruism is the dominating motive.

Different implications for the pro-environmental behavior of pure and impure altruism are also noticed. Hartmann et al. (2008) use two online surveys to show that warm glow has a stronger influence on pro-environmental intentions than pure altruism. One possible explanation for the different importance of warm glow and pure altruism in donation and pro-environmental behaviors is that donations can improve other people's welfare directly, while pro-environmental behavior can

only improve other people's welfare indirectly. Pure altruism should be strong when we can easily see the improvement of others from our behaviors easily.

In the literature on modeling parental love, however, implications of the two modeling methods have not been studied before. Usually one kind of the two methods is assumed to represent parental love without comparison to the other kind. Warm glow is used more often due to its relatively simple tractability. For example, in the literature of fertility under hyperbolic discounting, warm glow preferences is adopted to represent parental love [Wrede (2011); Wigniolle (2013)](Wrede, 2011; Wigniolle, 2013). In macroeconomics, Galor and Zeira (1993) uses warm glow as parental love to explore the role of wealth distribution through investment in human capital. In Section 3.5, we will use this paper as an example to show the differences between warm glow and pure altruism.

### 3.3 The Model

Every individual is assumed to live two periods: youth and adulthood. Bequests and fertility decisions are made in adulthood. Consistent with Becker (1974), we assume that an individual will have children without marriage. A family, therefore, consists of one parent and her children. The number of children is also allowed to be continuous. Wages in two periods are taken as given. A perfect financial market is available. Agents can save and borrow as much as they want at the same interest rates. Bequest taxes are paid by the parent.

#### 3.3.1 The Case of Warm Glow

An agent tries to maximize

$$U = u(c_1) + \delta u(c_2) + \delta \Phi(n)u(b')$$

subject to

$$(1 - \tau_1)w_1 + b + \frac{(1 - \tau_2)w_2(1 - nx) + ns - nb'/(1 - \tau)}{R} = c_1 + \frac{c_2}{R}$$

where  $\tau, \tau_1, \tau_2$  are the tax rates on bequests and wages in period one and two, respectively.

Taxes are introduced to explore the different effects of warm glow and pure altruism on bequests



and fertility. If the parent wants each of her children to receive a bequest  $b'$ , the parent needs to leave total bequests of  $nb'/(1 - \tau)$ . The only cost of raising children is time. One child needs  $x$  percentage of the parent's total time.  $s$  is the subsidy for one child.  $\delta$  is the time discounting rate and  $R$  is the interest rate. The altruism parameter is  $\Phi(n)$  which is the same as the case of pure altruism. In order to compare warm glow and pure altruism, we need to keep the remaining settings the same. Moreover, without altruism parameter, the only benefit of having children is the subsidy. This is not a reasonable case. The parent gains utility from giving bequests.

Let  $\lambda_1$  be the Lagrangian multiplier associated with the budget constraint. FOCs are

$$c_1 : u'(c_1) = \lambda_1 \quad (3.1)$$

$$c_2 : \delta u'(c_2) = \frac{\lambda_1}{R} \quad (3.2)$$

$$n : \delta \Phi'(n) u(b') = \lambda_1 \left( \frac{(1 - \tau_2) w_2 x - s + b'/(1 - \tau)}{R} \right) \quad (3.3)$$

$$b' : \delta \Phi(n) u'(b') = \frac{\lambda_1 n}{R(1 - \tau)} \quad (3.4)$$

We need some assumptions before going on.

**Assumption 1.** *The functions  $u(\cdot)$  and  $\Phi(\cdot)$  show constant elasticities of substitution. Specifically,  $\varepsilon(u) \equiv \frac{cu'(c)}{u(c)} = 1 - \gamma$ ,  $\varepsilon(\Phi) \equiv \frac{n\Phi'(n)}{\Phi(n)} = 1 - \epsilon$ . Moreover,  $0 < \tau, \tau_1, \tau_2, \gamma, \epsilon < 1$ .*

Generally speaking, this assumption states that tax rates, the elasticities of consumption and altruism function are positive but less than one.

**Assumption 2.**  $1 - \epsilon > 1 - \gamma$ .

This assumption states that the elasticity of the altruism function is greater than the elasticity of the utility function. As we will see shortly, this assumption is needed to guarantee positive bequests.

**Assumption 3.**  $(1 - \tau_2)w_2x > s$ .

This assumption states that the costs of raising children should be greater than the subsidies. Without this assumption, there will be corner solutions for fertility since agents will spend all the time having children instead of working. In fact, fertility will be  $\frac{1}{x}$  if this assumption is violated, a trivial situation.

Combining (3) and (4), we have

$$\frac{\Phi'(n)u(b')}{\Phi(n)u'(b')} = \frac{(1 - \tau_2)(1 - \tau)w_2x - (1 - \tau)s + b'}{n}$$

We can rewrite it as

$$\left(\frac{n\Phi'(n)}{\Phi(n)}\right) \left(\frac{u(b')}{b'u'(b')}\right) = \frac{1 - \epsilon}{1 - \gamma} = \frac{(1 - \tau_2)(1 - \tau)w_2x - (1 - \tau)s + b'}{b'}$$

Rearranging the equation above, we have the solution for the bequest.

**Proposition 1.** *Bequest under warm glow preferences can be determined by*

$$b'_{wg} = (1 - \tau) \frac{1 - \gamma}{\gamma - \epsilon} [(1 - \tau_2)w_2x - s] \quad (3.5)$$

Assumptions 1-3 guarantee that the bequest above is positive. Several corollaries arise from Proposition 1.

**Corollary 1.** *Under warm glow preferences, the bequests given are independent of both the bequests received and wages in the first period.*

Corollary 1 shows that bequests lack persistence and thus do not have accumulative effects across generations. Under warm glow preferences, bequest giving is similar to consumption, but with an altruism multiplier which is an increasing function of fertility. As will be shown in Corollary 3, an agent with more bequests and higher wages in the first period will choose to have more children, rather than to leave a higher bequest to each child, to improve her lifetime welfare.

**Corollary 2.** *The optimal bequests under warm glow preferences are a function of 7 variables and parameters. The effects of these variables and parameters on the bequests are shown as follows where “+” represents positive effects and “-” represents negative effects.*

$$b'_{wg} = b \left( \begin{array}{c} \gamma, \epsilon, \tau, \tau_2, w_2, s, x \\ (-) (+) (-) (-) (+) (-) (+) \end{array} \right)$$

We can have the following implications from Corollary 2. First, the bequests are positively related to the elasticity of the utility function, which is  $1 - \gamma$ , and negatively related to the elasticity of altruism function which is  $1 - \epsilon$ . A higher elasticity of the utility function implies higher marginal utility of consumption, and thus a higher level of optimal consumption. Therefore, the optimal bequests will be high when the elasticity of the utility function is high, since bequests play a role similar to that of consumption. Similarly, a higher elasticity of altruism function implies higher fertility, which is preferred to higher bequests.

Second, wages in the second period are positively related to bequests. Higher wages in the second period imply higher opportunity costs of raising children. So, the substitution effect is negative. The income effect of higher wages in the second period is positive and will increase fertility. The total effect on fertility is ambiguous. However, the effect of higher wages in the second period on bequests is positive all the time. Namely, parents will choose to leave a higher bequest to each child when they can make more money in the second period. Finally, higher  $s$  or lower  $x$  means lower costs of raising children and thus will reduce bequests since fertility will increase.

Now let's move on to solve for fertility under warm glow preferences. We need specific forms of utility function and altruism parameters to achieve this goal. Assuming that  $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$  and  $\Phi(n) = \alpha \frac{n^{1-\epsilon}}{1-\epsilon}$ ,  $\alpha > 0$ , we can figure out the solution of fertility in the following steps.

Substituting consumption into the budget constraint, we have

$$c_1 + \frac{c_2}{R} = (\lambda_1)^{-\frac{1}{\gamma}} \left[ 1 + (\delta R)^{\frac{1}{\gamma}} R^{-1} \right] = (1 - \tau_1)w_1 + b + \frac{(1 - \tau_2)w_2(1 - nx) + ns - nb'/(1 - \tau)}{R}$$

$$\Rightarrow (\lambda_1)^{-\frac{1}{\gamma}} = \alpha_1 \left[ (1 - \tau_1)w_1 + b + \frac{(1 - \tau_2)w_2(1 - nx) + ns - nb'/(1 - \tau)}{R} \right] \quad (3.6)$$

where  $\alpha_1 = 1/\left[1 + (\delta R)^{\frac{1}{\gamma}} R^{-1}\right]$ .

Rewriting (4) we have

$$(\lambda_1)^{-\frac{1}{\gamma}} = \left[ \frac{1 - \epsilon}{(1 - \tau)\alpha\delta R} \right]^{\frac{1}{\gamma}} b' n^{\frac{\epsilon}{\gamma}} \quad (3.7)$$

Combining the two equations above with the solution for bequests, we have the following proposition regarding fertility under warm glow preferences.

**Proposition 2.** *Fertility under warm glow preferences is unique and determined by the following equation,*

$$\frac{(1 - \tau_1)w_1R + bR + (1 - \tau_2)w_2}{(1 - \tau_2)w_2x - s} - \frac{1 - \epsilon}{\gamma - \epsilon} n_{wg} = (1 - \tau) \frac{1 - \gamma}{\gamma - \epsilon} \left[ 1 + (\delta R)^{-\frac{1}{\gamma}} R \right] \left[ \frac{1 - \epsilon}{(1 - \tau)\alpha} \right]^{\frac{1}{\gamma}} (n_{wg})^{\frac{\epsilon}{\gamma}} \quad (3.8)$$

The left-hand-side is a downward line with a positive intercept. The right-hand-side is an upward curve from the origin. Therefore, there is a unique solution for optimal fertility. Unlike the relatively simple expression for optimal bequests, optimal fertility becomes much more complex.

**Corollary 3.** *Fertility under warm glow preferences is determined by 13 variables and parameters. The effects of these variables and parameters on bequests are shown as follows, where “+” represents positive effects, “-” represents negative effects, and “?” represents ambiguous effects.*

$$n_{wg} = n \left( \begin{array}{cccccccccccccc} \gamma, & \epsilon, & \tau, & \tau_1, & w_1, & \tau_2, & w_2, & s, & x, & \alpha, & \delta, & R, & b \\ (?) & (?) & (-) & (-) & (+) & (?) & (?) & (+) & (-) & (+) & (+) & (+) & (+) & (+) \end{array} \right)$$

*Proof.* The effect of  $\gamma$  on fertility is ambiguous. However, if  $\delta R \leq 1$  and  $n_{wg} > 1$ , then  $\frac{\partial n_{wg}}{\partial \gamma} > 0$ .

This is illustrated in figure 2.

The effect of  $\epsilon$  on fertility is also ambiguous. However, if  $n_{wg} > 1$ , then  $\frac{\partial n_{wg}}{\partial \epsilon} < 0$ . This is because both  $\frac{1 - \epsilon}{\gamma - \epsilon}$  and  $\frac{(1 - \epsilon)^{1/\gamma}}{\gamma - \epsilon}$  are increasing functions of  $\epsilon$ . When  $n_{wg} > 1$ ,  $(n_{wg})^{\frac{\epsilon}{\gamma}}$  is also increasing with  $\epsilon$ . This is illustrated in figure 2.

The effects of  $w_2$  and  $\tau_2$  are also ambiguous. But it is easy to see that fertility will be negatively related to  $w_2$  and positively related to  $\tau_2$  when there is no subsidy.

Other certain signs can be determined by similar methods. □

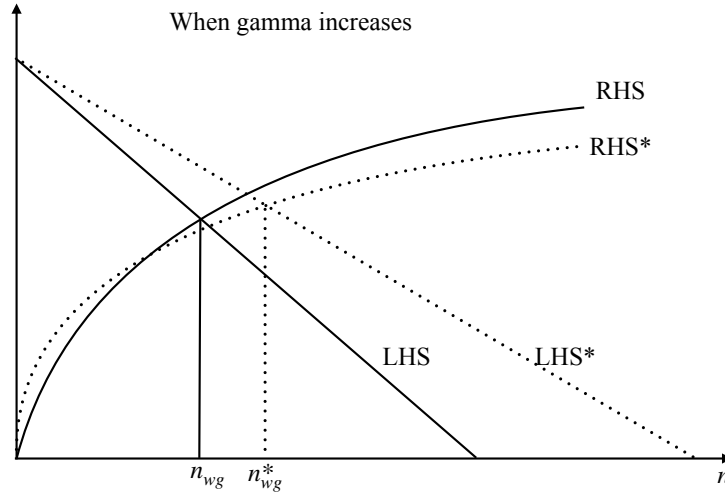


Figure 3.2 The Illustration of  $\frac{\partial n_{wg}}{\partial \gamma} > 0$

In steady state where  $b = b'$ , we will have the following proposition and corollary.

**Proposition 3.** *Steady-state fertility under warm glow preferences is unique and determined by the following equation,*

$$(1 - \tau) \frac{1-\gamma}{\gamma-\epsilon} R + \frac{(1-\tau_1)w_1R+(1-\tau_2)w_2}{(1-\tau_2)w_2x-s} - \frac{1-\epsilon}{\gamma-\epsilon} n_{wgss} = (1 - \tau) \frac{1-\gamma}{\gamma-\epsilon} \left[ 1 + (\delta R)^{-\frac{1}{\gamma}} R \right] \left[ \frac{1-\epsilon}{(1-\tau)\alpha} \right]^{\frac{1}{\gamma}} (n_{wgss})^{\frac{\epsilon}{\gamma}}$$

**Corollary 4.** *Steady-state fertility under warm glow preferences is determined by 12 variables and parameters,*

$$n_{wgss} = n \left( \begin{array}{cccccccccccc} \gamma, & \epsilon, & \tau, & \tau_1, & w_1, & \tau_2, & w_2, & s, & x, & \alpha, & \delta, & R \\ (?) & (?) & (-) & (-) & (+) & (?) & (?) & (+) & (-) & (+) & (+) & (+) \end{array} \right)$$

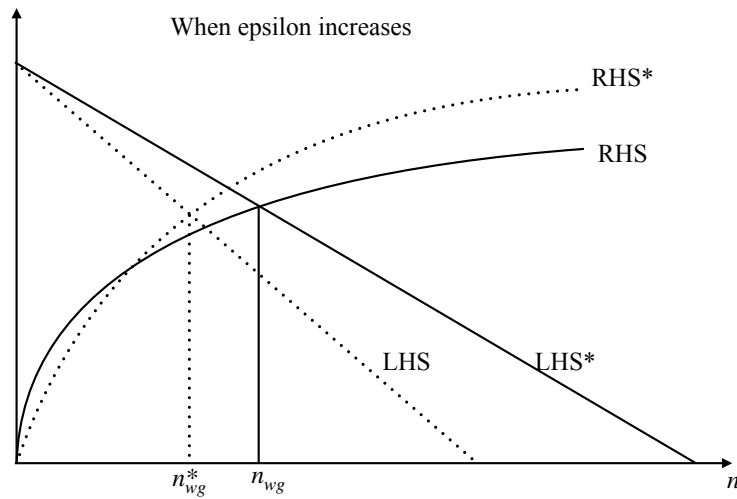


Figure 3.3 The Illustration of  $\frac{\partial n_{wg}}{\partial \epsilon} < 0$

Again, steady-state fertility under warm glow preferences will be negatively related to wages in the second period if there is no subsidy. However, the effects of  $\gamma$  and  $\epsilon$  can not be determined even if we restrict the optimal fertility to exceed unity.

We can see that in the case of warm glow, bequests have closed-form solutions and do not depend on the functional forms of the utility function and the altruism parameter. On the contrary, fertility under warm glow preferences does not have a closed-form solution and depends on the specific forms of the utility function and the altruism parameter.

### 3.3.2 The Case of Pure Altruism

We return to the general case without specific forms of the utility function and the altruism parameter. The young agent's value function is

$$V(b) = \max_{c_1, c_2, n, b'} \{u(c_1) + \delta u(c_2) + \delta \Phi(n)V(b')\} \quad (3.9)$$

subject to the same budget constraint with the case of warm glow,

$$c_1 + \frac{c_2}{R} = (1 - \tau_1)w_1 + b + \frac{(1 - \tau_2)w_2(1 - nx) + ns - nb'/(1 - \tau)}{R} := I$$

The same assumptions in the case of warm glows are still needed here. Let  $\lambda_2$  be the Lagrangian multiplier associated with the budget constraint. Optimal consumption choices satisfy

$$\lambda_2 = u'(c_1) = \delta R u'(c_2) \quad (3.10)$$

The optimal fertility and bequests conditions are

$$n : \lambda_2 \left( \frac{(1 - \tau)w_2x - s + b'/(1 - \tau)}{R} \right) = \delta \Phi'(n)V(b') \quad (3.11)$$

$$b' : \frac{\lambda_2 n}{R(1 - \tau)} = \delta \Phi(n)V'(b') \quad (3.12)$$

From the Envelope Theorem we have

$$V'(b) = \lambda_2 \quad (3.13)$$

So in steady states where  $b = b'$ , equation (13) and the bequest condition determine fertility

$$\frac{\Phi(n)}{n} = \frac{1}{(1 - \tau)\delta R} \quad (3.14)$$

We can see that fertility under pure altruism depends on the altruism parameter, bequest tax rate, time discounting factor and interest rate. It is independent of the utility function and of wages in both periods. We add the same functional forms of  $u(\cdot)$  and  $\Phi(\cdot)$ , as in the last section, from now on.

**Proposition 4.** *Steady-state fertility under pure altruism can be expressed as,*

$$n = \left[ \frac{(1 - \tau)\alpha\delta R}{1 - \epsilon} \right]^{\frac{1}{\epsilon}} \quad (3.15)$$

**Corollary 5.** *Steady-state fertility under pure altruism is determined by five parameters,*

$$n_{pass} = n \left( \begin{matrix} \epsilon, \tau, \alpha, \delta, R \\ (+), (-), (+), (+), (+) \end{matrix} \right)$$

Compared to steady-state fertility under warm glow preferences, we find differences and similarities.

**Corollary 6.** *Wages and the utility function will affect the steady-state fertility under warm glow preferences but have no influence on the steady-state fertility under pure altruism.*

This is the characteristic difference between warm glow and pure altruism with regard to their implications for fertility. Parents with pure altruism can foresee their children's lifetime utility perfectly, which implies steady-state fertility is determined by few fundamental parameters and renders wages and utility functions irrelevant in determining fertility.

**Corollary 7.** *Steady-state fertility is positively related to interest rates and negatively related to bequest tax rates in both cases.*

This is the characteristic similarity between warm glow and pure altruism in their implications for fertility. Raising kids is a kind of investment and interest rates are the returns for this kind of investment. Unsurprisingly, higher interest rates imply higher fertility. With higher bequest tax rates, parents are discouraged from raising children since the returns of this kind of investment will decrease.

Now we can move on to solve for bequests under pure altruism.

Substituting equation (10) into (9)

$$V(b) = \max_{n,b'} \left\{ \alpha_1^{-\gamma} \frac{I^{1-\gamma}}{1-\gamma} + \delta \frac{\alpha n^{1-\epsilon}}{1-\epsilon} V(b') \right\} = \max_{c_1, n, b'} \left\{ \frac{1}{\alpha_1} \frac{c_1^{1-\gamma}}{1-\gamma} + \delta \frac{\alpha n^{1-\epsilon}}{1-\epsilon} V(b') \right\} \quad (3.16)$$



Combining (16) with FOCs, we have

$$b : V'(b) = c_1^{-\gamma} \quad (3.17)$$

$$b' : c_1^{-\gamma} \frac{n}{R(1-\tau)} = \delta \frac{\alpha n^{1-\epsilon}}{1-\epsilon} V'(b') \quad (3.18)$$

$$n : c_1^{-\gamma} \left( \frac{(1-\tau_2)w_2x - s + b'/(1-\tau)}{R} \right) = \delta \alpha n^{-\epsilon} V(b') \quad (3.19)$$

From (17) and (18) we have

$$c_1^{-\gamma} = (1-\tau) \delta R \frac{\alpha n^{-\epsilon}}{1-\epsilon} (c_1')^{-\gamma} \quad (3.20)$$

From (19) we have

$$\begin{aligned} \frac{\alpha n^{1-\epsilon}}{1-\epsilon} V(b') &= \frac{\alpha n^{1-\epsilon}}{1-\epsilon} c_1^{-\gamma} \left( \frac{(1-\tau_2)w_2x - s + b'/(1-\tau)}{R} \right) \frac{1}{\delta \alpha n^{-\epsilon}} \\ &= \frac{c_1^{-\gamma}}{\delta(1-\epsilon)} \left( \frac{(1-\tau_2)w_2x - s + b'/(1-\tau)}{R} n \right) \end{aligned} \quad (3.21)$$

Substituting optimal consumption levels into the budget constraint, we have

$$c_1 + \frac{c_2}{R} = \frac{c_1}{\alpha_1} = (1-\tau_1)w_1 + b + \frac{(1-\tau_2)w_2(1-nx) + ns - nb'/(1-\tau)}{R} \quad (3.22)$$

By combining the two equations above, we have

$$\frac{\alpha n^{1-\epsilon}}{1-\epsilon} V(b') = \frac{c_1^{-\gamma}}{\delta(1-\epsilon)} \left( (1-\tau_1)w_1 + b + \frac{(1-\tau_2)w_2}{R} - \frac{c_1}{\alpha_1} \right) \quad (3.23)$$

From (19) and (20) we also have

$$V(b) = \frac{(c_1^0)^{-\gamma}}{\delta \alpha n^{-\epsilon}} \left( \frac{(1-\tau_2)w_2x - s + b'/(1-\tau)}{R} \right) = \frac{(1-\tau)R(c_1)^{-\gamma}}{1-\epsilon} \left( \frac{(1-\tau_2)w_2x - s + b'/(1-\tau)}{R} \right) \quad (3.24)$$

where the superscript of zero denotes the variables of the agent's parent. Rearranging the equation above we have

$$V(b) = \frac{c_1^{-\gamma}}{1-\epsilon} (1-\tau) \left( (1-\tau_2)w_2x - s + b'/(1-\tau) \right) \quad (3.25)$$

Substituting (23) and (25) into (16), we have

$$\frac{c_1^{-\gamma}}{1-\epsilon}(1-\tau) \left( (1-\tau_2)w_2x - s + b'/(1-\tau) \right) = \frac{1}{\alpha_1} \frac{c_1^{1-\gamma}}{1-\gamma} + \frac{c_1^{-\gamma}}{1-\epsilon} \left( (1-\tau_1)w_1 + b + \frac{(1-\tau_2)w_2}{R} - \frac{c_1}{\alpha_1} \right) \quad (3.26)$$

In equilibrium, we have

$$c_1^* = \alpha_1 \frac{1-\gamma}{\gamma-\epsilon} \left( (1-\tau) \left[ (1-\tau_2)w_2x - s \right] - (1-\tau_1)w_1 - \frac{(1-\tau_2)w_2}{R} \right) \quad (3.27)$$

Substituting consumption in the first period and fertility into (22), we have bequests in steady state

$$b_{pass} = \frac{R}{(1-\tau)R - \left[ \frac{(1-\tau)\alpha\delta R}{1-\epsilon} \right]^{\frac{1}{\epsilon}}} \left[ \begin{array}{c} \frac{1-\gamma}{\gamma-\epsilon}(1-\tau) \left[ (1-\tau_2)w_2x - s \right] \\ -\frac{1-\epsilon}{\gamma-\epsilon} \left[ (1-\tau_1)w_1 + \frac{(1-\tau_2)w_2}{R} \right] + \frac{(1-\tau_2)w_2x-s}{R} \left[ \frac{(1-\tau)\alpha\delta R}{1-\epsilon} \right]^{\frac{1}{\epsilon}} \end{array} \right]$$

Rearranging the variables, we have the following proposition,

**Proposition 5.** *Steady-state bequests under pure altruism can be expressed as,*

$$b_{pass} = \frac{R}{(1-\tau)R - \left[ \frac{(1-\tau)\alpha\delta R}{1-\epsilon} \right]^{\frac{1}{\epsilon}}} \left[ \begin{array}{c} (1-\tau_2)w_2 \left\{ \frac{1-\gamma}{\gamma-\epsilon}(1-\tau)x - \frac{1-\epsilon}{\gamma-\epsilon} \frac{1}{R} + \frac{x}{R} \left[ \frac{(1-\tau)\alpha\delta R}{1-\epsilon} \right]^{\frac{1}{\epsilon}} \right\} \\ -\frac{1-\epsilon}{\gamma-\epsilon}(1-\tau_1)w_1 - s \left\{ \frac{1-\gamma}{\gamma-\epsilon}(1-\tau) + \frac{1}{R} \left[ \frac{(1-\tau)\alpha\delta R}{1-\epsilon} \right]^{\frac{1}{\epsilon}} \right\} \end{array} \right]$$

**Corollary 8.** *Steady-state bequests under pure altruism is a function of 12 parameters and variables,*

$$b_{pass} = b \left( \begin{array}{cccccccccccc} \gamma, & \epsilon, & \tau, & \tau_1, & w_1, & \tau_2, & w_2, & s, & x, & \alpha, & \delta, & R \\ (-) & (+) & (-) & (+) & (-) & (?) & (?) & (-) & (+) & (+) & (+) & (?) \end{array} \right)$$

Compared to the bequests under warm glow preferences, we find differences and similarities.

**Corollary 9.** *Interest rates and wages in the first period will affect the steady-state bequests under pure altruism, but will not affect the bequests under warm glow preferences.*

This is the difference between warm glow and pure altruism in their implications for bequests. It should be noted that the bequests given are independent of the bequests received. So the concept of steady-state bequests do not make sense under warm glow preferences.

**Corollary 10.** *Bequests under both warm glow and pure altruism are positively related to the elasticity of the consumption function and the time percentage of raising a child; and are negatively related to the elasticity of altruism function, bequest tax rates and subsidy.*

These are the similarities between warm glow and pure altruism in their implications for bequests. A higher elasticity of consumption function implies higher bequests since bequests can be regarded as consumption in our models. Parents will leave a higher bequest to each child to compensate for having fewer children due to higher raising costs. On the contrary, parents will also leave a lower bequest to each child in order to have more children due to the higher elasticity of altruism function.

### 3.4 An Example

In this section, we will use an example to show the differences between warm glow and pure altruism. Specifically, we will use pure altruism to replace warm glow in a model and compare the new results with the previous ones. The benchmark model is from Galor and Zeira (RES 1993; GZ hereafter). The GZ model includes two variables, education investment and bequest.

Being consistent with GZ, each individual has one parent and one child. Every individual lives two periods. It is also assumed that individuals consume only in the second period of life. There are also two major assumptions in the GZ model.

**Assumption 4.** *Credit markets are imperfect. The interest rates for borrowers,  $i$ , are higher than that for lenders,  $r$ .*

This assumption is critical to generating the dynamics of bequests in the GZ model, but does not play such an important role in our revised model with pure altruism, shown later.

**Assumption 5.** *Investment in human capital is indivisible.*

The implication of this assumption is that an individual will not invest in education to become a skilled worker until her wealth reaches some threshold level.

The utility function which is needed to be maximized in GZ is,

$$u = \alpha \ln c + (1 - \alpha) \ln b$$

$b$  is the bequest the individual gives to her child. So it is warm glow here that represents the parent's love. If we use pure altruism to replace warm glow, the value function we need to maximize becomes

$$V(x) = \alpha \ln c + (1 - \alpha)V(b)$$

$x$  is the bequest the individual receives from her own parent. We can consider  $1 - \alpha$  as the parent's altruism parameter. The higher the altruism parameter is, the more bequests the parent will leave to her children.

We have three different cases.

### 3.4.1 Case One: Unskilled Workers

Unskilled workers will not invest in human capital. They are lenders in the credit market because they can work in the first period and, by assumption, consumption only happens in the second period. We use  $V_1$  to denote the value function of unskilled workers. Currently, we assume that the children of unskilled workers will still be unskilled workers. The maximization problem is

$$\max_{c,b} V_1(x) = \alpha \ln c + (1 - \alpha)V_1(b)$$

subject to

$$I_1 = c + b = (x + w_n)(1 + r) + w_n$$

Substituting the budget constraint into the objective function, the maximization problem becomes

$$\max_b V_1(x) = \alpha \ln(I_1 - b) + (1 - \alpha)V_1(b)$$

FOCs w.r.t.  $b$  and  $x$  are

$$\frac{\alpha}{I_1 - b} = (1 - \alpha)V_1'(b)$$

$$V_1'(x) = \frac{\alpha}{I_1 - b}(1 + r)$$

Combining the two equations above, we have

$$\frac{V_1'(b)}{V_1'(x)} = \frac{1}{(1 - \alpha)(1 + r)}$$

**Lemma 1.**  $V_1(x)$  is strictly increasing and concave.

*Proof.*  $V_1'(x) = \frac{\alpha}{I_1 - b}(1 + r) > 0$ ,  $V_1''(x) = -\frac{\alpha}{(I_1 - b)^2}(1 + r)^2 < 0$ . □

**Proposition 6.** *The dynamics of bequests of unskilled workers depend on the interest rates of lending and the altruism parameter. Specifically speaking,*

1. when  $(1 - \alpha)(1 + r) < 1 \Rightarrow b < x$ . Bequests will keep decreasing generation by generation.
2. when  $(1 - \alpha)(1 + r) = 1 \Rightarrow b = x$ . Bequests will not change across generations.
3. when  $(1 - \alpha)(1 + r) > 1 \Rightarrow b > x$ . Bequests will keep increasing generation by generation.

The results are quite different from that in GZ. In GZ for unskilled workers, there exists a unique low-level steady-state equilibrium when  $(1 - \alpha)(1 + r) < 1$  which is the only situation discussed. The low-level equilibrium is also locally stable. A small deviation from this equilibrium will be eliminated automatically. In our model, however, there exists infinitely multiple low-level steady-state equilibria when  $(1 - \alpha)(1 + r) = 1$ . The equilibria are neither stable nor unstable locally. A small deviation from any equilibrium will neither be eliminated nor be augmented.

It can also be noticed that bequests will keep increasing in the third case. When bequests reach some threshold level, the optimal choice will be to invest in human capital and be skilled workers. A low-level equilibrium is not a trap for unskilled workers if interest rates are high enough or they love their children so much that they give a disproportionately high weight on their children's welfare. After the accumulation of several generations of unskilled workers, a generation of skilled workers will finally come out.

In the GZ model, however, this interesting case is ruled out because it is assumed  $(1 - \alpha)(1 + r) <$

1.

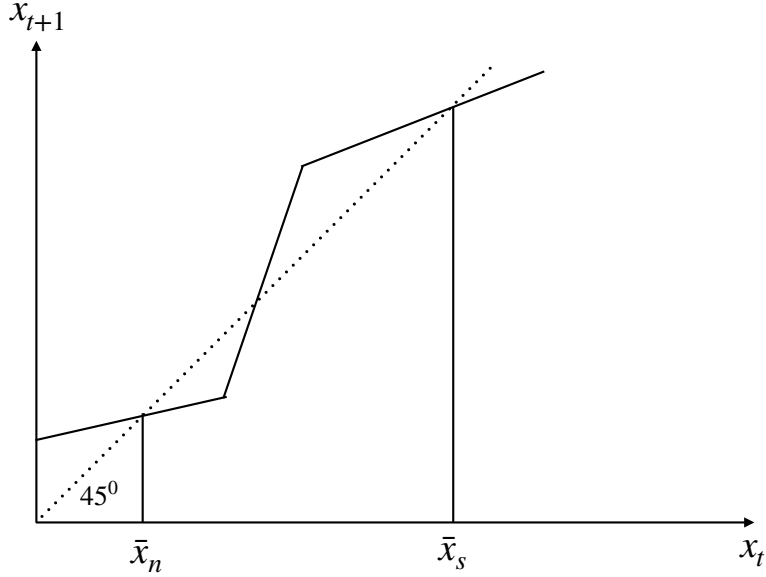


Figure 3.4 The Dynamics of Bequests in GZ Model

### 3.4.2 Case Two: Skilled Workers as Borrowers

If an individual decides to invest in human capital in the first period, she will become a skilled worker in the second period. The indivisible investment in human capital is  $h$ . If the individual receives a bequest less than the required investment, she needs to borrow money in credit market. Again, we assume that the child of a skilled worker will also work as a skilled worker when she grows up. The maximization problem is

$$\max_{c,b} V_2(x) = \alpha \ln c + (1 - \alpha)V_2(b)$$

subject to

$$I_2 = c + b = w_s + (x - h)(1 + i)$$

where  $w_2$  is the wage of skilled workers.

Substituting the budget constraint into the objective function, the maximization problem becomes

$$\max_b V_2(x) = \alpha \ln(I_2 - b) + (1 - \alpha)V_2(b)$$

FOCs w.r.t.  $b$  and  $x$  are

$$\frac{\alpha}{I_2 - b} = (1 - \alpha)V_2'(b)$$

$$V_2'(x) = \frac{\alpha}{I_2 - b}(1 + i)$$

Combining the two equations above, we have

$$\frac{V_2'(b)}{V_2'(x)} = \frac{1}{(1 - \alpha)(1 + i)}$$

**Lemma 2.**  $V_2(x)$  is strictly increasing and concave.

*Proof.*  $V_2'(x) = \frac{\alpha}{I_2 - b}(1 + i) > 0$ ,  $V_2''(x) = -\frac{\alpha}{(I_2 - b)^2}(1 + i)^2 < 0$ . □

**Proposition 7.** *The dynamics of bequests of skilled workers as borrowers depend on the interest rates of borrowing and the altruism parameter. Specifically speaking,*

1. when  $(1 - \alpha)(1 + i) < 1 \Rightarrow b < x$ . *Bequests will keep decreasing.*
2. when  $(1 - \alpha)(1 + i) = 1 \Rightarrow b = x$ . *Bequests will not change.*
3. when  $(1 - \alpha)(1 + i) > 1 \Rightarrow b > x$ . *Bequests will keep increasing.*

The results here are also different from that in GZ. In GZ for skilled workers who need borrowing, there exists a unique, but unstable, medium-level steady-state equilibrium when  $(1 - \alpha)(1 + i) > 1$  which is assumed and the only case discussed in GZ. A tiny deviation from this medium-level equilibrium will make a big difference: a positive deviation will lead to the locally stable high-level equilibrium while a negative deviation will lead to the locally stable low-level equilibrium.

Our results show that the dynamics of bequests are more rich. When  $(1 - \alpha)(1 + i) < 1$ , bequests will decrease across generations and finally the new generation will choose to be unskilled workers. When  $(1 - \alpha)(1 + i) = 1$ , bequests will keep the same and every generation will borrow to invest in human capital to be skilled workers. When  $(1 - \alpha)(1 + i) > 1$ , bequests will increase and finally

the new generation will have enough money to cover the investment of education. At that time, the individual will be a skilled worker and a lender at the same time.

### 3.4.3 Case Three: Skilled Workers as Lenders

If an individual receives enough bequest, she can not only invest in human capital, but also lend to the rest. It is also assumed that the child of a skilled worker as a lender will also be a skilled worker. The maximization problem is

$$\max_{c,b} V_3(x) = \alpha \ln c + (1 - \alpha)V_3(b)$$

subject to

$$I_3 = c + b = w_s + (x - h)(1 + r)$$

Substituting the budget constraint into the objective function, the maximization problem becomes

$$\max_b V_3(x) = \alpha \ln(I_3 - b) + (1 - \alpha)V_3(b)$$

FOCs w.r.t.  $b$  and  $x$  are

$$\frac{\alpha}{I_3 - b} = (1 - \alpha)V_3'(b)$$

$$V_3'(x) = \frac{\alpha}{I_3 - b}(1 + r)$$

Combining the two equations above, we have

$$\frac{V_3'(b)}{V_3'(x)} = \frac{1}{(1 - \alpha)(1 + r)}$$

**Lemma 3.**  $V_3(x)$  is strictly increasing and concave.

*Proof.*  $V_3'(x) = \frac{\alpha}{I_3 - b}(1 + r) > 0$ ,  $V_3''(x) = -\frac{\alpha}{(I_3 - b)^2}(1 + r)^2 < 0$ . □

**Proposition 8.** *The dynamics of bequests of skilled workers as lenders depend on the lending interest rates and the altruism parameter. Specifically speaking,*

1. when  $(1 - \alpha)(1 + r) < 1 \Rightarrow b < x$ . Bequests will keep decreasing.



2. when  $(1 - \alpha)(1 + r) = 1 \Rightarrow b = x$ . *Bequests will not change.*
3. when  $(1 - \alpha)(1 + r) > 1 \Rightarrow b > x$ . *Bequests will keep increasing.*

In the GZ model where  $(1 - \alpha)(1 + r) < 1$  is assumed, the dynamic is that bequests of skilled workers as lenders will lead to a locally stable high-level equilibrium. In my model, however, there is no equilibrium when  $(1 - \alpha)(1 + r) < 1$ . On the contrary, bequests will keep decreasing across generations and finally the new generation has to borrow to invest on human capital, as discussed in the previous case. When  $(1 - \alpha)(1 + r) = 1$ , bequests will stay the same, meaning that the child will replicate her parent's life style. When  $(1 - \alpha)(1 + r) > 1$ , bequests will increase across generations. The next generation will still be skilled workers as lenders. The only difference is that the next generation will have more money to lend.

### 3.5 Conclusions

We explored the implications of warm glow and pure altruism for fertility and bequests when they are used to model parental love. We find that the two kinds of preferences have dramatically different implications for fertility and bequests. In the case of warm glow preferences, bequests have a relatively simple expression and is independent of wages in the first period. Fertility is positively related to wages in the first period. The relationship between fertility and wages in the second period is ambiguous.

In the case of pure altruism, fertility has a relatively simple expression and is independent of wages in both periods. Bequests are negatively related to wages in the first period. The relationship between bequests and wages in the second period is ambiguous. In the extended example, we show that the dynamics of wealth will be different if we use pure altruism to replace warm glow to model parental love.

Both warm glow and pure altruism are used to model parental love towards their children. The attributes of the two kinds of love, however, are not the same. Warm glow represents selfish love. Parents love their children from their own perspective. They care about how much they have sacrificed for their children. Pure altruism represents selfless love. Parents love their children from

the perspective of their children. They care about whether their children have a good life. Because of the huge differences between the two kinds of love, it is not a surprise to find that warm glow and pure altruism have dramatically different implications for fertility and bequests. The same behavior with different motivations produces different consequences.

## CHAPTER 4. FERTILITY, PURE ALTRUISM AND HYPERBOLIC DISCOUNTING

### 4.1 Introduction

In traditional models of economics, we assume that people discount future utilities at the same rates. For example, if we assume the discounting rate is  $\delta$  in a discrete-time model, then period- $t$  utility (denoted as  $u(t)$ ) is discounted at the rate  $\delta^t u(t)$  at the beginning  $t = 0$ . This form of discounting is known as exponential discounting. It is first proposed by Samuelson (1937). If a person exhibits exponential discounting preferences, she is able to make optimal plans for every future period at the beginning and will not change her plans afterwards. We say that this person's behavior is time-consistent. The assumption of constant time discounting is so strong that Samuelson himself stated like this, "This third assumption, unlike the previous two, is in the nature of an hypothesis, subject to refutation by the observable facts." At the end of Samuelson (1937), he warned again that, "In conclusion, any connection between utility as discussed here and any welfare concept is disavowed." Although exponential discounting is contested both normatively and descriptively, it is so elegant and easily tractable that it is accepted by the mainstream economics quickly and becomes the standardized assumption or even axiom, to some extent, in modern economics. Samuelson's warning has been forgotten.

In reality, however, people's behaviors are not always time-consistent. The assumption of constant time discounting is rejected by observable facts, as Samuelson guessed. Another kind of discounting, hyperbolic discounting, has been proposed to explain time-inconsistent behavior [Phelps and Pollak (1968); Laibson (1997); O'Donoghue and Rabin (1999)]. Technically speaking, an agent who discounts hyperbolically gives a fixed discount for all the future utilities based on exponential discounting. Intuitively, such an agent shows a changing, instead of constant, time discounting and behaves as present-biased. Under hyperbolic discounting, an agent in the second period will change

her plans made in the first period if she is naive, meaning that she is not able, at the beginning, to foresee her optimization problem in the second period. In this case, we say the agent has a self-control problem. If the agent is sophisticated, she will foresee her optimal second-period decision in the first period. Using this information, she will make such plans in the first period such that there is no incentive for her to change her plans in the second period.

Hyperbolic discounting is widely used in many fields of economics. For example, it is used to explain saving decisions [Laibson et al. (1998); Diamond and Köszegi (2003); Cao and Werning (2018)], addictive behavior [Gruber and Köszegi (2001)] and fertility patterns across generations [Gobbi and Goni (2016)]. It is also used to study public policies to deal with the issue of global warming [Karp (2005)]. In political game theory, hyperbolic discounting has also gained popularity [Cao and Werning (2016)].

However, there is little research on fertility within one generation under hyperbolic discounting. Wrede (2011) studies fertility decisions under hyperbolic discounting and assumes that women can choose to have children in the first two periods of a three-period life. With financial constraints, fertility decreases due to hyperbolic discounting. Without financial constraints, however, the author finds that the effect of hyperbolic discounting on fertility becomes ambiguous, depending on whether children are considered as consumption goods or investment goods. When children are considered as consumption goods, sophisticated mothers will choose to have more children in both periods. When children are considered as investment goods, hyperbolic discounting will reduce the fertility of sophisticated mothers in the second period. Parents are assumed not to be altruistic in this paper.

Wigniolle (2013) studies the quantity-quality trade-off under hyperbolic discounting. This paper shows that people in developed countries and developing countries react differently. Specifically speaking, hyperbolic discounting reduces fertility in developed countries and increases fertility in developing countries. The results are obtained under very strong assumptions: the cost of quantity with respect to family income is relatively higher, and the cost of quality relatively lower

Table 4.1 The Position of This Paper in Literature

	Exponential Discounting	Hyperbolic Discounting
Warm Glow	1: Special Case of 2	2: Wrede (2011); Wigniolle (2013)
Pure Altruism	3: Barro and Becker(1988)	4: This paper

in developed countries. Parents care about how much education their children will receive. Also, no financial market exists, since there is no saving behavior.

The two papers above both use warm glow to represent parental love towards their children. That is, in their model, parents do not care about their children's lifetime utility. Parents with warm glow preference care about only how many bequests they leave to their children. Parents with pure altruism preference, on the contrary, care about their children's welfare and try to improve their children's lifetime utility by leaving bequests. The mechanisms of warm glow and pure altruism are illustrated in Figure 3.1 in chapter 3. Although the two kinds of preferences require parents to leave bequests to their children, their implications are quite different.

Becker and Barro (1988), which is the cornerstone of modern fertility economics, incorporates pure altruism into their model. I will follow this paper and use pure altruism as the benchmark to explore this issue. However, hyperbolic discounting is not considered in Becker and Barro (1988). It is necessary to explore fertility decisions under both pure altruism and hyperbolic discounting. The position of this paper in the literature is illustrated in Table 1. The aim of this paper is to fill in the hole of case 4. As far as I know, this is the first paper to consider fertility decisions under hyperbolic discounting with pure altruism preferences. Besides fertility, I also take educational investment in children as endogenous. I find that parents with hyperbolic discounting preference and full access to financial markets will have lower fertility. This conclusion is in complete contradiction to Wrede (2011) which finds that fertility will decrease without financial markets and it is ambiguous with financial markets. Contrary to Wigniolle (2013), I also find that hyperbolic discounting reduces fertility in all countries. Finally, I find that educational investment will also be influenced by hyperbolic discounting.

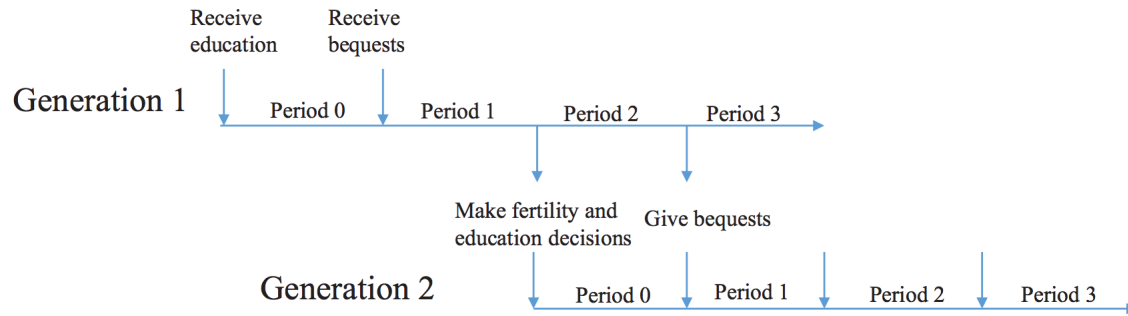


Figure 4.1 The Order of Decisions

The paper is organized as follows. Section 4.2 will present the model. I will discuss two cases: the naive agent's problem and the sophisticated agent's problem. Section 4.3 will give an easy example to demonstrate the two cases. Section 4.4 concludes.

## 4.2 The Model

I assume that an agent lives four periods: childhood, young, middle-aged and old. It is assumed that a child receives education investment from her parent, but does not make any decisions. Children are assumed to have no consumption for simplicity since they make no decisions. The childhood period is denoted as period zero.

For simplicity, a family is assumed to include only one parent and her children. So in this model, the replacement fertility rate is one instead of two. When the agent is young, she receives a bequest from her parent and makes saving decisions. When middle-aged, she makes fertility and saving decisions, as well as decisions for her level of education investment in her children. In this period, her children are born and in childhood. When old, she makes bequest decisions for all of her children. I also assume that all the children provide the same amount of utility to their parent and receive the same amount of bequests. Consumption decisions are also made in every period. All the decisions are made at the beginning of each period. The order of decisions is illustrated in Figure 4.2.

From now on I will denote childhood as period 0 and youth as period 1 and so on. Following the practice of Galor and Zeira (1993), period 0, in which human capital is formed, is ignored in the analysis. The agent's wage in the first period is  $h(e)$  where  $e$  is the educational investment she receives. I assume  $h'(e) > 0, h''(e) < 0$ . The wage in the second period is  $(1 + g)h(e)$ , where  $g$  is the growth rate of wages. The agent does not have income in period 3 since she is in retirement. The agent's problem is to choose her consumption levels, her number of children, the level of educational investment she will provide for her children and the level of bequests for her children to maximize her lifetime utility. I also assume that children's utility enters parent's utility function when bequests are given. This is also the time when children are independent and start to work.

The value function of an agent in period  $t$  is denoted as  $V_t(a_t)$  where  $a_t$  is the agent's wealth at the beginning of period  $t$ .  $a'_t$  is the wealth of the agent's children at the beginning of their period  $t$ . Time costs of raising  $n$  children are  $(1 + g)h(e)nx$  which are the only costs for parents.  $x$  is the percentage of a parent's total time she spends to raise one child. The altruism parameter for  $n$  children is  $\Phi(n)$ . It is assumed to be strictly increasing and concave. I denote the agent in period  $t$  as Self  $t$ . So the three Bellman equations are

Self 1:

$$V_1(e, b) = \max_{c_1, c_2, c_3, n, e', b'} \{u(c_1) + \beta\delta u(c_2) + \beta\delta^2 u(c_3) + \beta\delta^2 \Phi(n)V_1(e', b')\} \quad (4.1)$$

Self 2:

$$V_2(e, a_2) = \max_{c_2, c_3, n, e', b'} \{u(c_2) + \beta\delta u(c_3) + \beta\delta \Phi(n)V_1(e', b')\} \quad (4.2)$$

Self 3:

$$V_3(e', n, a_3) = \max_{c_3, b'} \{u(c_3) + \Phi(n)V_1(e', b')\} \quad (4.3)$$

#### 4.2.1 Model 1: A Naive Agent's Commitment Problem

If a naive agent makes all the optimal decisions in the first period and the decisions are implemented in the following periods, I refer to the sequence of decisions as a commitment. In the

commitment problem, Self 1 makes all the choices.

$$V_1(e, b) = \max_{c_1, c_2, c_3, n, e', b'} \{u(c_1) + \beta\delta u(c_2) + \beta\delta^2 u(c_3) + \beta\delta^2 \Phi(n)V_1(e', b')\} \quad (4.4)$$

subject to

$$c_1 + \frac{c_2}{R} + \frac{c_3}{R^2} = h(e) + b + \frac{(1+g)h(e)(1-nx) - ne'}{R} - \frac{nb'}{R^2} := I \quad (4.5)$$

I denote the right-hand side as  $I$  and denote the solutions in steady states as  $c_1^*, c_2^*, c_3^*, n^*, e^*, b^*$ .

**Proposition 9.** *In the commitment problem, fertility is determined by the following equation*

$$\beta\delta^2 R^2 \frac{\Phi(n^*)}{n^*} = 1 \quad (4.6)$$

*Proof.* Let  $\lambda$  be the Lagrangian multiplier associated to the budget constraint. Optimal consumption choices satisfy

$$\lambda = u'(c_1) = \beta\delta R u'(c_2) = \beta\delta^2 R^2 u'(c_3) \quad (4.7)$$

The optimal education, fertility and bequest conditions read

$$\frac{\lambda n}{R} = \beta\delta^2 \Phi(n) \frac{\partial V_1(e', b')}{\partial e'} \quad (4.8)$$

$$\lambda \left( \frac{(1+g)h(e)x + e'}{R} + \frac{b'}{R^2} \right) = \beta\delta^2 \Phi'(n) V_1(e', b') \quad (4.9)$$

$$\frac{\lambda n}{R^2} = \beta\delta^2 \Phi(n) \frac{\partial V_1(e', b')}{\partial b'} \quad (4.10)$$

From the Envelope Theorem we have

$$\frac{\partial V_1(e, b)}{\partial e} = \lambda \left[ 1 + \frac{(1+g)(1-nx)}{R} \right] h'(e) \quad (4.11)$$

$$\frac{\partial V_1(e, b)}{\partial b} = \lambda \quad (4.12)$$

So in steady states, the condition above and bequest condition decide fertility

$$1 = \beta\delta^2 R^2 \frac{\Phi(n)}{n} \quad (4.13)$$

□



**Corollary 11.** *The existence of hyperbolic discounting reduces naive agents' fertility since  $\frac{dn^*}{d\beta} > 0$ .*

*Proof.*  $\beta \downarrow \Rightarrow \frac{\Phi(n)}{n} \uparrow \Rightarrow n \downarrow$ . Hyperbolic discounting means that  $\beta < 1$ . So fertility also decreases.  $\square$

**Corollary 12.** *Fertility in the commitment problem is independent of utility function and is determined by four parameters: hyperbolic discounting rate, time discounting rate, interest rate and altruism parameter.*

We can also solve for the optimal education investment.

**Proposition 10.** *In the commitment problem, education in steady state can be determined by the following equation*

$$h'(e^*) \left[ \frac{1}{R} + \frac{(1+g)(1-n^*x)}{R^2} \right] = 1 \quad (4.14)$$

*Proof.* The two conditions w.r.t. education give the following equation in steady state

$$\beta\delta^2\Phi(n) \left[ 1 + \frac{(1+g)(1-nx)}{R} \right] h'(e) = \frac{n}{R} \quad (4.15)$$

Rearranging the equation, we have

$$\beta\delta^2R^2\frac{\Phi(n)}{n} \left[ \frac{1}{R} + \frac{(1+g)(1-nx)}{R^2} \right] h'(e) = 1 \Rightarrow \left[ \frac{1}{R} + \frac{(1+g)(1-nx)}{R^2} \right] h'(e) = 1$$

$\square$

**Corollary 13.** *Higher wage growth rates lead to higher fertility and more education investment since  $\frac{de}{dg} > 0$ .*

In order to solve for education investment, we need a specific utility function. Let's assume that  $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$  and denote  $\epsilon(c) = \frac{u'(c)c}{u(c)}$ ,  $\epsilon(n) = \frac{\Phi'(n)n}{\Phi(n)}$ . From equation (7) and the budget constraint, we have

$$c_1 = \frac{I}{1 + (\beta\delta R)^{\frac{1}{\gamma}} R^{-1} + (\beta\delta^2 R^2)^{\frac{1}{\gamma}} R^{-2}} = \alpha_1 I \quad (4.16)$$

$$c_2 = \frac{(\beta\delta R)^{\frac{1}{\gamma}} I}{1 + (\beta\delta R)^{\frac{1}{\gamma}} R^{-1} + (\beta\delta^2 R^2)^{\frac{1}{\gamma}} R^{-2}} = (\beta\delta R)^{\frac{1}{\gamma}} \alpha_1 I \quad (4.17)$$

$$c_3 = \frac{(\beta\delta^2 R^2)^{\frac{1}{\gamma}} I}{1 + (\beta\delta R)^{\frac{1}{\gamma}} R^{-1} + (\beta\delta^2 R^2)^{\frac{1}{\gamma}} R^{-2}} = (\beta\delta^2 R^2)^{\frac{1}{\gamma}} \alpha_1 I \quad (4.18)$$

where  $\alpha_1 = \frac{1}{1 + (\beta\delta R)^{\frac{1}{\gamma}} R^{-1} + (\beta\delta^2 R^2)^{\frac{1}{\gamma}} R^{-2}}$  is the percentage of income that will be consumed in period

1. Substituting the three equations into (4)

$$V_1(e, b) = \max_{c_1, n, e', b'} \left\{ \alpha_1^{-\gamma} \frac{I^{1-\gamma}}{1-\gamma} + \beta\delta^2 \Phi(n) V_1(e', b') \right\} = \max_{c_1, n, e', b'} \left\{ \frac{u(c_1)}{\alpha_1} + \beta\delta^2 \Phi(n) V_1(e', b') \right\} \quad (4.19)$$

We have three new FOCs which are given as following

$$b : \frac{\partial V_1(e, b)}{\partial b} = u'(c_1) \quad (4.20)$$

$$b' : u'(c_1) \frac{n}{R^2} = \beta\delta^2 \Phi(n) \frac{\partial V_1(e', b')}{\partial b'} \quad (4.21)$$

$$n : u'(c_1) \left( \frac{(1+g)h(e)x + e'}{R} + \frac{b'}{R^2} \right) = \beta\delta^2 \Phi'(n) V_1(e', b') \quad (4.22)$$

From (20) and (21) we have

$$u'(c_1) = \beta\delta^2 R^2 \frac{\Phi(n)}{n} u'(c_1) \quad (4.23)$$

From (23) we have

$$\begin{aligned} \Phi(n) V_1(e', b') &= u'(c_1) \left( \frac{(1+g)h(e)x + e'}{R} + \frac{b'}{R^2} \right) \frac{\Phi(n)}{\beta\delta^2 \Phi'(n)} \\ &= \frac{u(c_1)}{c_1} \left( \frac{(1+g)h(e)nx + ne'}{R} + \frac{nb'}{R^2} \right) \frac{\epsilon(c_1)}{\beta\delta^2 \epsilon(n)} \end{aligned} \quad (4.24)$$

From the budget constraint and consumption levels we have

$$c_1 + \frac{c_2}{R} + \frac{c_3}{R^2} = \frac{c_1}{\alpha_1} = h(e) + b + \frac{(1+g)h(e)}{R} - \left( \frac{(1+g)h(e)nx + ne'}{R} + \frac{nb'}{R^2} \right) = I \quad (4.25)$$

By combining the two equations above, we have

$$\Phi(n)V_1(e', b') = \frac{u(c_1)}{c_1} \left( h(e) + b + \frac{(1+g)h(e)}{R} - \frac{c_1}{\alpha_1} \right) \frac{\epsilon(c_1)}{\beta\delta^2\epsilon(n)} \quad (4.26)$$

From (22) and (23) we also have

$$V_1(e, b) = \frac{u'(c_1^0)}{\beta\delta^2\Phi'(n)} \left( \frac{(1+g)h(e^0)x + e}{R} + \frac{b}{R^2} \right) = \frac{R^2 u'(c_1)}{\epsilon(n)} \left( \frac{(1+g)h(e^0)x + e}{R} + \frac{b}{R^2} \right) \quad (4.27)$$

where the superscript of zero denotes the variables of the agent's parent. Rearranging the equation above we have

$$V_1(e, b) = \frac{\epsilon(c_1)}{\epsilon(n)} \left( (1+g)h(e^0)xR + eR + b \right) \frac{u(c_1)}{c_1} \quad (4.28)$$

Substituting (24) and (28) into (19), we have

$$\frac{\epsilon(c_1)}{\epsilon(n)} \left( (1+g)h(e^0)xR + eR + b \right) \frac{u(c_1)}{c_1} = \frac{u(c_1)}{\alpha_1} + \frac{u(c_1)}{c_1} \left( h(e) + b + \frac{(1+g)h(e)}{R} - \frac{c_1}{\alpha_1} \right) \frac{\epsilon(c_1)}{\epsilon(n)} \quad (4.29)$$

In equilibrium, we have

$$c_1^* = \alpha_1 \frac{\epsilon(c_1^*)}{\epsilon(n^*) - \epsilon(c_1^*)} \left( (1+g)h(e^*)xR + e^*R - h(e^*) - \frac{(1+g)h(e^*)}{R} \right) \quad (4.30)$$

If educational investment is fixed and the elasticities of consumption and fertility are constant, we can see that the existence of hyperbolic discounting increases the consumption in the first period and reduces the consumption in the second and third periods.

From (25) and the equation above, we have the optimal amount of bequest for each child

$$b^* = \frac{R^2}{R^2 - n^*} \left[ \left( (1+g)h(e^*)xR + e^*R \right) \left( \frac{\epsilon(c_1^*)}{\epsilon(n^*) - \epsilon(c_1^*)} + \frac{1}{R^2} \right) - \frac{\epsilon(n^*)}{\epsilon(n^*) - \epsilon(c_1^*)} \left( h(e^*) + \frac{(1+g)h(e^*)}{R} \right) \right] \quad (4.31)$$

#### 4.2.2 Model 2: A Sophisticated Agent's Time-Consistent Problem

One characteristic of a naive agent with hyperbolic discounting preferences is that the agent will change her first-period optimal plans since the agent is not able to foresee the future in the first period. However, a sophisticated agent can foresee the future and prevent changes in her decisions

in later periods. In order to make time-consistent plans in the first period, we must consider the problem backwards. First we consider Self 3's choices and then we take these choices as given and consider Self 2's optimal choices. Then, we take Self 2 and 3's choices as given and consider Self 1's optimal choices. In this way, the agent will not change her optimal plans. We call a solution gained by the backward method as a time consistent solution. In this problem, Self 3 chooses  $c_3$  and  $b'$  given  $e', a_3$  and  $n$ ; Self 2 chooses  $c_2, n, e'$  and  $a_3$  given  $a_2$  and  $e$ ; and Self 1 chooses  $c_1$  and  $a_2$  given  $e$  and  $b$ . The problems are described as follows:

Self 3:

$$W_3(e', a_3, n) = \max_{c_3, b'} \{u(c_3) + \Phi(n)W_1(e', b')\} \quad (4.32)$$

subject to

$$c_3 = a_3 - nb' \quad (4.33)$$

Denote the solutions to this problem  $c_3 = c_3^W(e', a_3, n)$  and  $b = b^W(e', a_3, n)$ .

Self 2 solves:

$$W_2(a_2, e) = \max_{c_2, a_3, n, e'} \{u(c_2) + \beta\delta W_3(e', a_3, n)\} \quad (4.34)$$

subject to

$$c_2 = a_2 + (1 + g)h(e)(1 - nx) - ne' - \frac{a_3}{R} \quad (4.35)$$

Denote the solutions to this problem  $c_2 = c_2^W(a_2, e)$ ,  $a_3 = a_3^W(a_2, e)$  and  $n = n^W(a_2, e)$ .

Self 1 solves:

$$W_1(e, b) = \max_{c_1, a_2} \{u(c_1) + \beta\delta W_2(a_2, e)\} \quad (4.36)$$

subject to

$$c_1 = h(e) + b - \frac{a_2}{R} \quad (4.37)$$

**Proposition 11.** *In the time consistent problem, fertility is determined by the following equation*

$$(\beta\delta R)^2 \frac{\Phi(n^W)}{n^W} = 1 \quad (4.38)$$

*Proof.* Self 3. The first order condition w.r.t. bequest is

$$u'(c_3)n = \Phi(n) \frac{\partial W_1(e', b')}{\partial b'} \quad (4.39)$$

The Envelope Theorem conditions are

$$\frac{\partial W_3(e', a_3, n)}{\partial e'} = \Phi(n) \frac{\partial W_1(e', b')}{\partial e'} \quad (4.40)$$

$$\frac{\partial W_3(e', a_3, n)}{\partial a_3} = u'(c_3) \quad (4.41)$$

$$\frac{\partial W_3(e', a_3, n)}{\partial n} = -u'(c_3)b' + \Phi'(n)W_1(e', b') \quad (4.42)$$

Self 2. The first order conditions w.r.t.  $a_3$ ,  $e'$  and  $n$  are

$$\frac{u'(c_2)}{R} = \beta\delta \frac{\partial W_3(e', a_3, n)}{\partial a_3} \quad (4.43)$$

$$u'(c_2)n = \beta\delta \frac{\partial W_3(e', a_3, n)}{\partial e'} \quad (4.44)$$

$$u'(c_2)(1+g)h(e)x = \frac{\partial W_3(e', a_3, n)}{\partial n} \quad (4.45)$$

The Envelope Theorem gives the following condition

$$\frac{\partial W_2(a_2, e)}{\partial a_2} = u'(c_2) \quad (4.46)$$

$$\frac{\partial W_2(a_2, e)}{\partial e} = u'(c_2)(1+g)h'(e)(1-nx) \quad (4.47)$$

Self 1. The first order condition w.r.t.  $a_2$  and Envelope Theorem read

$$\frac{u'(c_1)}{R} = \beta\delta \frac{\partial W_2(a_2, e)}{\partial a_2} \quad (4.48)$$

$$\frac{\partial W_1(e, b)}{\partial e} = u'(c_1)h'(e) + \beta\delta \frac{\partial W_2(a_2, e)}{\partial e} \quad (4.49)$$

$$\frac{\partial W_1(e, b)}{\partial b} = u'(c_1) \quad (4.50)$$

From the FOCs and Envelope Theorem conditions w.r.t.  $a_2$  and  $a_3$ , we have the equation to describe consumption

$$u'(c_1) = \beta\delta R u'(c_2) = (\beta\delta R)^2 u'(c_3) \quad (4.51)$$

In steady states, we also have

$$u'(c_3)n = \Phi(n) \frac{\partial W_1(e, b)}{\partial b} = \Phi(n)u'(c_1) \quad (4.52)$$

Combining the two conditions above to eliminate consumption, we have

$$(\beta\delta R)^2 \frac{\Phi(n)}{n} = 1$$

□

**Corollary 14.** *Fertility in the time consistent problem is also independent of utility function and is determined by four parameters: hyperbolic discounting rate, time discounting rate, interest rate and altruism parameter.*

We can see that in both the commitment problem and the time consistent problem, fertility is determined by the altruism, interest rate, time discounting and hyperbolic discounting parameters. Wrede (2011) shows that, with a perfect financial market, the effects of hyperbolic discounting on fertility are ambiguous, depending on whether the children are considered as consumption or investment goods. The effects are negative when children are considered only as investment goods. The two results are consistent in the sense that pure altruism is essentially the same as considering children as investment goods. In pure altruism, a child's lifetime utility will enter her parent's utility function. Therefore, a parent stops sacrificing time and money after her children become independent, but the parent can continue to receive utility from her children thereafter. So in pure altruism, there is a time lag between giving resources and obtaining utility. The key feature of investment is exactly the same. So pure altruism is the same as considering children as investment goods for parents.

In the warm glow preference of Wigniolle (2013), the effects of hyperbolic discounting are different in developed and developing countries. Hyperbolic discounting tends to reduce fertility in developed countries but increase fertility in developing countries. Both our paper and Wigniolle (2013) assume that time costs are the only costs of raising children. The difference is that Wigniolle (2013) also assumes a non-wage income and that the cost of quantity of children is relatively higher with respect to family income in developed countries since women there will have more opportunities in the labor markets. The positive effects of hyperbolic discounting on fertility in

developing countries are possible only in the case that the parent gives her children zero education investment at the optimum. This case is quite rare in most developing countries.

We can use wage growth rates to distinguish between developed and developing countries since wage growth rates are higher in developing countries; when we do so, our model shows that hyperbolic discounting reduces fertility in both developed and developing countries. I think this result is more consistent with reality because fertility rates are decreasing in the majority of developing countries.

Last, there is no corner solution for fertility in my model. This is because the solutions are in steady states. Parents have the same life patterns as their children. If there are no children, equilibrium solutions in steady states will also disappear. Although many individuals choose not to have children, we can consider the agent in this model as a representative agent of the whole population. Then fertility in the model is the average of the whole society and will not have a corner solution.

**Proposition 12.** *In the time consistent problem, education is determined by the following equation*

$$h'(e^W) \left[ \frac{1}{R} + \frac{(1+g)(1-n^W x)}{R^2} \right] = 1 \quad (4.53)$$

*Proof.* Proof Combine the four conditions about education, we have

$$\begin{aligned} u'(c_2) &= \beta\delta \frac{\partial W_3(e', a_3, n)}{\partial e'} = \beta\delta\Phi(n) \frac{\partial W_1(e', b')}{\partial e'} = \beta\delta\Phi(n) \left[ u'(c_1)h'(e) + \beta\delta \frac{\partial W_2(a_2, e)}{\partial e} \right] \\ &\Rightarrow u'(c_2) = \beta\delta\Phi(n) [u'(c_1)h'(e) + \beta\delta u'(c_2)(1+g)h'(e)(1-nx)] \end{aligned} \quad (4.54)$$

Combining this equation with the equation describing the marginal utility of consumption, we have

$$1 = (\beta\delta R)^2 \frac{\Phi(n^W)}{n^W} h'(e) \left[ \frac{1}{R} + \frac{(1+g)(1-n^W x)}{R^2} \right] \quad (4.55)$$

$$1 = h'(e) \left[ \frac{1}{R} + \frac{(1+g)(1-n^W x)}{R^2} \right]$$

□

Again, there is no corner solution for education. In Wigniolle (2013), it is possible to have corner solutions for education. The reason is that the author assumes a free and compulsory public education. If we consider the whole society as a representative individual and consider the optimal education investment as the sum of public and private education investment, there is no corner solution. Public education subsidies are taxed from individuals.

**Corollary 15.** *Result In the time consistent problem, fertility is lower while education is higher:  $n^W < n^*$ ,  $e^W > e^*$ .*

*Proof.* Dividing the two equations from the two propositions about fertility, we have

$$\beta \frac{\Phi(n^W)}{n^W} = \frac{\Phi(n^*)}{n^*} \Rightarrow \frac{\Phi(n^W)}{n^W} > \frac{\Phi(n^*)}{n^*} \Rightarrow n^W < n^*$$

Education is determined by the same function form.

$$n^W < n^* \Rightarrow 1 - n^W x > 1 - n^* x \Rightarrow h'(e^W) < h'(e^*) \Rightarrow e^W > e^*$$

Again, let's assume the same utility function as in the model of commitment problem. With similar calculation, we have

$$c_1^W = \alpha_2 \frac{\epsilon(c_1^W)}{\epsilon(n^W) - \epsilon(c_1^W)} \left( (1+g)h(e^W)xR + e^W R - h(e^W) - \frac{(1+g)h(e^W)}{R} \right) \quad (4.56)$$

$$b^W = \frac{R^2}{R^2 - n^W} \left[ \left( (1+g)h(e^W)xR + e^W R \right) \left( \frac{\epsilon(c_1^W)}{\epsilon(n^W) - \epsilon(c_1^W)} + \frac{1}{R^2} \right) - \frac{\epsilon(n^W)}{\epsilon(n^W) - \epsilon(c_1^W)} \left( h(e^W) + \frac{(1+g)h(e^W)}{R} \right) \right] \quad (4.57)$$

$$\text{where } \alpha_2 = \frac{1}{1 + (\beta\delta R)^{\frac{1}{\gamma}} R^{-1} + (\beta\delta R)^{\frac{2}{\gamma}} R^{-2}}. \quad \square$$

Wigniolle (2013) also finds that fertility is not continuous in the entire range of possible fertility. For example, a small change in wages or education costs may lead to a jump in fertility. Our model does not predict this kind of discontinuity in fertility.

### 4.3 An Example

Let's assume that  $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ ,  $\Phi(n) = \frac{n^{1-\epsilon}}{1-\epsilon}$ ,  $h(e) = e^\sigma$ ,  $0 < \gamma, \epsilon, \sigma < 1$ . We will focus on the analysis of fertility and education since bequest is too complicated.



#### 4.3.1 Model 1: A Naive Agent's Commitment Problem

First, we can calculate fertility from the first proposition

$$\beta\delta^2 R^2 \frac{\Phi(n^*)}{n^*} = \beta\delta^2 R^2 \frac{(n^*)^{-\epsilon}}{(1-\epsilon)} = 1 \Rightarrow n^* = \left( \frac{\beta\delta^2 R^2}{1-\epsilon} \right)^{1/\epsilon}$$

It is easy to see that the equilibrium fertility is below the replacement rate if

$$\beta\delta^2 R^2 < 1 - \epsilon \Rightarrow \beta < \frac{1 - \epsilon}{\delta^2 R^2}$$

Then we can have education after fertility is determined,

$$\begin{aligned} h'(e^*) \left[ \frac{1}{R} + \frac{(1+g)(1-n^*x)}{R^2} \right] &= \sigma(e^*)^{\sigma-1} \left[ \frac{1}{R} + \frac{(1+g)(1-n^*x)}{R^2} \right] = 1 \\ \Rightarrow e^* &= \left[ \frac{\sigma}{R^2} \left( R + (1+g) \left( 1 - x \left( \frac{\beta\delta^2 R^2}{1-\epsilon} \right)^{1/\epsilon} \right) \right) \right]^{\frac{1}{1-\sigma}} \end{aligned}$$

#### 4.3.2 Model 2: A Sophisticated Agent's Time-Consistent Problem

The fertility of sophisticated agents is determined by the following equation

$$(\beta\delta R)^2 \frac{\Phi(n^W)}{n^W} = (\beta\delta R)^2 \frac{(n^W)^{1-\epsilon}}{1-\epsilon} = 1 \Rightarrow n^W = \left( \frac{\beta\delta R}{1-\epsilon} \right)^{2/\epsilon}$$

In this case, the equilibrium fertility is below the replacement rate if

$$\beta\delta R < 1 - \epsilon \Rightarrow \beta < \frac{1 - \epsilon}{\delta R}$$

Education is given by

$$e^W = \left[ \frac{\sigma}{R^2} \left( R + (1+g) \left( 1 - x \left( \frac{\beta\delta R}{1-\epsilon} \right)^{2/\epsilon} \right) \right) \right]^{\frac{1}{1-\sigma}}$$

### 4.3.3 Discussion

We can see that fertility is determined by interest rates, discount rates and the altruism parameter in both the commitment problem and the time consistent problem. This result is consistent with Becker and Barro (1988). The reason why absolute wages don't appear in the optimal fertility solutions is that time costs are the only costs of raising children. So the costs of raising children relative to family income are the same. Taxes will not be effective commitment devices to guarantee the fertility of naive agents, which are higher than those of sophisticated agents. If we include a fixed cost of raising children, taxation will be effective. Actually, a tax on childlessness was imposed for a long time in the Soviet Union and in other Communist countries. For many other countries with low fertility rates such as Japan and northern European countries, families with more babies can enjoy tax exemptions and various other subsidies. The reason for governments to engage in such interventions in family planning is that the social benefits of producing additional children are higher than the private benefits. So, interventions are needed to achieve the social optimum.

Interest rates play a key role in determining fertility rates. The result is quite straightforward. In a world without financial constraints, pure altruism implies that parents regard children as investment goods. Interest rates, being the rate of return on investments, will consequently play a key role in determining fertility. The conclusion is that higher interest rates will lead to higher fertility rates. The conclusion is consistent with reality. Interest rates and fertility rates in developing countries are both usually higher than those in developed countries.

## 4.4 Conclusions

In this paper, I have investigated the fertility decisions of agents with hyperbolic discounting and pure altruism preferences. I have also included education investment and bequest decisions. I consider the problem in two situations: the commitment problem of naive agents and the time-consistent problem of sophisticated agents. First, I find that, in both cases, fertility decreases because of hyperbolic discounting. The reason is that hyperbolic discounting discounts more than exponential discounting and the parent gains utility only after raising children.

I also find that sophisticated agents will choose to have fewer children and invest more into each of them, compared to naive agents. In the commitment problem, Self 1 makes all the choices. So there is no hyperbolic discounting between the second and third periods. That is to say, in the commitment problem, the time gap between the costs and benefits of raising children is smaller than that in the time consistent problem. When parents are sophisticated, which means that they can imagine the future, they will choose to have fewer children.

There are many reasons to explain the decline in fertility. For example, the increasing costs of raising children, widespread adoption of higher education of women and more accessible and affordable contraception methods. This paper tries to look at the issue from a new perspective and finds that it can explain the decline in fertility theoretically. I believe that the idea discussed in this paper illuminates one of the many reasons for the fertility decline. Also this paper contributes to the methodology of solving time inconsistent problems in the framework of pure altruism. Interestingly, the paper also finds that having children is essentially similar to investment for parents with pure altruism preferences.

There are several natural extensions of the paper. First, credit constraints can be introduced to relax the assumption of perfect financial markets. Second, the costs of raising children can be relaxed from linear to concave with respect to the number of children because the marginal time costs of raising children are decreasing; especially if the age gaps between children are small. Also, non-time costs can be included, too. These extensions may be suitable for future inquiry.

## BIBLIOGRAPHY

- Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97(6):1447–1458.
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.
- Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy*, 82(6):1063–1093.
- Becker, G. S. (1976). Altruism, egoism, and genetic fitness: Economics and sociobiology. *Journal of Economic Literature*, 14(3):817–826.
- Becker, G. S. and Barro, R. J. (1986). Altruism and the economic theory of fertility. *Population and Development Review*, 12:69–76.
- Becker, G. S. and Barro, R. J. (1988). A reformulation of the economic theory of fertility. *The Quarterly Journal of Economics*, 103(1):1–25.
- Becker, G. S., Kominers, S. D., Murphy, K. M., and Spenkuch, J. L. (2018). A theory of intergenerational mobility. *Journal of Political Economy*, 126(S1):S7–S25.
- Becker, G. S. and Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6):1153–1189.
- Behrman, J. and Taubman, P. (1985). Intergenerational earnings mobility in the united states: some estimates and a test of becker’s intergenerational endowments model. *The Review of Economics and Statistics*, 67(1):144–151.
- Cao, D. and Werning, I. (2016). Dynamic savings choices with disagreements. *National Bureau of Economic Research*, No. w22007.
- Cao, D. and Werning, I. (2018). Saving and dissaving with hyperbolic discounting. *Econometrica*, 86(3):805–857.
- Chan, T. W. and Boliver, V. (2013). The grandparents effect in social mobility: Evidence from british birth cohort studies. *American Sociological Review*, 78(4):662–678.
- Chan, T. W. and Boliver, V. (2014). Social mobility over three generations in finland: A critique. *European Sociological Review*, 30(1):13–17.

- Crumpler, H. and Grossman, P. J. (2008). An experimental test of warm glow giving. *Journal of Public Economics*, 92(5–6):1011–1021.
- Diamond, P. and Köszegi, B. (2003). Quasi-hyperbolic discounting and retirement. *Journal of Public Economics*, 87(9–10):1839–1872.
- Erola, J. and Moisiso, P. (2007). Social mobility over three generations in finland, 1950–2000. *European Sociological Review*, 23(2):169–183.
- Fehr, E. and Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism—experimental evidence and new theories. *Handbook of the Economics of Giving, Altruism and Reciprocity*, 1:615–691.
- Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *The Review of Economic Studies*, 60(1):35–52.
- Gobbi, P. E. and Goni, M. (2016). Childless aristocrats, fertility, inheritance, and persistent inequality in britain (1550-1950). Working Paper.
- Gruber, J. and Köszegi, B. (2001). Is addiction “rational”? theory and evidence. *The Quarterly Journal of Economics*, 116(4):1261–1303.
- Harbaugh, W. T., Mayr, U., and Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*, 316(5831):1622–1625.
- Hartmann, P., Eisend, M., Apaolaza, V., and D’Souza, C. (2008). Warm glow vs. altruistic values: How important is intrinsic emotional reward in proenvironmental behavior? *Journal of Environmental Psychology*, 52:43–55.
- Hodge, R. W. (1966). Occupational mobility as a probability process. *Demography*, 3(1):19–34.
- Karp, L. (2005). Global warming and hyperbolic discounting. *Journal of Public Economics*, 89(2–3):261–282.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478.
- Laibson, D. I., Repetto, A., Tobacman, J., Hall, R. E., Gale, W. G., and Akerlof, G. A. (1998). Self-control and saving for retirement. *Brookings papers on economic activity*, 1998(1):91–196.
- Lindahl, M., Palme, M., Massih, S. S., and Sjögren, A. (2015). Long-term intergenerational persistence of human capital an empirical analysis of four generations. *Journal of Human Resources*, 50(1):1–33.

- Lucas, R. E. and Kerr, S. P. (2013). Intergenerational income immobility in finland: contrasting roles for parental earnings and family income. *Journal of Population Economics*, 26(3):1057–1094.
- Mare, R. and Song, X. (2014). Social mobility in multiple generations. *UCLA CCPR Population Working Papers*.
- Møllegaard, S. and Jæger, M. M. (2015). The effect of grandparents' economic, cultural, and social capital on grandchildren's educational success. *Research in Social Stratification and Mobility*, 42:11–19.
- O'Donoghue, T. and Rabin, M. (1999). Doing it now or later. *American economic review*, 89(1):103–124.
- Ottoni-Wilhelm, M., Vesterlund, L., and Xie, H. (2017). Why do people give? testing pure and impure altruism. *The American Economic Review*, 107(11):3617–33.
- Peters, H. E. (1992). Patterns of intergenerational mobility in income and earnings. *The Review of Economics and Statistics*, 74(3):456–466.
- Phelps, E. S. and Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, 35(2):185–199.
- Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, 4(2):155–161.
- Shiue, C. H. (2016). Social mobility in the long run: An analysis with five linked generations in china, 1300-1900. *Working paper*.
- Simon, H. A. (1993). Altruism and economics. *The American Economic Review*, 83(2):156–161.
- Warren, J. R. and Hauser, R. M. (1997). Social stratification across three generations: New evidence from the wisconsin longitudinal study. *American Sociological Review*, 62(4):561–572.
- Wigniolle, B. (2013). Fertility in the absence of self-control. *Mathematical Social Sciences*, 66(1):71–86.
- Wrede, M. (2011). Hyperbolic discounting and fertility. *Journal of Population Economics*, 24(3):1053–1070.
- Zeng, Z. and Xie, Y. (2014). The effects of grandparents on children's schooling: Evidence from rural china. *Demography*, 51(2):599–617.

## APPENDIX. PROOF OF CHAPTER 2

### .1 The Proof of AR(1) Case

Plugging (1)(2)(4) into (5), we have

$$U_i = (1 - \alpha)\log(y_{i,t-1} - I_{i,t-1}) + \alpha u + \alpha\theta p \log I_{i,t-1} + \alpha p e_{it} \quad (.1)$$

FOC w.r.t. education investment can be written as

$$\frac{\partial U_i}{\partial I_{i,t-1}} = -\frac{1 - \alpha}{y_{i,t-1} - I_{i,t-1}} + \frac{\alpha\theta p}{I_{i,t-1}} = 0 \quad (.2)$$

Optimal investment can be obtained from equation (11)

$$I_{i,t-1} = \frac{\alpha\theta p}{1 - \alpha(1 - \theta p)} y_{i,t-1} \quad (.3)$$

Substituting (12) into (2), we have

$$S_{it} = \theta \log \frac{\alpha\theta p}{1 - \alpha(1 - \theta p)} y_{i,t-1} + e_{it} \quad (.4)$$

From equation (4), we have

$$\log y_{i,t-1} = u + p S_{i,t-1} \quad (.5)$$

Plugging (14) into (13)

$$S_{it} = \theta \log \frac{\alpha\theta p}{1 - \alpha(1 - \theta p)} + \theta u + \theta p S_{i,t-1} + e_{it} \quad (.6)$$

Lagging (15) by one generation and multiplying it by  $\lambda$

$$\lambda S_{i,t-1} = \lambda \theta \log \frac{\alpha\theta p}{1 - \alpha(1 - \theta p)} + \lambda \theta u + \lambda \theta p S_{i,t-2} + \lambda e_{i,t-1} \quad (.7)$$

Subtracting (16) from (15)

$$S_{it} - \lambda S_{i,t-1} = (1 - \lambda) \left[ \frac{\alpha\theta p}{1 - \alpha(1 - \theta p)} + \theta u \right] + \theta p S_{i,t-1} - \lambda \theta p S_{i,t-2} + e_{it} - \lambda e_{i,t-1} \quad (.8)$$

Combining (3) and rearranging the equation above, we can get

$$S_{it} = \delta + (1 - \lambda)\theta \left[ u + \frac{\alpha p}{1 - \alpha(1 - \theta p)} \right] + (\lambda + \theta p)S_{i,t-1} - \lambda\theta p S_{i,t-2} + v_{it} \quad (.9)$$

So, under the assumption of AR(1) process,  $S_{it}$  can be written as

$$S_{it} = \Omega_1 + (\lambda + \theta p)S_{i,t-1} - \lambda\theta p S_{i,t-2} + v_{it}$$

where  $\Omega_1 = \delta + (1 - \lambda)\theta \left[ u + \frac{\alpha p}{1 - \alpha(1 - \theta p)} \right]$ .

## .2 The Proof of AR(2) Case

Lagging (15) by one generation and multiplying it by  $\lambda_1$

$$\lambda_1 S_{i,t-1} = \lambda_1 \theta \log \frac{\alpha \theta p}{1 - \alpha(1 - \theta p)} + \lambda_1 \theta u + \lambda_1 \theta p S_{i,t-2} + \lambda_1 e_{i,t-1} \quad (.10)$$

Lagging (15) by two generations and multiplying it by  $\lambda_2$

$$\lambda_2 S_{i,t-1} = \lambda_2 \theta \log \frac{\alpha \theta p}{1 - \alpha(1 - \theta p)} + \lambda_2 \theta u + \lambda_2 \theta p S_{i,t-2} + \lambda_2 e_{i,t-1} \quad (.11)$$

Subtracting (19) and (20) from (15)

$$S_{i,t} = (1 - \lambda_1 - \lambda_2) \left[ \theta \log \frac{\alpha \theta p}{1 - \alpha(1 - \theta p)} + \theta u \right] + (\lambda_1 + \theta p)S_{i,t-1} + (\lambda_2 - \lambda_1 \theta p)S_{i,t-2} - \lambda_2 \theta p S_{i,t-2} + e_{it} - \lambda_1 e_{i,t-1} - \lambda_2 e_{i,t-1} \quad (.12)$$

Combining (7) and (21), we have

$$S_{i,t} = \delta + (1 - \lambda_1 - \lambda_2)\theta \left[ u + \log \frac{\alpha p}{1 - \alpha(1 - \theta p)} \right] + (\lambda_1 + \theta p)S_{i,t-1} + (\lambda_2 - \lambda_1 \theta p)S_{i,t-2} - \lambda_2 \theta p S_{i,t-3} + v_{it} \quad (.13)$$

So, under the assumption of AR(2) process,  $S_{it}$  can be written as

$$S_{i,t} = \Omega_2 + (\lambda_1 + \theta p)S_{i,t-1} + (\lambda_2 - \lambda_1 \theta p)S_{i,t-2} - \lambda_2 \theta p S_{i,t-3} + v_{it}$$

where  $\Omega_2 = \delta + (1 - \lambda_1 - \lambda_2)\theta \left[ u + \log \frac{\alpha p}{1 - \alpha(1 - \theta p)} \right]$ .