Essay in Family Economics and Media Economics in China

by

Yang Yue

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2017 by the
Graduate Supervisory Committee:

Daniel Silverman, Chair
Natalia Kovrijnykh
Gregory Veramendi

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

Family economics uses economic concepts such as productions and decision making to understand family behavior. Since Gary Becker introduced household decisions into family, economists began to place emphasis on the rule of families on labor supply, human capital investment, and consumption. In a household, the members choose the optimal time allocations between working, housework and leisure, and money between consumption of different members and savings. One-Child policy and strong inter-generational connections cause unique family structure in China. Households of different generations provide income transfer and labor support to each other. Households consider these connections in their savings, labor supply, human capital investment, fertility and marriage decisions. Especially, strong intergenerational relationships in China are one cause of the high level of young female labor supply and high saving rate. I will investigate the rules of intergenerational relationships on household economic behavior.

Affirmative Action allocates college seats to a separate group. To evaluate the distribution effects of AA on discrete groups, we need to study household's strategic reactions on the rule of college seats allocation. The admission system of National College Entrance Examination (NCEE) in China is a type of AA. That distributes college seats by regions. I will use the rapid expansion of Chinese college enrollment as a natural experiment to check the households' reaction on AA and college expansion.

Media economics utilizes economic empirical and theoretical tools to figure out the social, cultural, and economic issues in media industries. The impact of online piracy on genuine products sales is under debate, because people cannot find representing proxies to evaluate piracy levels. I will use Chinese data to study the effects of online piracy on theater revenue.

ACKNOWLEDGMENTS

I would like to extend our warm appreciations and gratitude to the following persons for making this project possible. Their supports, inspirations, criticisms, comments and suggestions contribute a valuable strength to make this paper presentable and knowledgeable.

I would to acknowledge our helpful and over-supportive committee, Prof. Daniel Silverman, Prof. Natalia Kovrijnykh, Prof. Gregory Veramendi, and Prof. Matthew Wiswall, gave us the ideas to find the best solutions to our dilemma and taught us on how to make a good paper. I would also like to thank them for motivating me to do our best shot as well as for her pieces of advice on how to improve it.

To my dear parents, wife and children, who are supporting me of having group activities outside and inside the campus and for lending us financial support for the printing matters. Moreover, they showed their love and gave us words of encouragement that pushed us to do these well. I'm also immensely grateful to my friends and classmates for their comments on my manuscript and presentations. Any errors are my own and should not tarnish the reputations of these esteemed persons.

To those who are not mentioned but whose help are truly contributing in adding some significant information will not be forgotten, I thank you.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1.

INTERGENERATIONAL CONTRACTS AND FEMALE LABOR SUPPLY

The strong intergenerational relationships in China are one reason of the high level of young female labor supply and high saving rate. In Chapter 1, I examine how intergenerational relationships between parents and grandparents affect females' labor supply and households savings. I develop a non-altruistic dynamic contract model using economic benefits such as a bequest, coinsurance, and cheaper care service to sustain such relationships in the face of long term incentive problems. I then estimate the parameters of the model using Chinese household surveys. I evaluate the labor and income reallocation effect throughout the relationships. I find that intergenerational relationships increase the labor supply of younger females by 32% but reduce the labor supply of older females by 21%, while increasing older females' household savings by 13%. My policy experiments produce the following predictions: delaying retirement age reduces the labor supply of young females; raising inheritance taxes increases the labor supply of young females and savings of both parents and grandparents. Therefore, I find that public policy affects the households attached to the target group through intergenerational relationships.

1.1 Introduction

This article examines the influence of intergenerational relationships on females' labor supply decisions and households' reactions to a number of policies. Intergenerational relationships build economic connections between households of various generations. In the relationships, households of different generations provide income transfer, child care and elder care to each other. These activities affect households' labor supply and saving decisions by redistributing labor and income across generations. Furthermore, economic ties matter in evaluating the effects of government policies. Through intergenerational relationships, public policies are not only affecting the target groups, but also affecting the households attached to the target groups. In countries without strong social welfare programs[1], intergenerational relationships are the main way to provide elder support, child care and family insurance. For example, 66% of the Chinese elderly (aged 65 and over) provides child care to their grandchildren (Wu et al., 2014), 45% of the Chinese elderly live with their children and 22% of their income come from their children (National Survey Research Center, 2014). In these countries, intergenerational relationships substantially reshape households' decisions on savings and labor supply, as well as policy implications of various public policies.

I construct a theoretical framework to analyze the incentive problems in intergenerational relationships. In an intergenerational relationship, households exchange income and labor service in different periods. The exchange is not balanced within each period. Without incentives to keep households committed to the relationship at each stage, the sequential exchange will not happen from the beginning. For example, grandparents take care of their grandchildren to exchange parents' elder support in the future. However, when grandparents are old, parents refuse to support grandparents, if they cannot get benefits from it. Without the confirmation of payback, grandparents will not help parents from the

[1] Table 5 in the Appendix compares intergenerational relationships across countries.

beginning. Considering intergenerational relationships provide incentives for the sequential exchange is a prerequisite for evaluating the impact of intergenerational relationships on household behavior. I build a dynamic contract model using economic benefits to keep the households in the relationships.

I quantify the effect of intergenerational relationships on females' labor supply over the life cycle. The relationships reallocate labor across generations by affecting households' child and elder care decisions. Furthermore, grandparents' help free young parents from child care and enable them to stay in their jobs. However, to assist parents, grandparents may work less and leave their jobs early. Moreover, as grandparents grow old, parents provide elder care service to grandparents and work less. An intergenerational relationship has various effects on females' labor supply at various stages. I use a structural model to estimate the labor reallocation effects of each stage of the relationship.

I measure the spillover effects of public policies through intergenerational relationships. The spillover effects can change the policy implications of public policies. For example, a government delays the mandatory retirement age to increase seniors' labor supply. Then, grandparents, who used to take care of their grandchildren, need to get back to work. As the result, parents reduce their working time to take care of their children. Thus, delaying mandatory retirement age reduces the labor supply of young females through intergenerational relationships. Ignoring the economic connections between generations could lead to incomplete forecasts of the impact of some policy change. I conduct policy experiments to evaluate various public policies with the existence of intergenerational relationships.

The first contribution of the article is to develop a non-altruistic dynamic contract framework that uses economic benefits to sustain long term intergenerational relationships. The contract model is built by adding labor supply and preference for child and elder care, as well as bequest to the Kocherlakota (1996) environment[2]. The dynamic contract framework solves the incentive problems of the relationships and creates Pareto improvements for the two households by exchanging income and labor service. The economic benefits from labor allocation, risk sharing, and bequests give households utility values higher than those of outside options at any stage. Furthermore, the Pareto gains provide incentives for the households to remain in the relationships. The model also reveals the effects of incentive problems on household behaviors. In the contract without incentive problems, households have the same behaviors as that in an altruistic model. The contact gives a constant utility weight on each household and allocates the endowment according to the weight. In the paper, I use the contract without incentives problem to show households' behaviors in altruistic models. I can use the dynamic contract framework to derive of the rules of income and labor service reallocation in intergenerational relationships. On the basis of the rule, I identify the effect of intergenerational relationships on savings and labor supply. These expressions enable the formal identification proof and promote the estimation of the structural parameters.

The second contribution of the article is to quantify the effect of intergenerational

[2] Kocherlakota (1996) built a dynamic risk sharing problem between two risk averse agents living in infinite horizon and facing idiosyncratic income shocks. In the paper, risk sharing is limited by two-sided lack of commitment to the insurance contract.

relationships on the female labor supply in China. I implement the identification strategy and estimate the model using data from Chinese household surveys. I choose these surveys for three major reasons. First, China has strong intergenerational relationships. Observing and estimating the influence of relationships on households' decisions is easy. Second, several major social policy changes have occurred in recent decades. These changes provide enough variations to identify the effects of economic condition changes on intergenerational relationships is easy. Finally, these surveys have detailed information about income transfer, child care and elder care. The information enables us to identify the economic connections between households. I can use this information to measure the extent to how households conduct child and elder care, as well as transfer decisions. The results show that grandparents taking care of their grandchildren increases the labor supply of parents by 32%, but decreases the labor supply of grandparents by 21% in the earlier stage of the relationships. In providing elder care to grandparents, parents reduce their labor supply by 13% at the later stage of the relationships. The wage structure in China contributes to the strong effects of intergenerational relationships on labor supply. Given that younger generations have a higher wage rate, the total income of two households increases when grandparents take care of their grandchildren and free parents to work. I also compare the differences of households' behavior between the contracts with and without incentives problem. I found a bad income shock for grandparents, for example, can increase the transfer and support from parents to grandparents in coming periods in the contract without incentive problems, but decreases that in the contract with incentive problems.

The third contribution of the article is to evaluate the spillover effects of various policies through intergenerational relationships. My policy experiments reveal the effect of social welfare programs on female labor supply and savings as well as on households' transfer, elder and child care decisions. I impose a 20% subsidy on the child care and elder care services from market. I find that 20% of child care subsidy increases grandparents' labor by 41% and 20% elder care subsidy increases parents' labor supply by 13%. I then increase the inheritance tax from 0% in benchmark to 30%. The direct effects of inheritance tax reduce grandparents' bequests in exchange for parents' labor service and income transfer. Parents reduce the income transfer and labor support to grandparents. Grandparents increase their savings to maintain the bequest incentives. The change increases grandparents' savings by 20% and increases parents' labor supply by 9%. Finally, I delay the mandatory retirement age from 60 in benchmark to 65. I find that pushing back the mandatory retirement age reduces the parents' labor by only 8% at the delaying period when grandparents have a lower wage rate than parents. The wage structure determines that grandparents are always the child care providers before and after the policy change. Then I change the wage structure by reducing the wage gap between parents and grandparents. I find that delaying mandatory retirement age decreases the labor supply of parents by 43% at the delaying period. With the new wage structure, grandparents no longer provide child care. Therefore, the spillover effects through intergenerational relationships change the implications of these public policies.

1.2 Literature

Motives of Intergenerational Relationships Since Barro (1974) and Becker (1974), researchers mainly use altruism[3] to address the incentive problem in intergenerational relationships. Altruism sustains long term intergenerational relationships, but cannot explain several phenomena. For example, parents account for the relative economic positions of their children and transfer wealth to or share their inheritance with their children unequally (Schanzenbach & Sitkoff 2008). Although altruistic parents are expected to give more to their less well-off children, bequests tend to divide equally among siblings (McGarry 2001). As altruism is morally charged, it should be independent of institutions and economic factors. Many studies have found that family ties in countries with weak social welfare programs tend to be stronger than those countries with strong programs (Bonsang 2007; Hank & Buber 2008). Altruistic models fail to explain the relationship of intergenerational ties with these economic factors. I use a non-altruistic model to emphasize the functions of observable economic factors on intergenerational relationships.

Some studies use non-altruistic forces to address the incentive problem in intergenerational relationships. Researchers have argued that households enforce the intergenerational ties through a self-enforcing constitution[4] (Cigno 2006), demonstration effect (Jellal & Wolff

[3] Altruism indicates that a parent (child) can derive utility from the consumption of his child (parent).

[4] Cigno (2006) relies on social norms explain long-term relationships. Families can be viewed as communities governed by self-enforcing constitutions. In the OLG framework, if people do not support their parents, their children will be not willing to support them. In this environment, the dominant strategy of individuals is to provide transfer and help to other people.

2005) or nurtured altruism (Stark & Zhang 2002). However, these studies ignore the direct economic benefits from intergenerational relationships. Direct economic benefits can sustain intergenerational relationships. Intergenerational relationships improve labor allocation efficiency by allowing those who are more productive to work (Geurts, van Tilburg, Poortman & Dykstra 2015). Furthermore, assets after unexpected death are an important source of bequest (Lockwood 2014). Accidental bequest has no direct effect on parents' utility, but is used in exchange for children's transfer and support. The Pareto improvement through risk sharing, cheap care service and bequest creates economic benefits for intergenerational relationships. My model combines these direct economic benefits to sustain intergenerational relationships.

<u>Female Labor Supply Affected by Intergenerational Relationships</u> The provision of child and elder care reduces females' labor supply. Mothers face the problem of reconciling work and child care responsibilities. Grandparents may substitute for mothers in doing child care work and thus mothers become free to work. Grandparents' assistance strongly increases younger parents' labor supply (Posadas & Vidal-Fernandez 2012; Compton & Pollak 2014, Battistin, Nadai &Padula, 2015). Within intergenerational relationships, the question of who provides child and elder care is primarily determined by individuals' health and income conditions. Working grandparents use more money to subsidize their grandchildren rather than provide care directly (Luo, LaPierre Hughes & Waite, 2014). Grandparents with newborn grandchildren are more likely to provide care for their grandchildren, and

married grandparents are also more likely to work and to provide financial help (Ho 2015). In the later periods of intergenerational relationships, children provide income support and physical care to their parents. In most developing countries, most of the elderly population receives financial support from their adult children (Hamaaki, Hori & Murata 2014). Parents who provide care for grandparents cause the great reduction of female labor supply at midlife (Johnson & Sasso 2006). The literature focuses mostly on the effects of intergenerational relationships in a single period. Conversely, this article determines the influence of intergenerational relationships on females' labor supply throughout the life cycle.

Formal care service and social welfare programs can substitute for intergenerational relationships by providing the same service to households. The expansion of public child care provokes a large positive effect on maternal employment (Bauernschuster & Schlotter 2015). Married women's labor supply decreased with the ascending cost of formal child care (Blau & Kahn 2005). Elder care giving decreases female work intensity and the impact decreases after the launch of the market oriented elder care insurance (Sugawara & Nakamura 2014). The Temporary Assistance for Needy Families program undermines intergenerational support (McDonald & Armstrong 2001). The implementation of National Health Insurance in Taiwan in 1995 decreased the likelihood of intergenerational coresidency (Hsieh, Chou, Liu & Lien 2015). In this article, formal care service and social welfare programs affect the outside options of intergenerational relationships.

1.3 Background

China has a high female labor force participation rate. The high female labor force participation rate is a legacy of the Communist Party's rule that women are equal to men in all spheres of life (Yu & Liu, 2010). Yet, since the 1980s, the transition to market economy has widened the gender income gap[5]. Between 1990 and 2014, Chinese females' labor force participation rates declined from 77% to 64%. But, the rate is still higher than the world average of 50% (World Bank 2016).

Strong intergenerational connections increase the labor supply of young females (Chen & Liu 2009). Grandparents who care their grandchildren are common in most families in China (Chen, Liu & Mair 2011). Grandparents provided child care to grandchildren in 35% of family setups in rural China (Silverstein, Cong, & Li 2006). In 2010, about 66% of people older than 60 years of ages have provided care for their grandchildren (Melenberg & Zheng 2012). The provision of child care by grandparents affects both parents' and grandparents' labor supply decisions. For example, the participation of daughters' in the labor force is one major reason why grandmothers provide child care (Chen, Liu, & Mair 2011). Traditionally, children in China bear the ultimate responsibility for taking care of their aging parents (Chen & Liu 2009). An adult child faces criminal charges for refusing

[5] According to National Bureau of Statistics of China, in 1990, the average female's wage rate was 78% in rural area and 79% in urban area of that of the average male. In 2013, the ratio declined to 67% in urban areas and 57% in rural areas. .

to support an aged parent[6]. In China, elders in most areas do not have a formal safety net. The majority depends exclusively on their children for support (Cong & Silverstein 2012). About 45% of people older than 60 years live with their children and 22% of their income comes from their children (National Survey Research Center, 2014).

The low fertility rate caused by the one child policy[7] contributes to the high female labor participation rate. In a unique "four-two-one" family[8], the only child receives child care from four grandparents in childhood, and also bears the responsibility of supporting two parents and, sometimes, four grandparents in their old age. Since the enforcement of the one child policy in 1980s, the fertility rate in China decreased from 3 in 1980 to 1.6 in 2015, which is lower than the world average of 2.6 (World Bank 2015). The low fertility rate reduces females' burden of child care and causes a high female labor market participation rate.

Weak institutionalized care and social welfare programs also contribute to the high female labor participation rate. After China's economic transition in the 1980s, publicly funded care[9] and elder care[10] is largely eliminated, and market care service is either too

[6] The Chinese constitution of 1982 proclaims the obligation of adult children to support their elderly parents.

[7] The one-child policy, introduced in 1979, only allowed families to have one child each. Since 1984, a rural family can have a second child if the firstborn is female (Chen, Jin & Yue 2010). Since 2014, all couples can have second children. See Figure 6 in the Appendix for the details of the fertility rate change.

[8] In a "Four-two-one" family, the child is the only child for two parents and the only grandchild for four grandparents.

[9] The number of publicly funded kindergartens dropped from about 150,000 (an 83% market shares) in 1998 to about 43,000 (a 24% market share) by in 2012 (National Bureau of Statistics 2016).

[10] China has just about 2% of people ages sixty-five and older living in residential care facilities. (Feng, Liu, Guan & Mor 2012).

expensive for most households to afford or suffer from low quality (Zhang & Maclean 2012). In 2003, accordance with the National Research Center on Aging, only 2% of the population aged 65 and over use institutionalized care. In addition, only 46% of urban employees were covered by a pension plan in 2004 (Trinh 2006), and only 12% of the rural labor force participated in the old age social insurance programs scheme in 2006 (Wang 2006). Without strong institutionalized care and social welfare programs, households can only rely on intergenerational relationships to provide child care and elder support.

1.4 Theoretical Model

In this section, I present a dynamic contract model of inter-household decision making. Households can live independently, or join a contract through the mutual provision of a series of state contingent income transfers, elder care, and child care to each other. I define the case without interactions as the autarky case, and that with interactions as contracts. Households remain in the contract, purely because of economic benefits they can gain. In contrast to the existing literature, the model does not take altruism into consideration. Ignoring altruism does not mean denying the importance of altruism, but doing so highlights the functions of observable economic factors in intergenerational relationships.

The model has two agents: parent and grandparent. The parent's household has one child, who makes no decisions and needs child care service. To simplify, I only look at the female supply decisions, and take the male's labor supply and income as exogenous. For notational convenience, I denote the age of the parent and the calendar year by t. I assume the grandparent is 6 periods older than the parent. At period t, the parent's age is t and the grandparent's age is t+6. An agent lives for 20 periods at most. At age t, an agent's death rate is ϱ_t. At time t, the parent's death rate is ϱ_t and the grandparent's death rate is ϱ_{t+6}. In simplifying the model, the parent's death rate is 0 before period 14. The parent will not die before the grandparent. The agent needs child care from age 1 to 4. From age 11 to 20, the agent needs elder care. The agent retires after period 9. After age 10, the agent no longer provides elder and child care as well as work[11].

<u>Preferences</u> Household i has preference on consumption c_t^i, leisure l_t^i, child care hours K_t^i, and elder care hours N_t^i. Child care service can come from the parent k_t^p, the grandparent k_t^g, or an outside service k_t^m, with $K_t^i = k_t^p + k_t^g + k_t^m$. Elder care service can come from parent n_t^p, and outside service n_t^o, with $N_t^i = n_t^p + n_t^m$. η represents the preference over leisure. α_t represents the preference over child care. γ_t represents the preference over elder care. α_t and γ_t, the tastes over child care and elder care service, change over time. α_t equals to 0 after age 4. γ_t equals to 0 when the household is younger than age 10. At time t, the parent's utility parameters are $\theta_t, \eta_t, \alpha_t$ and γ_t; the grandparent's utility parameters are, θ_{t+6}, η_{t+6}, α_{t+6}, and γ_{t+6}. Agent i's current-period utility function at age t is:

$$U_{it}(c_t^i, l_t^i, K_t^i, N_t^i) = \ln c_t^i + \eta \ln l_t^i + \alpha_t \ln K_t^i + \gamma_t \ln N_t^i \qquad (1..1)$$

[11]Retired agent can still provice child care at age 10.

Budget Sets Households take the price of outside child care and elder care service p_t^k and p_t^n, wage rate w_t^i, and the interest rate R_t, as given. The budget constraint $\left(BC_t^i\right)$ is:

$$c_t^i + s_{t+1}^i + p_t^k k_t^m + p_t^n n_t^m + T_t^i \leq R_t s_t^i + w_t^i h_t^i + \epsilon_{kt}^i, \forall i \in \{p, g\} \tag{BC_t^i}$$

The money endowment is from asset $R_t s_t^i$, wage income $w_t^i h_t^i$, which is determined by wage rate w_t^i and working hours h_t^i, and a random income shock[12] ϵ_{lt}^i. ϵ_{jt}^i with probability π_{tj}^p, $\sum_{j=1}^{J} \pi_{jt}^p = 1$, and $\epsilon_{jt}^p \in \left\{\epsilon_{1t}^p, ..., \epsilon_{Jt}^p\right\}$; ϵ_z^g with probability π_{zt}^g, $\sum_{z=1}^{Z} \pi_{zt}^g = 1$, and $\epsilon_{zt}^g \in \left\{\epsilon_{1t}^g, ..., \epsilon_{Zt}^g\right\}$. I assume, $w_t^p + \epsilon_{1t}^p > 0$,and $w_t^g + \epsilon_{1t}^g > 0$. The assumption avoids corner solutions. Household i spends the money on consumption c_t^i, saving for the next period s_{t+1}^i, outside child care k_t^o and elder care n_t^o, and net transfer to the other household T_t^i.

Time Allocation Each household's overall time is 1. Household i can spend time on work h_t^i, leisure l_t^i, elder care n_t^i and child care k_t^i. The time constraint $\left(TC_t^i\right)$ is:

$$1 = h_t^i + l_t^i + n_t^i + k_t^i, \forall i \in \{p, g\}. \tag{TC_t^i}$$

Autarky Case I define the en-ante value function in autarky case and call this autarky value $V_{it}^{aut}(s_t^i)$. After the income shock is realized, each household makes an optimal allocation of current-period consumption, individual savings s_{t+1}^i, carried on to the next period, labor-force participation decision h_t^i, child care decisions k_t^i and k_t^o, and elder care decisions n_t^i and n_t^o, which solves the following optimally constrained problem. Define the set of decisions

[12] In this paper, income shocks are from the uncertainty of not working income and health spending. For example, a man has a certain probability of getting a disease every year. If he gets sick, he needs to pay for medical treatment, which is a bad income shock for him. If he is healthy, he doesn't need to spend money on medical treatment, which is a good income shock for him.

made as $\Omega_{ti} = \left\{s^i_{t+1}, h^i_t, n^i_t, k^i_t, n^o_t, k^o_t\right\}$. The Bellman equation of the autarky case is[13]:

$$v_t(s^i_t, \epsilon^i_{kt}) = \max_{\Omega_{ti}} U_{it}(c^i_t, l^i_t, K^i_t, N^i_t) + \beta\left(1 - \varrho^i_{t+1}\right) V^{aut}_{t+1}(s^i_{t+1}), \forall i \in \{p, g\} \tag{1..2}$$

subject to budget constraint BC^i_t and time constraint TC^i_t. In this case, the solution to the problem above yields the following expected value function for the household at the beginning of period t. The autarky value is:

$$V^{aut}_{it}(s^i_t) = \sum_{j=1}^{J} \pi^i_{jt} v_t(s^i_t, \epsilon^i_{jt}) \tag{1..3}$$

The Optimal Contract A contract determines on bequest, punishment, income transfer, child and elder care, savings and consumption. A contract can end in two cases. If one side breaks the rule of the contract, the contract will end as a punishment. If the grandparent dies in the contract, the contract will end. At the same time, the parent receives the grandparent's savings as a bequest and lives in the autarky case. In the contract, each household leaves the relationship at any time. I call this the no-commitment case. I also define a special case that neither household can leave after both households join in the contract. I call this the full-commitment case. No incentive problems exist in this case. I use full-commitment contract as a benchmark to show the first-order importance of the incentive problems.

In the initial period t=0 of a contract, both households choose to either join in the contract or live alone. If households cannot reach an agreement, each household will live in the autarky case forever after period 0. If they reach an agreement, they go to the next

[13] ϱ^i_{t+1} is the death rate of agent i, with $\varrho^p_{t+1} = \varrho_{t+1}$ and $\varrho^i_{t+1} = \varrho_{t+7}$.

period. Before period t=1 and onwards, the death shock happens. If the grandparent dies, the parent receives the all the savings of the grandparent and lives in the autarky case. If neither household dies, both households enter the contract of the period. Taking the income as given, the contract gives the promised support and transfer to each other, as well as makes the consumption, time allocation and savings decisions. (see Figure 1 in the Appendix for the timing of the contract)

The rule of income and labor service allocation is adjusted to an allocation that lies on the Pareto frontier. The state space is comprised of each household's asset s_t^i and the grandparent's promised value G_t. I define the parent's value function $P_t(s_t^p, s_t^g, G_t)$. In solving for the optimal contract, one maximizes $P_t(s_t^p, s_t^g, G_t)$ subject to delivering at least G_t the grandparent. Following Spear and Srivastava (1987), one can rewrite the sequence problem corresponding to the optimal contract in recursive form, with the promised value as a state variable, and continuation values to the grandparent as control variables. Essentially, the promised value summarizes the previous history of the play.

<u>No-commitment Case</u> A contract determines the net transfer T^i_{tjz}, with $T^p_{tjz} = -T^g_{tjz}$, child care support k^g_{tjz}, and elder care support n^p_{tjz} in the state with income shocks ϵ^p_j and ϵ^g_z. I define the set of decisions made on support and transfer as $\Gamma_{tjz} = \left\{ n^p_{tjz}, k^g_{tjz}, T^p_{tjz} \right\}$. According to the contract, contracts makes an optimal allocation of current-period consumption c^p_{tjz} and c^g_{tjz}, as well as the time allocations to each household such as labor-force participation decision h^i_{tjz}, child care decisions k^i_{tjz} and k^m_{tjz}, elder care decisions n^i_{tjz} and n^m_{tjz}. I define

the set of decisions made on care giving as $\digamma_{tjz} = \left\{h^g_{tjz}, n^m_{tjz}, h^p_{tjz}, k^p_{tjz}, k^m_{tjz}\right\}$. In addition, the contract determines the individual savings s^i_{t+1jz}, and the promised value G^{t+1}_{jz}. I define the set of decisions made on state variables as $\Omega_{tjz} = \left\{s^g_{t+1jz}, s^p_{t+1jz}, G^{t+1}_{jz}\right\}$. The Bellman equation is:

$$P^t(s^p_t, s^g_t, G^t) = \max_{\{\digamma_{tjz}, \Omega_{tjz}, \Gamma_{tjz}\}} \sum_{j=1}^{J}\sum_{z=1}^{Z} \pi^p_{jt}\pi^g_{zt}[U_{pt}(c^p_{tjz}, l^p_{tjz}, K^p_{tjz}, N^p_{tjz})$$

$$+\beta\left(1-\varrho_{t+7}\right)P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz}) + \beta\varrho_{t+7}V^{aut}_{pt+1}(s^p_{t+1jz} + s^g_{t+1jz})]$$

subject to budget constraints BC^p_t and BC^g_t, time constraints TC^p_t and TC^g_t, the promise keeping constraint at period t:

$$\sum_{j=1}^{J}\sum_{z=1}^{Z} \pi^p_{jt}\pi^g_{zt}\left[U_{gt}(c^g_{tjz}, l^g_{tjz}, K^g_{tjz}, N^g_{tjz}) + \beta\left(1-\varrho_{t+7}\right)G^{t+1}_{jz}\right] \geq G^t \qquad (PK^g_t)$$

the incentive constraints[14] given any income shocks at period t:

$$U_{gt}(c^g_{tjz}, l^g_{tjz}, K^g_{tjz}, N^g_{tjz}) + \beta\left(1-\varrho_{t+7}\right)G^{t+1}_{jz} \geq v^{aut}_{gt}(s^g_t, \epsilon^g_{zt}), for \forall j, z \qquad (IC^g_t)$$

and

$$U_{pt}(c^p_{tjz}, l^p_{tjz}, K^p_{tjz}, N^p_{tjz}) + \beta\left(1-\varrho_{t+7}\right)P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})$$

$$+\beta\varrho_{t+7}V^{aut}_{pt+1}(s^p_{t+1jz} + s^g_{t+1jz}) \geq v^{aut}_{pt}(s^p_t, \epsilon^p_{jt}),\ \ for\ \forall j, z \qquad (IC^p_t)$$

the participation constraints at period t+1:

$$G^{t+1}_{jz} \geq V^{aut}_{gt+1}(s^g_{t+1jz}), for \forall j, z \qquad (PC^g_{t+1})$$

[14] The incentive constraints are indeed the ex-post participation constraints. To simplify the notation, I call these constraints as incentive constraints.

and

$$P^{t+1}(s^{p}_{t+1jz}, s^{g}_{t+1jz}, G^{t+1}_{jz}) \geq V^{aut}_{pt+1}(s^{p}_{t+1jz}),\ \ for\ \forall j, z \qquad (PC^{p}_{t+1})$$

The promise-keeping constraint (PK^{g}_{t}) ensures that the contract delivers the promised level of discounted utility to the grandparent. It plays the role of a law of motion for the state variables. The incentive constraint for household i (IC^{i}_{t}) is the incentive compatibility constraint ensuring that the household i gets a higher ex-post utility value from the contract than it could from the autarky case. The participation constraint for household i (PC^{i}_{t+1}) is the incentive compatibility constraint ensuring that household i gets a higher ex-ante utility value from the contract, than it could from the autarky case in the next period.

Full-commitment Case In a full-commitment case, no incentive problems exist. The Bellman equations of no-commitment contract doesn't have participation constraints, and incentive constraint.

Characterization of the Optimal Contract

Full-commitment Case This section characterizes the optimal contract. λ_t is the Lagrangian multiplier associated with promise keeping constraint of the grandparent household. Using the envelope theorem, I get:

$$G^{t} : \frac{\partial P^{ft}(s^{p}_{t}, s^{g}_{t}, G^{t})}{\partial G^{t}} = -\lambda_t \qquad (ET1)$$

The first order conditions for the optimal contract problem are:

$$T_t^p \quad : \quad \frac{c_{tjz}^g}{c_{tjz}^p} = \lambda_t \tag{$FOC2$}$$

$$G_{jz}^{t+1} \quad : \quad \frac{\partial P^{f,t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}} = -\lambda_t \tag{$FOC3$}$$

Proposition 1 *In the full-commitment case, both households fully share the income risk. The consumption ratio is a constant through the all periods.* $\frac{c_{tjz}^g}{c_{tjz}^p} = \lambda_t$, *with* λ_t *is a constant, for* $\forall t \in [1, T]$, $\forall \epsilon_{jt}^p$ *and* ϵ_{zt}^g*.(see Appendix for proof.)*

Two households fully share risks in full-commitment case. Equation ET1 defines the ex-ante income and labor service allocation rule. Equation FOC2 defines the ex-post income and labor service allocation in the state with income shocks ϵ_{zt}^g and ϵ_{jt}^p. Equation FOC3 defines the ex-ante income and labor service allocation rule of at time t+1. The three equations show that the full-commitment case always gives constant utility weights $\frac{\lambda_t}{\lambda_t+1}$ to the grandparent's utility, and $\frac{1}{\lambda_t+1}$ to the parent's utility at any state of any period. In addition, $\lambda_t = \lambda_1$ for $\forall$ t. The utility weight on each term is a constant over time. To give an intuition of the result, consider two households maximizing their ex-ante utilities. The contract allows the household in a relatively good state to transfer its income and labor service to the household that is in a relatively bad state. Two households smooth their consumption by allocating a constant portion of the income and labor service to each other. Fully risk sharing increases both households' ex-ante utility values.

An optimal contract with full-commitment is equivalent to a single household problem with special altruistic preferences. In the single household problem, the parent and the grandparent have the same preference for the parent's utility relative to the grandparent's utility. The preference for each household becomes:

$$\nu_t = U_{pt}(c^p_{tjz}, l^p_{tjz}, K^p_{tjz}, N^p_{tjz}) + \lambda_1 U_{gt}(c^g_{tjz}, l^g_{tjz}, K^g_{tjz}, N^g_{tjz}) \tag{1..4}$$

The incentive problems in intergenerational relationships cause the differences of households' behaviors between the intergenerational contract model and altruistic models. Without the incentive problems, households in an optimal contract behave like altruistic households. The only difference between the contracts with full-commitment and altruistic problems is on the initial utility weight λ_1. λ_1 Is endogenous chosen in contract, but exogenous given in the problem with altruistic households. In this paper, I use the optimal contract with full-commitment to demonstrate households' behavior with altruistic preference.

No-commitment Case I characterize the optimal contract using the first order conditions and envelope theorem. λ_t is the Lagrangian multiplier associated with the grandparent household's promise-keeping constraint. $\pi^p_{jt}\pi^g_{zt}\theta_{jz}\left(1-\varrho_{t+7}\right)$ are the Lagrangian multipliers associated with grandparent's participation constraint. $\pi^p_{jt}\pi^g_{zt}\mu_{jz}\left(1-\varrho_{t+7}\right)$ are the Lagrangian multipliers associated with the grandparent household's participation constraint. $\pi^p_{jt}\pi^g_{zt}\varkappa_{jz}$ are the Lagrangian multipliers associated with the parent household's participation constraint. $\pi^p_{jt}\pi^g_{zt}\omega_{jz}$ are the Lagrangian multipliers associated with the parent household's incentive constraints. Using the envelope theorem, I get:

$$G^t : \frac{\partial P^t(s^p_t, s^g_t, G^t)}{\partial G^t} = -\lambda_t \tag{$ET2$}$$

From the first order conditions (see Appendix for details) for the optimal contract problem, I get:

$$T_t^p \quad : \quad \frac{c_{tjz}^g}{c_{tjz}^p} = \frac{\lambda_t + \varkappa_{jz}}{1 + \omega_{jz}} \qquad (FOC4)$$

$$G_{jz}^{t+1} \quad : \quad \frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}} = -\frac{\lambda_t + \varkappa_{jz} + \theta_{jz}}{1 + \omega_{jz} + \mu_{jz}} \qquad (FOC5)$$

Equation ET2 defines the ex-ante rule of income and labor allocation. Equation FOC3 gives the ex-post rule of income and labor allocation, when the income shocks are ϵ_{zt}^g and ϵ_{jt}^p. Equation FOC5 defines the law of motion of the rule of income and labor allocation at time t+1. Different from the full-commitment case, the utility weights are no longer fixed. Income shocks change the ex-post utility weights of the current period and the ex-ante utility weights of the future. I obtain the optimal choices vectors $\{\digamma_{tjz}, \Omega_{tjz}, \Gamma_{tjz}\}$ from the optimal contract. I define the ex-post utility values of the optimal contract, given the income shock ϵ_{jt}^p and ϵ_{zt}^g, as follows:

$$\begin{aligned} p_t\left(s_t^p, s_t^g, \epsilon_j^p, \epsilon_z^g, G^t\right) &= U_{pt}(c_{tjz}^p, l_{tjz}^p, K_{tjz}^p, N_{tjz}^p) + \beta\varrho_{t+7} V_{pt+1}^{aut}(s_{t+1jz}^p + s_{t+1jz}^g) \\ &\quad + \beta\left(1 - \varrho_{t+7}\right) P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1}) \end{aligned} \qquad (1..5)$$

and

$$g_t\left(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p, G^t\right) = U_{gt}(c_{tjz}^g, l_{tjz}^g, K_{tjz}^g, N_{tjz}^g) + \beta\left(1 - \varrho_{t+7}\right) G_{jz}^{t+1} \qquad (1..6)$$

Proposition 2 *A contract is a process of Pareto improvements. At least one household is better off ex-post in the contract. Two incentive constraints cannot be bind at the same time.*

At least one household is better off ex-ante in the contract. Two participation constraints cannot be bind at the same time. (see Appendix for proof).

Households in the contract can gain surplus by benefiting from bequest and labor allocations. Even without any transfer or support, a bequest increases the parents' expected utility. The Pareto gain from the bequest benefit means at least one household can get a higher utility value from the contract. Because of the incentive constraints, each household's utility value is no worse than the autarky case. As one incentive constraint is binding, the other will get all the surpluses from the contract and has a higher utility level.

Proposition 3 *If the market service price is lower than that of both the parent's and the grandparent's wages, such that* $p_t^k = \min\{w_t^p, w_t^g, p_t^k\}$ *or* $p_t^n = \min\{w_t^p, w_t^g, p_t^n\}$*, then households only get care service from market, with* $n_t^p = k_t^p = k_t^g = 0$*. If the parent's wage is the lowest,, then the parent is the primary elder care (child care) provider, with* $n_t^p > 0$ $(k_t^p > 0)$*. if the grandparent's wage is the lowest, then the grandparent is the primary child care provider, with* $k_t^g > 0$*. (see Appendix for proof)*

Wage structure and market service price determine the methods to be used to help the other household. Households choose the most inexpensive way to take care of elders or children. If the market service price is higher than that of both wages, all the care service will from the market. A household with wages higher than those of the service price uses pecuniary transfer rather than labor to help the other household. If a household provides child or elder care to help the other household, the intergenerational relationship directly reduces the household's working hour. In the contract, both households choose the most inexpensive way to provide child or elder care. If both wages are higher than those of outside service price, two households will use income transfer only to help each other. In one period, transfer and support exist at the same time. One household uses income transfer in exchange for the other household's labor support.

Proposition 4 *In an optimal contract, the expected substitution rate of marginal utility of the next period is equal to the substitution rate of the marginal utility of current period. For* $\forall\ \epsilon^g_{zt}$ *and* ϵ^p_{jt}*, there is:*

$$\frac{c^g_{tjz}}{c^p_{tjz}} = \sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^p_{t+1mn}} / \sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^g_{t+1mn}}$$

(See Appendix for proof)

This proposition shows the trade-off of today's and future utilities. The right term of the equations is the expected substitution rate of marginal utility of period t+1. The left term is the substitution rate of the marginal utility at time t. Two numbers are equal in the optimal contract. However, as one household has relatively low marginal utility and the other has relatively high marginal utility, the two households agree to exchange income and labor service with a relatively low marginal utility for the other's income and labor service until the substitution rates of marginal utility are equal for two periods.

Proposition 5 *If both households' income shocks are neither too large nor too small, two households will fully share the risk. Given* $\epsilon^p_{jt}(\epsilon^g_{zt})$*,* $\exists\ \epsilon^g_{kt}$ (ϵ^p_{lt}) *for* $\forall\ \epsilon^g_{zt} < \epsilon^g_{kt}$ $(\epsilon^p_{jt} > \epsilon^p_{lt})$*,* *with IC_t^p binding, and $\exists\ \epsilon_{lt}^g$ (ϵ_{kt}^p) for $\forall\ \epsilon_{zt}^g < \epsilon_{kt}^g$ $(\epsilon_{jt}^p > \epsilon_{kt}^p)$, with IC_t^g binding. Between the two extreme values, both households fully share the risk and no incentive constraints are binding. (See Appendix for proof)*

The proposition defines the marginal value separating the income shocks leading to full risk sharing and partial risk sharing. When the parent's income shock is fixed, a unique marginal value exists with the parents' incentive constraint binding and the consumption ratio equaling to λ_t. If the shock is larger than the marginal value, keeping the consumption ratio equal to λ_t will cause the parents' utility value lower than the one in autarky case. When the parent's income shock is fixed, another unique marginal value exists, with the grandparent incentive constraint binding and the ex-post consumption ratio equaling to λ_t. If the shock is smaller than the marginal value, keeping the consumption ratio equal to λ_t will cause the grandparents' utility value to be lower than that in the autarky case. In this case, the parent receives all the surplus of the contract. Both households fully share the income risk between the two extreme values. If the income shock ratio $\epsilon_{zt}^g/\epsilon_{jt}^p$ is too small, the parent's incentive constraint is binding and the grandparent receives all the surpluses. If the income shock ratio $\epsilon_{zt}^g/\epsilon_{jt}^p$ is too large, the grandparent's incentive constraint is binding and the parent get all the surplus. If the income shock ratio $\epsilon_{zt}^g/\epsilon_{jt}^p$ is neither too large nor too small, no incentive constraints bind. and two households fully share the risk.

Proposition 6 *In the no-commitment cases, a large income shock for the grandparents or a small income shock for the parents will cause a small consumption ratio $\frac{c_{tjz}^g}{c_{tjz}^p}$. Fixed ϵ_{jt}^p, if*

$\epsilon^g_{mt} > \epsilon^g_{nt}$, *then* $\frac{c^g_{tjm}}{c^p_{tjm}} \geq \frac{c^g_{tjn}}{c^p_{tjn}}$. *Fixed* ϵ^g_{zt}, *if* $\epsilon^p_{kt} > \epsilon^p_{lt}$, *then* $\frac{c^g_{tkz}}{c^p_{tkz}} \leq \frac{c^g_{tlz}}{c^p_{tlz}}$.*(See Appendix for proof)*

As one side receives a small income shock, the other household helps as much as it can, until the incentive constraint binding. A small income shock will reduce the household's consumption share, if it causes the other household's incentive constraint to be binding. A small income shock ratio $\epsilon^g_{zt}/\epsilon^p_{jt}$ causes a small consumption ratio $\frac{c^g_{tjz}}{c^p_{tjz}}$. The consumption ratio determines the transfer and support intensity households give to each other. A high consumption ratio means the grandparent receives a large share of the total endowment.

Proposition 7 *In the optimal contract, a small income shock from one household cause low discount utility values and consumption of both households. If* $\epsilon^g_{mt} > \epsilon^g_{nt}$, *then*

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \geq g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

and

$$p_t\left(s^p_t, s^g_t, \epsilon^p_{jt}, \epsilon^g_{mt}, G^t\right) \geq p_t\left(s^p_t, s^g_t, \epsilon^p_{jt}, \epsilon^g_{nt}, G^t\right).$$

If $\epsilon^p_k > \epsilon^p_l$, *then*

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{kt}, G^t\right) \geq g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{lt}, G^t\right),$$

and

$$p_t\left(s^p_t, s^g_t, \epsilon^p_{kt}, \epsilon^g_{zt}, G^t\right) \geq p_t\left(s^p_t, s^g_t, \epsilon^p_{lt}, \epsilon^g_{zt}, G^t\right).$$

(See Appendix for proof)

Households fully share income risk in the cases with no incentive constraint binding in contracts. In these cases, a small income shock from one side causes low utility values and consumption of both households. Households partially share income risk in the cases with one incentive constraint binding. In these cases, the household with a relatively large income shock helps the household with the relatively small one as much as it can until the incentive constraint binding. A small income shock form one household causes a low utility value and consumption level of the household, and leaves the others unchanged.

1.5 Estimation

In this section, I discuss the identification of several key parameters of the model. The key parameters related to households' choices are wage rate w_t^i, non-working income shock v_{it}, utility weight on leisure η, utility weight on child care α_t, utility weight on elder care weight γ_t and death rate ϱ_t. In the estimation, I assume that the working income rate and non-working income are exogenously given. I also assume in the model that only one grandparent household exists for each parent household. Due to the model's complexity, the arguments are mainly heuristic. Using the detailed time allocation and consumption information from my data, I identify the parameters of the preference. With information on work and income, I identify the parameters for income process based on households' characteristics. Then, I identify the parameters of value function in the following manner. I first discuss the identification on the law of motion in consumption ratio. Given the law of motion of consumption ratio across the period, the contract makes decisions in dividing the income and time into the current utility gain and savings for the future. I use an interpolation method to obtain the parameters for the value functions.

Income Process In the data, income comes from working income and non-working income. Households make decisions on working time. I define w_t^i as females' individual working income rate. I normalize the overall time equal to 1[15]. The individual's working income I_{it} is the annual overall income from working, and it consists of income from wages, agricultural activities and business. I define the income rate $w_{it} = I_{it}/h_{it}$. The income rate follows the rules like.

$$\ln w_{it} = \Psi X_{it} + D_{year} + D_{region} + \zeta_{it}, \tag{1..7}$$

with X_{it} as the control variable, consisting of education level edu_{it}, age age_{it}, and age squared age_{it}^2. D_{year} are the year dummies that capture the time trend. D_{region} are region dummies that control the regional fixed effects. I use ordinary linear least squares regression to obtain Ψ, from which I can derive everyone's income rate prediction at various ages. I get the working income rate, which satisfies the following:

$$\ln w_{it} = \ln w_{i0} \times \left(\pi_1 \times age_{it} + \pi_2 \times age_{it}^2\right), \tag{1..8}$$

where π_1 and π_2 are from the estimated Ψ. w_{0t} is determined by education, gender, year and region.

In the non-working income part, I treat health spending as one kind of negative non-working income. Overall household health spending H_{it} is defined as overall spending in the previous year on health care service and medicine. I_{ht} is the working income from a husband and other household members. Non-working income N_{it} is defined as the overall

[15] I define the working time as h_{it} =(annual work month/12)×(daily work hour/14)×(weekly work day/7).

household income from pension, subsidy, and another non-work income source. I define net non-working income as follows:

$$P_{it} = N_{it} - H_{it} + I_{ht}. \tag{1..9}$$

Using the information on health spending, non-working income, and individual working income, I can obtain each household's net non-working income distribution at various ages from the data. Non-working income for the individual is as follows:

$$P_{it} = \Phi X_{it} + D_{year} + D_{region} + \varepsilon_{it}, \tag{1..10}$$

I use ordinary least squares regression to obtain Φ and the distribution of ε_{it}, from which I derive each age group's non-working income rate prediction at various ages.

Preference Parameters I estimate the utility function parameters of preference, which consists of utility weight on leisure η, child care α_t, and elder care γ_t. I define the opportunity cost of time spending on child care w_{kt} or elder care w_{nt}, which is the last marginal unit of time spending on child or elder care. I normalize the child care and elder care time to 1[16]. I get the spending on child care $\pi_{it}^K = \sum_j k_t^j w_{kt}$ and elder care $\pi_{it}^N = \sum_j n_t^j w_{nt}$, j is the household member. The value of time spending is $\pi_{it}^L = w_{it}^l l_t^j$, with $w_{it}^l = w_{it}$ if i spends on working or the last unit of elder/child care is from i; $w_{it}^l = w_t^j$, if i provides child care, but the last unit is provided by j; $w_{it}^l = p_t$, if i provides child care but the last unit is provided

[16] For example, child care provided by j is k_t^j =(annual child care month/12)×(daily child care hour/14)×(weekly child care day/7).

by the market. I define household i's overall endowment[17] spending on period t is:

$$\begin{aligned} E_t^i &= R_t s_t^i + \sum_h w_{ht}^l + \epsilon_l^i - s_{t+1}^i - T_t^i \\ &\quad + \sum_h \left(n_h^i w_{nt} + k_h^i w_{kt}\right) - \sum_l \left(n_l^i w_{nt} + k_l^i w_{kt}\right) \end{aligned} \tag{1..11}$$

Here, h is the household member belonging to the household i, and l is the household member not belonging to the household i. The first part, $R_t s_t^i + \sum_h w_{ht}^l + \epsilon_l^i$ is the household endowment before transfer and support, s_{t+1}^i is the saving for the next period. T_t^i is the net transfer from household i to the other household. $\sum_h \left(n_h^i w_{nt} + k_h^i w_{kt}\right) - \sum_l \left(n_l^i w_{nt} + k_l^i w_{kt}\right)$ is the net child care and elder care support from household i to the other household. The overall household consumption π_{ir}^C is the household's overall spending on food, clothes, transportation, durable goods, utility, fuel, entertainment, education, beauty, and other consumption goods. Define the vector of spending $\pi_{it} = \left(\pi_{it}^C, \pi_{it}^K, \pi_{it}^N, \pi_{it}^L\right)'$ and the parameters vector $\Delta_{it} = (1, \alpha_{it}, \gamma_{it}, \eta_i)'$, I get:

$$\pi_{it} = \frac{\Delta_{it} E_t^i}{1 + \eta + \alpha_t + \gamma_t} \tag{1..12}$$

Define the vector of spending $y_i = \left(\pi_{it}^K, \pi_{it}^N, \pi_{it}^L\right)'$ and the parameters vector $\Lambda_{it} = (\alpha_{it}, \gamma_{it}, \eta_i)'$, I get:

$$y_i = \Lambda_{it} \pi_{ir}^C \tag{1..13}$$

To capture preference heterogeneity, I use a random coefficient model to estimate parameter distribution. Namely, the parameter vector Λ_{it} is specified as $\alpha_{it} = \alpha_t + \xi_{it}^{\alpha}$,

[17] Endowment consists both time and money values. I get the value of the time endowment by using the overall time times the opportunity cost of the time.

$\eta_i = \eta + \xi_{it}^{\eta}$, and $\gamma_{it} = \gamma_t + \xi_{it}^{\gamma}$, where $\Lambda_t = (\alpha_t, \gamma_t, \eta)'$ is a vector of constants, and $\xi_{it'} = (\xi_{it}^{\alpha}, \xi_{it}^{\eta}, \xi_{it}^{\gamma})'$ is a vector of stationary random variables with zero means and constant variance–covariance. I use two steps generalized least squares regression (GLS) to get the preference parameters (see Appendix for details). I get

$$\tilde{\Lambda}_{it}^{g} = \left(x_{ii}^{T} \tilde{w}_i^{-1} x_i \right)^{-1} x_i^{T} \tilde{w}_i^{-1} y_i. \tag{1..14}$$

Here, the weight $\tilde{w}_i$ equals to the variance $\tilde{\Sigma}_t$ estimated in the first step.I then get the sample mean $\tilde{\Lambda}_t^{g}$and sample variance Σ_t^g of $\tilde{\Lambda}_{it}^{g}$. $\tilde{\Lambda}_t^{g}$captures the average utility weight in each utility term and Σ_t^g captures the heterogeneous preference distribution.

However, heterogeneity increases the state variable dimensions and causes the curse of dimensionality for the dynamic problem. If Λ_{it} is continuous, the state variable space is infinite and impossible to solve. To reduce its dimensions and simplify the problem, I assume Λ_{it} as discrete rather than continuous in the simulation. I define two types of Λ_{it} in this simplified version of heterogeneous preference. Namely, the parameter vector $\Lambda_{it} = (\alpha_{it}, \gamma_{it}, \eta_i)'$ is specified as $\alpha_{it} \in \left\{\alpha_{it}^1, \alpha_{it}^2\right\}$,$\eta_i \in \left\{\eta_{it}^1, \eta_{it}^2\right\}$ and $\gamma_{it} \in \left\{\gamma_{it}^1, \gamma_{it}^2\right\}$. The two types are specified as $\Lambda_{it}^1 = \left(\alpha_{it}^1, \eta_{it}^1, \gamma_{it}^1\right)'$ and $\Lambda_{it}^2 = \left(\alpha_{it}^2, \eta_{it}^2, \gamma_{it}^2\right)'$. I define $\bar{\Lambda}_t^{i} = \left(\frac{\alpha_t^i}{1+\eta^i+\alpha_t^i+\gamma_t^i}, \frac{\gamma_t^i}{1+\eta^i+\alpha_t^i+\gamma_t^i}, \frac{\eta^i}{1+\eta^i+\alpha_t^i+\gamma_t^i}\right)$, for $i \in \{1, 2, g\}$. I draw N (N=10,000) pair of η_i^g, α_{it}^g and γ_{it}^g from the parameter distribution of the estimation. I define the sample means $\mu_t^i = (\mu_{\alpha t}^i,\ \mu_{\gamma t}^i,\ \mu_{\eta t}^i)$, the sample variances $\sigma_t^i = \left(\sigma_{\alpha t}^i, \sigma_{\gamma t}^i, \sigma_{\eta t}^i\right)$ and the sample covariance $q_t^i = \left(q_{\alpha\eta t}^i, q_{\alpha\gamma t}^i, q_{\alpha\gamma t}^i\right)'$, for $\bar{\Lambda}_t^{i}$. The probability that i' is type 1 is ρ and that i is type 2 is

$1-\rho$. I get the moments:

$$\mu_t^g = \rho\mu_t^1 + (1-\rho)\,\mu_t^2 \tag{1..15}$$

$$\sigma_t^g = \rho\left(\sigma_t^1\right)^2 + (1-\rho)\left(\sigma_t^2\right)^2 \tag{1..16}$$

and

$$q_{\alpha\eta t}^g = \rho^2\sigma_{\alpha t}^1\sigma_{\eta t}^1 + (1-\rho)^2\,\sigma_{\alpha t}^2\sigma_{\eta t}^2 + \rho[(1-\rho)\,\sigma_{\alpha t}^1\sigma_{\eta t}^2 + \sigma_{\alpha t}^2\sigma_{\eta t}^1] \tag{1..17}$$

Using generalized moment method, I obtain the seven parameters (see Appendix for details). I use a two-type preference model for the following reasons. (1) The two-type model captures part of the preference heterogeneity. (2) The two-type model and continuous-type model have the average utility weight on leisure, child and elder care. (3) Adding more discrete types increases dimensions causes the calculation time and memory to grow exponentially with the dimensionality.

Discrete-type and continuous-type models have the same average spending share on leisure, child care and elder care. The time spending on child care is, for example, $\pi_{it}^K = \frac{E_t^i\alpha_t}{(1+\eta+\alpha_t+\gamma_t)w_{kt}}$, which is determined by two factors: the share of spending $\frac{\alpha_{it}}{1+\eta_i+\alpha_{it}}$, and the shadow price E_t^i/w_{kt}. The setting of preference type distribution cannot affect the average share spending on child care. I test the goodness of fit of the two-type model by comparing it with the one-type model and the continuous-type model.

First, I compare the simulation results from the two-type model with the ones from the one-type model. The one-type model spends more on child care and leisure than the two-type model. In autarky case, the average labor supply in one type is 0.25. The average labor supply in the two-type model is 0.25 in the first four periods. The difference between them is approximately 1%. The average labor supply after period 5 is 0.39 in the one-type model and 0.38 in the two-type model. In the no-commitment case, the average labor supply in the one-type model is 0.34, and that in the two-type model is 0.34 in the first four periods. The difference between them is approximately 1%. The average labor supply after period 5 is 0.33 in the one-type model and 0.34 in the two-type model. Reducing types causes the underestimation on the effects of intergenerational relationships on child and elder care, leisure and labor supply.

Second, I check the goodness of fit of the utility weight of the two-type model on the continuous type model. Therefore, I estimate the average weight on child care in the continuous distribution and in the two-type model with the estimation results. The utility weight on α_{it} satisfies a truncated lognormal distribution, and η_i also meets the truncated lognormal distribution. In this part, I use a F-test (see Appendix for details) to check the goodness of the two-type preference model. I obtain the following results: F_α is 0.69, F-value is 11,126 and p value is 0. Using the same method, I obtain the value of the F statistic on leisure $F_\eta = 0.73$, and on elder care F_γ at around 0.49-0.72. All the values have p values of less than 0.05. The two-type model fits the continuous type model well.

<u>Substitution of Care Service between Market and Household</u> The real price of outside service is not the entire opportunity cost of using an outside service. In this part, I calculate the

substitution rate of elder care provided by household members and outside service. The wage rate paid for the services is not the entirety of this service cost and service quality varies among different suppliers. With the information on child and elder care, I obtain the substitution rate of elder care provided by household members and outside service. I define an outside service using dummy o_t^i - if the household uses outside child or elder care service, the number is 1; if the household uses child or elder care service from household members but not from the outside market, the number is 0. The substitution rate is ι, I parametrize the choice as:

$$o_t^i = \begin{cases} 1, \text{ if } \iota p_t \leq w_t^i \\ 0, \ otherwise \end{cases} \tag{1..18}$$

I use maximum likelihood estimation method to estimate the probit model (see Appendix for details).

Death Rate The household death rate in period t is defined as that the probability that everyone dies in period t, when the household with at least one household member alive at period t-1. I use two steps to calculate the household death rate (see Appendix for details).

Parameters for Intertemporal Decisions In this section, I estimate the form of the value functions in the dynamic problems, by using an approximation method based on simulation and interpolation. I assume each period represents four years. Grandparents retire at age 4 and parents retire at age 10.[18]

[18] According to the Survey on Fertility and Birth Control in China the average age at which females have their first child had increased from 22 in 1991 to 28 in 2010. In the model, therefore, grandparents are 24

Autarky Case The dynamic model is solved by backward recursion[19]. First, I draw the asset value and random shocks. I create an asset, wage and price space by drawing 5,000 grids of age varying vectors. Then for each asset level, I use the Gauss-Hermite quadrature method[20] to draw 10 quadrature nodes of income shocks $(\epsilon_{1t}, ...\epsilon_{10t})$ from the estimated distribution from Equation 1..10. These nodes are chosen by dividing the support of the normal distribution into 10 equiprobable intervals and then finding the conditional means within each interval. The quadrature nodes $(\epsilon_{1t}, ...\epsilon_{10t})$, lie in the domain of normal distribution, and the quadrature weights $(W_{1t}, ...W_{10t})$ are assigned appropriately to the approximate of the expected value.

As the problems for the subsequent periods are fundamentally the same, I discuss only the problem in period t. The basic logic is as follows: I use the estimated expected value function from the previous period to set up the agent's objective function to solve the problem. After obtaining the solutions, I calculate the autarky value for each asset draw. Using these coefficients, I construct the estimated expected value function which is used to solve the t-1 period problem. Then I solve the problems in the same fashion backward until t=1. The idea is stated formally as laid out below.

Suppose I have already solved for the emax function for age t+1 and the functional form of the value function $V_{t+1}^{aut}(s_{t+1}^p)$ is already solved. Given the quadrature nodes, I calculate $v_t^{aut}(s_t^p, \epsilon_{kt})$ with respect to s_t^p and ϵ_{kt}. Furthermore, in integrating for each value

years older than parents. Parents' fertility age is 28 on average.

[19] This is the standard approach in this literature. For example, see Keane & Wolpin (1997).

[20] This method follows the Gauss-Hermite rule in Chapter 7 of Judd (1998).

of the shock vector, I get the optimal consumption, labor supply, elder care and child care to derive $v_t^{aut}(s_t^p, \epsilon_{kt})$. The autarky value is given by:

$$V_t^{aut}(s_t^p) = \sum_{k=1}^{10} v_t^{aut}(s_t^p, \epsilon_{kt}) W_{kt} \tag{1..19}$$

By solving each asset draw, I get the relationship between $V_t^{aut}(s_t^p)$ and s_t^p. Then I use a linear model to approximate the expected value function:

$$\bar{V}_t^{aut}(s_t^p) = \varpi_0^t + \varpi_1^t \log s_t^p + \varpi_2^t \log p + \varpi_3^t \log w_t^i + \xi_t \tag{1..20}$$

Using linear regression, I obtain the coefficients and the fit value. Using backward induction, I solve the optimization problem for every period. After solving the dynamic problem, I obtain a sequence of coefficients for value functions. I use such a simple regression form for two reasons: (1) This simple form captures the fact that the value function is concave; (2) There is a trade-off in choosing the form of regression. On one hand, making the regression more complex could possibly improve the predictable power of the regression; on the other hand, the complex form of regression would make the first-order conditions for t-1 period problem quite complex, which is very difficult to solve.

No-commitment Case In the no-commitment case, asset levels s_t^g and s_t^p, and consumption ratio λ_t (G^t)[21], are the state variables in period t. A contract places the utility weights $\frac{1}{\lambda_t+1}$ on the parent's utility and $\frac{\lambda_t}{\lambda_t+1}$ on the grandparent's utility before the realization of income shocks. After the income shocks, the contract allocates income and labor service

[21] λ_t defines the ex-ante share of the endowment of each household gets. Unique λ_t exists for each G^t, given s_t^g and s_t^p,

according to the ex-post utility weight, which is determined by the ex-ante utility weights and incentive constraints. In each state, the contract allocates total endowment according to the ex-post weights.

I follow the backward recursion method to estimate the value function. Since the problem for each period is fundamentally the same, I only discuss the problem in period t and assume that households know the form of the value functions in period t+1. In period t, I create an asset space by drawing 5,000 pairs of age overall varying assets $s_t^g + s_t^p$, consumption ratio λ_t, and wage rate. Then, for each draw, I use the Gauss-Hermite quadrature method to draw 10 quadrature nodes for the parent $(\epsilon_{1t}^p, ...\epsilon_{10t}^p)$ and 10 quadrature nodes for the grandparent $(\epsilon_{1t}^g, ...\epsilon_{10t}^g)$. For each income shock combination of ϵ_{lt}^p and ϵ_{mt}^g, the weight is $W_{lt}W_{mt}$. The basic logic is as follows: in step 1, I match s_t^g/s_t^p with $s_t^g + s_t^p$ and λ_t; in step 2, I set up the agent's objective function and obtain the optimal choices at each state; in step 3, I calculate the value function of the agent for each asset draw, and identify the form of value functions. The second and third steps follow the same estimation method in the autarky case.

The first step reduces the number of state variables from three to two by matching s_t^g/s_t^p, with $s_t^g + s_t^p$ and λ_t. In proposition 4, the ex-post marginal utility ratio of period t is equal to the ex-ante marginal utility ratio in period t+1, such that:

$$c_{tjz}^g/c_{tjz}^p = E\left(1/c_{t+1}^p\right)/E\left(1/c_{t+1}^g\right) \tag{1..21}$$

and

$$\lambda_t = E\left(1/c^p_{t+1}\right)/E\left(1/c^g_{t+1}\right) \tag{1..22}$$

which define the law of motion of the consumption ratio. A unique s^g_t/s^p_t exists for each pair of $s^g_t + s^p_t$ and λ_t. Using the two equations and incentive constraints, I can identify the ex-post utility weight in each state, and match s^g_t/s^p_t with $s^g_t + s^p_t$ and λ_t. The state variables of the contract become $s^g_t + s^p_t$ and λ_t. I then define the ex-post consumption ratio on the grandparent κ_{jzt}, so that

$$\kappa_{jzt} = \frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}}, \tag{1..23}$$

and

$$\kappa_{jzt} = c^g_{tjz}/c^p_{tjz}, \tag{1..24}$$

with income shocks ϵ^p_{kt} and ϵ^g_{kt}. In the optimal contract with s^g_t and s^p_t, a lower bound $\underline{\kappa}_{jzt}$ exists, with the grandparent's incentive constraint binding; and a higher bound $\bar{\kappa}_{jzt}$ exists, with the parent's incentive constraint binding. Between the two bounds, $\kappa_{jzt} = \lambda_t$. $\underline{\kappa}_{jzt}$ is strictly increasing and $\bar{\kappa}_{jzt}$ is strictly decreasing on s^g_t/s^p_t. Since κ_{jzt} is a function of s^g_t/s^p_t, I get:

$$m\left(s^g_t/s^p_t\right) = \sum_{l=1}^{10}\sum_{m=1}^{10}\left(W_{lt}W_{mt}\kappa_{jzt}/c^g_{tjz}\right)/\sum_{l=1}^{10}\sum_{m=1}^{10}\left(W_{lt}W_{mt}/c^g_{tjz}\right) - \lambda_t \tag{1..25}$$

Using the moment $m\left(s^g_t/s^p_t\right) = 0$, I use GMM method to find the unique s^g_t/s^p_t for each pair of $s^g_t + s^p_t$ and λ_t.

The second step is to solve the agent's objective function for each pair of ϵ_{kt}^{p} and ϵ_{kt}^{g}. First, without taking incentive constraints into consideration, I set the ex-post consumption ratio $\tilde{\kappa}_{jzt} = \lambda_t$. I use $\tilde{\kappa}_{jzt}$ to get the utility value weight on each household and use a first order approach to solve the problem. With the solution, I get $\tilde{g}_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g)$ and $\tilde{p}_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g)$. I then take the incentive constraints into consideration. If $\tilde{g}_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g) < v_{gt}^{aut}(s_t^g, \epsilon_{kt}^g)$, I set the grandparent's incentive constraint binding, and the parent gets all the Pareto gain. If $\tilde{p}_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g) < v_{pt}^{aut}(s_t^p, \epsilon_{kt}^p)$, I set the parent's incentive constraint binding, and the grandparent gets all the Pareto gain. Otherwise, $g_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g) = \tilde{g}_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g)$, and $p_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g) = \tilde{p}_t(s_t^p, s_t^g, \epsilon_{kt}^p, \epsilon_{kt}^g)$. In this step, I get $g_t(s_t^g, s_t^p, \lambda_t, \epsilon_z^g, \epsilon_j^p)$ and $p_t(s_t^g, s_t^p, \lambda_t, \epsilon_z^g, \epsilon_j^p)$ for each pair of ϵ_{kt}^p and ϵ_{kt}^g.

The third step is to calculate the ex-ante value functions for each pair of assets and consumption ratio, and get the function form. With the $g_t(s_t^g, s_t^p, \lambda_t, \epsilon_z^g, \epsilon_j^p)$ and $p_t(s_t^g, s_t^p, \lambda_t, \epsilon_z^g, \epsilon_j^p)$ from solved from the previous step, the value functions for each asset and consumption ratio draw is the following:

$$G^t\left(s_t^g, s_t^p, \lambda_t\right) = \sum_{l=1}^{10}\sum_{m=1}^{10}\left[g_t(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p) W_{lt} W_{mt}\right], \tag{1..26}$$

and

$$P^t\left(s_t^g, s_t^p, \lambda_t\right) = \sum_{l=1}^{10}\sum_{m=1}^{10}\left[p_t(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p) W_{lt} W_{mt}\right]. \tag{1..27}$$

I get the expected value functions for each pair of $s_t^g + s_t^p$ and λ_t. By solving each asset draw, I get the relationship between value functions and state variables, given the income shock distribution. I then use a log linear function to approximate the expected value function:

$$
\begin{aligned}
G^t\left(s_t^g, s_t^p, \lambda_t\right) &= \varphi_0^t + \varphi_1^t \log(s_t^g + s_t^p) + \varphi_2^t \log \lambda_t \\
&\quad + \varphi_3^t \log p + \varphi_4^t \log w_t^p + \varphi_5^t \log w_t^g + \xi_t^g \qquad (1..28)
\end{aligned}
$$

and

$$
\begin{aligned}
P^t\left(s_t^g, s_t^p, \lambda_t\right) &= \phi_0^t + \phi_1^t \log(s_t^g + s_t^p) + \phi_2^t \log \lambda_t \\
&\quad + \phi_3^t \log p + \phi_4^t \log w_t^p + \phi_5^t \log w_t^g + \xi_t^p \qquad (1..29)
\end{aligned}
$$

Using regression, I get the coefficients and fit values. Starting at t=20, I use backward recursion to solve the problems.

Full-commitment Case The full-commitment case chooses a constant utility weight on each household in every state. The estimation of a full-commitment case model is the same as the estimation of a no-commitment case model, but the consumption ratio is fixed. For each period, I create an asset space by drawing 5,000 pairs of age overall varying asset $s_t^g + s_t^p$, consumption ratio λ_t, wages, and price, and estimate the form of value functions.

1.6 Data

Data and Sample Selection I use data from the China Health and Nutrition Survey[22] (CHNS) and the China Health and Retirement Longitudinal Study (CHARLS)[23] to perform structural estimation.

The CHNS is a longitudinal survey project collected 11 waves since 1989. The survey took place over a 3-day period using a multistage, random cluster process to draw a sample of about 4,400 households with a total of 26,000 individuals in nine provinces that vary substantially in geography, economic development, public resources, and health indicators. In addition, detailed community data were collected through surveys on food markets, health facilities, and other social services. A multistage-random cluster process was used to draw the samples surveyed in each of the provinces. Counties in the 15 provinces[24] were stratified by income (low, middle, and high), and a weighted sampling scheme was used to randomly select 4 counties in each province. A provincial capital and a lower-income city were also selected when feasible. CHNS tracks the households that in previous samples. In the summary statistics, I divided the whole sample into 2 periods. Group 1 is the 2000-2004 sample and group 2 is the 2006-2011 sample. I use group 2 in my estimation. I restrict the

[22] CHNS data is an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention. The official website of CHNS is: http://www.cpc.unc.edu/projects/china

[23] CHARLS is a project of China Center for Economic Research. The data is based on the Health and Retirement Study and related aging surveys such as the English Longitudinal Study of Aging and the Survey of Health, Aging and Retirement in Europe. The official website of CHARLS is: http://charls.ccer.edu.cn/en

[24] The provinces are Beijing, Chongqing, Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, Shaanxi, Shandong, Shanghai, Yunnan, and Zhejiang.

analysis to females who are at least 20 years old. I use the CHNS data to identify the utility weight on child care and leisure, wage rate, and the non-working income distribution.

CHARLS is a biennial survey that aims to represent of the residents of China aged 45 and older, with no upper age limit. The national baseline sample size is 10,287 households and 17,708 individuals, covering 150 counties in 28 provinces. The baseline of the CHARLS pilot took place in two provinces in fall of 2008. The first national baseline wave was fielded from 2011 to 2012. Wave 2 was fielded in 2013. The household survey includes demographic background, household information, health status and functioning, health care and insurance, work information, household and individual income, expenditure and assets. I use CHARLS data to identify the utility weight on elder care, health spending distribution, and age patterns of transfer and support.

The death rate data are from the National Population and Reproductive Health Science Data Center[25]. I use the data of 2005 to get the death rate for the data before 2008 and the data of 2010 to get the death rate of the years following 2008 (see Appendix for details)..

Descriptive Statistics Table 1 describes the working and income information. The data on female working choices indicates that, between 2000 and 2011, the average working rate decreased from 72% to 62%. Figure 1 reports the average female working time by age for

[25] The National Population and Reproductive Health Science Data Center calculated death rate by using the data from the National Census of Population in 1982, 1990, 2000, and 2010. The center calculated death rate the average death rate by age for each gender in China. Data retrieve from: http://www.poprk.org/metadata/detail/254

various periods. The result indicates that females in urban areas decrease their working time before they reach the age of 50, which is much earlier than the legal retirement age. These households in the CHNS data are getting older and give a higher weight to the older population, which reduces the average employment rate. The first graph in Figure 2 shows the income rate distribution by age cohort. The development of education system enables the younger generations to receive better education than their parents. The younger generations are more productive and less likely to work in the agricultural sector. Therefore, they have higher incomes than the older generations. In the sample, 7,950 females between the age of 20 and 70 have both working time and working income information. Table 1 shows that the average working time is 0.33 and the average female income rate is about 50,000 yuan.

Table 1 contains summary statistics on transfer, child care and elder care (see Table 7 in the Appendix for details). About 15% of people older than 60 years old age take care of their grandchildren during the survey period. Only 19% of households use outside child care service. Figure 3 reports the change in transfer and support change by the grandparent's age. Both support and transfer from parent to grandparent increase as the grandparent gets older. Grandparents take care of their grandchildren before the age of 60 and are being taken care of by parents after age 70. The net value of transfer at age 45 is 0, and it increases to about 12,000 yuan at age 90.

Table 1 also describes the household characteristics (see Table 8 in the Appendix for details). For most households, the most important asset is house. The average asset[26] level is 423,270 Yuan in CHARLS and 455,223 Yuan in CHNS. The average consumption level is 30,316 Yuan in CHARLS and 31,774 Yuan in CHNS. CHARLS have more information about asset level and consumption[27]. Therefore, the consumption level in CHARLS is higher than that in CHNS. I use the mean and variance of the net non-working income[28] distribution in CHNS data to draw the net nonworking income in simulation. Using the information on non-working income, working income and health spending, I calculate the household net non-working income of 27,495 yuan on average.

Parameter Values This section reports the estimation results. The initial female working income rate is given in Table 2. In period 1, the average parent's wage rate is 55,457 yuan and the average grandparent's wage is 59,519 yuan. Grandparent's wage is higher than that of parents on average. After the second period, As wages grow, the average wage of parents is higher than that of grandparents. To draw the wage rate in the simulation, I estimate the correlation between each pair of wage rates w_{p0} and w_{g0} by matching the households of the parent and grandparent in the CHNS data. The correlation I find between w_{p0} and w_{g0} is 0.46. I use the wage growth rate from 2006 to 2011 to obtain the predicted wage rate in the benchmark simulation. From the regression of the working income rate (see Table

[26] Figure 4 in the Appendix reports asset levels by age. In the data, the asset level reaches its highest point at age 55.

[27] CHNS does not have information about financial assets and household spending on clothes and transportation. CHARLS has all the information.

[28] Table 2 in the Appendix reports the distribution of net non-working income of various age groups.

9 in the Appendix for details), the coefficient on age is 0.056; the coefficient on age square is -0.0007. The highest working income rate among all age cohorts in this period is age 40. I use the wage growth rate from 2000 to 2004 as the second wage structure to verify to the effect of the wage structure on labor supply. The coefficient on age is 0.082 and the coefficient on age square is -0.0008. The highest working income rate among all age cohorts in this period is age 50.

Table 2 presents the main estimation results of the preference parameters. The first set of columns shows the results of the random coefficient linear regression model specifications. The first and second columns are the mean and variance of the vector of constants Λ_t, and the third column is the vector of stationary random variables ξ_{it}. Table 2 shows that the utility weights on elder care increase by age. γ_t is 0.06 before 72, increases to 0.20 from 72 to 80, and is 0.52 after 85. The utility weight on child care is 0.31 and that on leisure is 0.88. The second and third set of columns shows two types of preference results from linear regression model specifications. Type 1 households have less weight in leisure, but more weight in child care and elder care than type 2. The estimation result shows that 42% of households are type 1 and 58% are type 2. So, in both no-commitment and full-commitment cases, 34% of parents and grandparents are type 1; 18% are both type 2; 24% have type 1 grandparents and type 2 parents; and 24% have type 1 parents and type 2 grandparents. Table 2 also reports that the substitution rate between outside child care service and household child service is 2.5, and the substitution rate between outside elder care service and household elder service is 1.9. I assume that the discount factor β is

0.97, which corresponds to a rate of time preference of 3% per year. I assume that the real interest rate R is 1.01, which corresponds to the average interest rate in China from 2006 to 2013 (World Bank 2016).

1.7 Life-Cycle Fit

In this part, I use simulation to predict households' decisions on savings and labor supply along the life-cycle. The parameters have been presented in the previous section. The simulation focuses on household choices within the intergenerational relationship. So, in the section and policy analysis section, I only consider the household behavior within the contract. The estimated results indicate the average number of the entire sample.

Simulation Procedure To create the simulation sample, I draw a random sample of 5,000 pairs of savings according to the parametrization described above. Using the simulation results, I offer some economic intuition related to the life-cycle profiles of household labor supply, transfer, support and savings from the model. From the CHNS data, I draw the initial wage rate w_0^p of parents using the wage rate distribution of females ages 20 to 23 and draw the initial wage rate w_0^g of grandparents using the wage rate distribution of females ages from 41 to 45. The correlation between w_0^g and w_0^p is 0.46. Using the parameters from the wage growth equation, I obtain individual i's predicted wage rate w_t^i at time t. The average real child care and elder care market service cost is 62,500 yuan[29]. The non-working income

[29] By calculation, I get the average child care price is 24,889 Yuan, and the average elder care price is 31,903 Yuan. The substitute rate between outside child care service and household child service is 2.5, and the substitution rate between outside elder care service and household elder service is 1.9. I assume both

is drawn according to the distribution of the net non-working income[30]. From the CHNS data, I draw parents' initial asset level s_0^p using the asset distribution of females ages 20 to 23 and draw grandparents' initial asset level s_0^g using the asset distribution of females ages from 41 to 45. In period 0, the average initial asset level of grandparents is 375,000 yuan, while the average initial asset level of parents is 75,000 yuan. The correlation between s_0^g and s_0^p is 0.53. For everyone in the sample, the simulation uses: three fixed individual characteristics (working income rate, non-working income distribution, and care service price), three initial state variables (parents' savings, grandparents' savings and consumption ratio).

The initial consumption ratio is defined by assuming both households divide the ex-ante Pareto gains equally before they join the contract:

$$\lambda_1 = \arg\max \left[G^1 - V_1^{aut}(s_1^p)\right]^{\frac{1}{2}} \left[P^1\left(s_1^g, s_1^p, G^1\right) - V_1^{aut}(s_1^g)\right]^{\frac{1}{2}} \tag{1..30}$$

In each period, I draw the income shock from the net non-working income distribution of each pair of households. I then solve the optimal contract, given the consumption ratio and the savings. Next, I find out the consumption share in period t according Equations 1..23 and 1..24. Using Equation 1..11, I get:

$$s_{t+1}^g + s_{t+1}^p = \frac{\beta A_1\left(E_t^p + E_t^g\right)}{\kappa_{jzt} A_2 + A_3} \tag{1..31}$$

Here $\beta A_1 = \beta\left[\left(1-\rho_{t+7}\right)\left(\kappa_{jzt}\varphi_1^t + \phi_1^t\right) + \rho_{t+7}^t \varpi_1^t\right]$ is the weight on overall saving; $\kappa_{jzt} A_2 = \kappa_{jzt}\left(1 + \alpha_{t+6} + \eta_{t+6} + \gamma_{t+6}\right)$ is the weight on the grandparents' utility level; $A_3 = 1 + \alpha_t +$

prices are 62,500 yuan.

[30] The distribution of net non-working income by age is shown in Table 2 in the Appendix.

$\eta_t + \gamma_t$ is the weight on parent's utility level. By Equations 1..12 and 1..13, I can get the spending on each term. By matching s^g_{t+1}/s^p_{t+1} with $s^g_{t+1} + s^p_{t+1}$ and κ_{jzt},according to Equations 1..21 and 1..22, I get s^g_{t+1} and s^p_{t+1}, and go to next period. Using the same method, I get the predicted values from period 1 to 20.

No-commitment Case Fits Data Better The no-commitment case model fits the data better than the full-commitment case and the autarky case. To illustrate this point, I compare the actual average working time, savings, child care, and elder care choices from the data predicted by the no-commitment model. Table 3 presents the actual and predicted values of the no-commitment case on labor and savings as well as other measures. The dynamic model reasonably predicts the working, leisure, transfer and saving choices. The chi-square goodness of fit tests does not reject the null hypothesis that these values are different.

The no-commitment case fits the data on labor supply. The autarky case fails to explain the age pattern of female supply. The decline in the female labor supply before retirement age does not occur in the autarky case. Figure 1 shows that the female working time of the rural[31] population aged 40-45 decreases from about from 0.4 to 0.1, which is about 5 years earlier than the legal retirement age. Figure 5 displays the simulation results on labor supply[32]. In the autarky case, the labor supply of grandparents decreases smoothly until the legal retirement age. In both the no-commitment and full-commitment

[31] I define rural households and urban households by the households' Hukou. A Hukou is resident recording systems required by law in China. Hukou officially identifies a person as a resident of an area, and rural or urban resident.

[32] Table 3 in the Appendix shows the simulated value of the labor supply by age.

cases, grandparents work less to provide more care for their grandchildren from period 1 to period 4[33], and parents work less to provide more elder care from period 5 to period 10[34]. The combination of the two effects causes the decline in labor supply before retirement age in the data. Labor support through intergenerational relationships changes the labor supply of both parents and grandparents. In addition, in the autarky case, parents provide child care and work less. The scenario is not shown in the data and the no-commitment case, because grandparents take care of grandchildren in both settings. Without intergenerational relationships, the autarky case cannot explain the age patterns of female labor supply. The no-commitment case under predicts the effect of intergenerational relationships on the labor supply before retirement age. The second graph in Figure 5 compares the age patterns of labor supply from the simulation and data. In reality, parents also need to take care of their own grandchildren. My model ignores the problems on grandchild caring on the part of parents. Without the grandchild care problem, the labor supply of parents from simulation is larger than that in the data at around age 50.

The no-commitment case fits the data on saving. The full-commitment case cannot make predictions on the savings of a single household. In the full-commitment case, which is free of incentive problems, only the overall savings and consumption ratio matter

[33] Figure 1 in the Appendix compares the female labor supply of the households with and without children. In this part, I define the female without children, as a household that does not have children and whose households' members were at least 35 years old before 2004. 1,465 females meet this standard and 9,415 females have at least one child.

[34] Figure 2 of Appendix compares the female labor supply of the households with and without parents alive. In this part, I define the female without a parent as a household that does not have parents or parents-in-law. 1,565 females meet this standard, while 6,923 females have at least one living parent or parent-in-law.

in the allocation of income and labor. Without a bequest tax, the saving allocation of the two households has no effects on working, child care, and elder care decisions in the full-commitment case. In the no-commitment case, a single household's saving determines the outside option of the contract and affect the endowment allocation between households; a unique savings-share rate exists in each state. Without a single household's saving information, I can only identify the labor supply, child care, and elder care decisions in full-commitment contract. The no-commitment case model predicts saving[35] better at the later stages of the relationship. The second graph of Figure 6 compares the age patterns of saving from simulation and data. The simulation results of the no-commitment case fit the data well in the later stages of the relationship. In the data, some parents cohabitate with the grandparents and thus their assets cannot be separated. Therefore, parents are given a high average asset level during the earlier periods of the relationship.

Choices of Transfer and Support across Periods Households' savings and wage rates determines the choices and intensity of support and transfer. Figure 7 illustrates the simulated net transfer and support value[36] by age in the no-commitment case. At the early stages, parents are less motivated to use transfer exchange bequests, because the grandparents' death rate is low. Without a bequest incentive in the first three periods, the net transfer and support values are only above zero. But as the grandparents grow older and are more

[35] Figure 4 in the Appendix shows the simulated value of saving by age.

[36] The net transfer and care support value are defined by net transfer plus labor support value. Labor support value is defined by the support hour time support opportunity cost.

likely to die, bequests are more likely to occur. During the first three periods, parents pay transfers to grandparents' child care. After period 4, bequest incentives are mainly changed by two forces: the grandparents' death rate, which increases by age, and asset level, which decreases by age after period 4. From period 4 to period 10, the first force dominates the second force. The bequest incentive increases and the net transfer value grows from period 4 to period 9. After period 10, the second force dominates the first force, and the net transfers and support value decreasing by age. The age patterns of endowment composition fit the actual data well, as showed in Figure 3. The endowment composition determines the intensity of support and transfer.

Figure 7 also indicates the composition change of transfer and support by age. The bequest incentives and coinsurance affect the overall net transfer and support value in the contracts. The composition of transfer and support changes across the periods. The wage structure determines whether the household chooses support or transfer to help one another. At the early stage of the relationship, because grandparents' wage rates are lower than those of parents and the outside care service rates, grandparents become the primary child care providers. Parents use transfer to pay for the support. After period 5, when grandparents need elder care and parents' wage is lower than the service price, parents use an outside elder care service and transfer in exchange for future bequests. After period 10, when parents are no longer able to provide elder care and work, they only use transfers in exchange for bequests. The age patterns of the intensity of transfer and support fit the actual data (see Figure 3).

Effect of Wage Structure on Labor Supply Wage structure affects labor supply by determining the choices of child and elder (see Table 10 in the Appendix for details). Figure 5 describes the simulated results of labor supply of the no-commitment, autarky, and full-commitment cases. Given the wage structure and service price, in the autarky case, parents provide child care by themselves and grandparents acquire elder care service from the market. In the no-commitment and full-commitment cases, parents use grandparents' child care service and grandparents use parents' elder care service. Figure 5 presents the difference in the labor supply between the full-commitment and no-commitment cases. In the full-commitment case, households provide more assistance to each other. Grandparents provide more child care and parents give more elder care in the full commitment case than the no-commitment case. In the full-commitment contract, the labor supply of parents is about 5% higher before period 4 and 5% lower after period 5 than that in the no-commitment case. Grandparents substitute the parents for child care in the contract, parents' labor supply increases by 26%, and grandparent' labor supply decreases by 19% in the first 4 periods, compared to the autarky case. However, when grandparents grow older, parents' labor supply is 13% less than that in the autarky case.

The wage structures influence labor supply by affecting the households' elder care and child care choices. To help understand how the wage structure changes the labor supply behavior, I add a new wage structure, which is presented in the second graph in Figure 2. In the new wage structure, parents become the primary child care providers (details see Table 10 in the Appendix). In the contracts, parents' labor supply does not vary significantly compared with that in autarky case in the first 3 periods. The wage structure determines who the primary care providers are, which affects whether or not the substitution affects on labor supply exist in intergenerational relationships. In a fast-growing economy, such as China and India, younger generations have higher wages rate than the older ones. In these countries, intergenerational relationships have greater effects on the labor supply than other countries.

Richer Gets More Transfer and Support Richer households have more resources for exchange in intergenerational relationships than poor households. Grandparents use bequests in exchange for parents' income and labor service in contracts. Bequest motives encourage grandparents to save more in the contracts than in the autarky cases. Figure 6 describes the simulated results of savings. Grandparents save more in the no-commitment case than in the autarky case. Besides, grandparents obtain a higher utility level of securing less expensive elder care, greater income and labor service from the intergenerational relationship, and therefore greater saving more money to save. Grandparents save, on average, 19% more in the no-commitment case than the autarky case (see Table 10 in the Appendix for details). Bequests discourage parents to save in the no-commitment case. Parents save 6% less on average in the no-commitment case than in the autarky case. In the no-commitment case, when grandparents die, bequests increase parents' savings in future. Bequests reduce the marginal utility of savings, and reduce parent's motivation to save. Bequests increase grandparents' but decrease parents' savings.

To confirm the effect of the incentive problem on household decisions, I conduct two experiments. In the first experiment, I change the average initial asset level of grandparents from 375,000 to 125,000 and leave parents' initial saving unchanged. Grandparents have small bequest in exchange for parents' income and elder care service. Figure 8 presents the change in transfer and labor supply brought about by the adjustment. The first figure shows the new predicted values of transfer and support. The average net transfer and support value decreases from around 15,000 to around 5,000. The transfer value is negative in some periods. Grandparents use both bequest and transfer in exchange for parents' elder care support, as bequests are not enough to meet the demand for elder care. Parents works more, when the grandparents have only a small savings. The second graph in Figure 8 indicates the effect of the asset change on labor supply. In the first 4 periods, grandparents provide more child care than the benchmark setting. Parents receive more child care from grandparents, thus increasing their working time. After period 4, the bequest incentive is weakened by low level asset level. Parent's working time increases by about 10%. In the second experiment, I leave the initial assets and other parameters unchanged. I select a pair of grandparent and parent by drawing the initial assets and wage rates level from the distributions. I set a fixed income and asset path before period 5. I assume 2 cases in period 5. In the no-commitment case, grandparents with good income shock obtains 13% greater elder care service from parents than those with a bad income shock. The good income shock increases the labor supply of parents by 7%. In the full-commitment case, the grandparent with good income shock gets 3% less elder care service from parent than those with bad income shock. The good income shock decreases the labor supply of parents by 2%.

The motive source affects households' transfer and support behavior in intergenerational relationships. In the no-commitment cases, coinsurance and bequest affect the amount of support and transfer that households provide to one another. These factors establish the intensity of mutual aid and the size of the effects on labor supply. Households transfer money and support one another in exchange for current or future income, and labor service of others. In poor economic conditions, households with smaller endowments exchange transfer and support from other households. Thus households with fewer assets obtain less support and transfer in the no-commitment contracts. The result is opposite to the prediction of the standard altruistic models, in which poorer households obtain more help in intergenerational relationships[37]. In the altruistic models, households help others to increase their own utility level by increasing other household's utility. Households with fewer assets obtain more help from the others. The differences in the motives for transfer and support cause the differing predictions of the two models.

<u>Heterogeneous Responses</u> Intergenerational relationships have a greater influence on households with more demand on care service than on those with less. Figure 9 shows the changes in labor supply change caused by intergenerational relationships between type 1×1 households (in which both households are type 1) and 2×2 households (in which both households are type 2). Intergenerational relationships have a more significant effect on the labor sup-

[37] Parents account both for the individual and relative economic position of their children and give them transfers or sharing of inheritance to their children unequally (Schanzenbach & Sitkoff 2008).

ply of type 1×1 households, because type 1 households need more child and elder care than type 2 households (see Table 10 in the Appendix for details). In periods 1 to 4, intergenerational relationships increase the labor supply of type 1×1 parents by 40%, increase that of type 2×2 parents by 2%, and decrease that of type 1×1 grandparents by 7%, and increases that of type 2×2 grandparents by 19% .

The demand for child care and elder care influences the effect of intergenerational relationships on labor supply. Intergenerational relationships have great impact on the labor supply of households with a high utility weight on child care or elder care.

1.8 Policy Analysis

To illustrate how household labor supply, savings, child care, elder care and transfer rates change as the public policies, I calculate the percentage changes in these measures as I impose the policies. The settings of death rate, income process, and wage structures are the same as the benchmark settings in the previous section.

Child Care and Elder Care Subsidies This section shows the effect of child care and elder care service subsidies on labor supply decisions. Keeping other parameters and settings constant, I compare the choices of households facing zero subsidy, 10% subsidy, and 20% subsidy on elder or baby care service purchased from the market and a 20% subsidy on both elder and child care service (see Table 11 in the Appendix for details).

The 20% care service subsidy has a much greater impact on labor supply than the

10% care service subsidy. Figure 10 illustrates the simulated effects of the subsidy on both kinds of care services. With a 10% care subsidy, the service price remains higher than most of females' wages. However, with a 20% subsidy, most parents' and grandparents' wages are higher than the price of outside service. The 10% care subsidy increases parents' and grandparents' labor supply by only 6%. The 20% care subsidy increases the grandparents' labor supply by 41%. Child care subsidy affects grandparents' labor supply in the earlier stages. Elder care subsidy affects labor supply at the later stages of the contract. The 20% child care service subsidy increases the labor supply of younger females by 24% in the autarky case and by 10% in the no-commitment case. The subsidy increases the labor supply of older females by 39% from period 1 to period 3. The 20% elder care service subsidy increases the labor supply of younger females by 13% in the no-commitment case after period 4.

Child care and elder care subsidies affect the labor supply of both households. When grandmothers are the primary child care providers, the low-cost market service can substitute for grandmothers in the provision of child care. The substitution increases the labor supply of old females. An elder care subsidy can affect the labor supply of parents. When the mothers are the primary elder care providers, the low-cost market service substitute for the mothers in the provision of elder care. Moreover, the labor supply of young females also increase. Child care and elder care service subsidies reduce the demand for support through intergenerational relationships by encouraging households to utilize the formal care service. The subsidies increase the females' labor supply, when the subsidized service price is lower than the care providers' wages.

Delay Mandatory Retirement Age This section presents the effects of delaying the mandatory retirement age on labor supply and saving decisions of households. Keeping other settings unchanged, I delay the retirement age from age 9 to 10. The mandatory retirement ages of females are between 55 and 60 years old in China (see Figure 7 in the Appendix for the details). In 2015, China's Ministry of Human Resources and Social Security announced a new retirement plan to delay the mandatory retirement age to 65 years old in the 2020s [38]. Therefor, female labor supply is predicted to change because of the new retirement plan (see Table 11 in the Appendix for details).

With the benchmark wage structure, retirement delay causes a insignificant effect on labor supply. In this part, the working income rate follows the growth rate of age as described in Table 4. Figure 11 shows the simulated effects on retirement delay. In the benchmark wage structure, grandparents' wage is lower than parents' wage and the service price. Before and after the policy is changed, grandmothers are always the primary child care providers. Delaying the retirement age has limited effects on the baby care choice, as well as the labor supply of grandparents and parents. The policy reduces the parents' labor only by 8% with the benchmark wage structure.

As I narrow the wage gap between parents and grandparents, the retirement delay

[38] The Chinese government plans to take pressure off the nation's increasingly strained pension system by gradually raising retirement ages for the nation's millions of workers between 2017 and 2022. The nation's Ministry of Human Resources and Social Security has declared that eligibility ages for men, women, urban workers and farmers will be raised in steps by adding "several months every year" to the age which pension payments can begin. (Shi, Xu, Zhang & Yao 2015)

causes a large effect on labor supply. The wage structure[39] has a small wage gap between young and old females. The retirement delay has different effects on the new wage setting. Delaying the mandatory retirement age remove grandmother from child care. The retirement delay decreases parents' labor supply and increases grandparents' labor supply in period 4. Delaying the retirement age reduces parents' labor supply by 43% in period 4.

The wage structure determines the effect of delaying the mandatory retirement age. In an economy with rapid human capital growth, younger generations have higher human capital level and wage rate than the elders. Elder females reduce their labor supply before the legal retirement age, to provide child care to their grandchildren. Delaying retirement age only has limited effects on females' labor supply. In an economy with slow human capital growth, the wide wage gap between old and young recedes. Delaying the retirement age removes the retired grandparents from child care, and parents decrease their labor supply to provide child care. Delaying the retirement age has great effects on reducing young females' labor supply.

Inheritance Tax This section states the effect of inheritance tax on household labor supply and saving decisions. In the benchmark setting, households face zero inheritance tax[40]. Keep the other parameters and setting unchanged, I estimate the households' responses with the 30% inheritance tax.

[39] The structure experiment uses the wage structure of 2000-2004, in which 50 years old has the highest income rate among all ages.

[40] China proposed inheritance tax law in 2004 but has not yet been able to introduce it due to widespread opposition.

Inheritance tax increases grandparents' savings in no-commitment cases. Figure 12 shows the simulated effects of inheritance tax on saving and labor supply. As the death rate before the period 4 is small, bequest is not likely to occur during these periods. The inheritance law only has few effects on households' saving labor supply behavior. However, after that period, with fewer bequest from grandparents, parents provide less elder care and transfer than the benchmark case. The influence of intergenerational relationships on labor supply is less than that in the benchmark case. The policy increases parents' labor supply by 9%. Taxation weakens the crowd out effects of bequest on saving, therefore parents save about 17% more than the benchmark case. Inheritance tax also reduces grandparents' expected consumption level by reducing the exchanged support and care service from parents. Grandparents save about 19% more than the benchmark.

By weakening the bequest incentives, inheritance taxes affects households' saving behavior and indirectly affects labor supply decisions. Bequests give incentives for parents to transfer money and provide elder care to grandparents. Inheritance tax reduces the net value of the bequest in exchange for parents' support and transfer. It causes less transfer and elder care from parents to grandparents. The tax indirectly increases the labor supply of parents at the later stage of the relationship. Grandparents increase their saving rate to maintain bequest incentives. It may offset part of the direct labor supply increasing effects caused by inheritance tax. In contrast to the no-commitment case, inheritance tax increases the intensity of transfer and support in the full-commitment case. In the full-commitment case, grandparents prefer to transfer all the money to parents before death to avoid inheritance tax. Without incentive problems, inheritance tax increases the transfer from grandparents to parents at the early stages and the transfer from parents to grandparents at the late stages. Grandparents save more to sustain the bequest incentive for parents to provide transfer and support after introducing the inheritance tax.

1.9 Conclusion

This article develops a non-altruistic dynamic contract to analyze how intergenerational relationships affect households' labor supply and saving behavior in the presence of idiosyncratic income shocks and death uncertainty. A distinguishing feature of the model is the use of economic benefits only to sustain intergenerational relationships. From closed form equilibrium allocations, it is straightforward to derive rule of income and labor service allocations within the relationship across time. This article adopts a first step toward showing how do the economic factors of intergenerational relationships affect households' behavior throughout a life-cycle. The framework can be extended to incorporate fertility decisions and life cycle human capital development. The theoretical framework sheds light on a range of questions in which marriage, child care, elder care and fertility are central to the analysis.

My empirical results suggest that intergenerational relationships increase young females' labor supply by 32% and decrease elder females' labor supply by 21% in China. The choices of support and transfer depend on wage structure and market care service price on extensive margins. Households' savings and wage rates determine the intensity of support and transfer of intensive margins. The rapid human capital level growth contributes to the strong intergenerational relations and high labor market participation rate in China. The article ignores human capital development and migration decisions for simplification purposes; this condition may underestimate the effects of intergenerational relationships on labor supply. Strong intergenerational relationships contribute the over-investment on children's education and large scale temporary migration in China. Studying the effects of intergenerational relationships on human capital investment and migration decisions is promising direction for future research.

A further step I take is to quantify the spillover effects of public policies through intergenerational relationships. I discover that child care subsidies increase the labor supply of grandparents and that elder care subsidies also increase the labor supply of parents. Inheritance taxes increase households' savings rate by reducing the bequest incentives. Delaying mandatory retirement age only has limited effects on female labor supply when grandmothers are the primary child care providers, but has big effects when mothers are the primary child care providers. These findings illustrate the importance of modeling intergenerational relationships and household decisions simultaneously. Ignoring the connections between households of different generations can lead to incomplete forecasts of the effects of some policy changes.

CHAPTER 2.

AFFIRMATIVE ACTION AND HOUSEHOLD EDUCATION INVESTMENT

To evaluate the distribution effects of AA on different groups, we have to understand household's strategic reactions on the rule of college seats allocation In Chapter 2, using a Chinese household survey and a dataset of college enrollment information, I exam how do households with heterogeneous talent and endowment to compete for different level of college seats, and check the impact of on human capital inequality between and within region. I investigate households' schooling investment responses to AA and college expansion. Using a Chinese household survey data and a college enrollment data, I examine the manner in which households with heterogeneous talents and endowments compete for separate levels of college seats. Using the number of new colleges and the number of college enrollment in the neighboring provinces as the instrument variables for the provincial enrollment ratio, I find that on average, higher enrollment ratios encourage households to invest in education. The effect of encouragement is greater for advantaged households with better education level and higher income level. In the long-term, college expansion has a significant effect on improving advantaged households' education level and income level. I find that AA and the college expansion widen the education and income disparity within each region.

2.1 Introduction

his article studies households' education investment responses to AA. AA in college admissions provides different educational opportunities to diverse groups. The existing literature has not considered the implications of AA for household's reactions to changes in educational opportunities. Universities allocate seats according to students' examination scores and other human capital outputs. Households invest money and time on education, accordance with their income, abilities, and competitors' conditions. The difference in educational opportunities as results of AA shapes households' education investment decisions. Within an AA targeted group, students and households have considerable differences in talent and endowment. Households in the same group may have different reactions to changes in educational opportunities. The reactions of households depend on the strategic moves of competitors, as well as the endowment and talent distribution within each group. Thus, this study addresses the following questions:

1. How do households' education investment decisions respond to the changes in educational opportunities or AA?

2. How do households respond to changes in educational opportunities change with respect to their endowment and talent?

3. How do the changes in educational opportunities or AA affect the education and income disparity within each group?

This study uses the college expansion in China after 1998 as a natural experiment to study the manner in which households' education investment decisions respond to changes in educational opportunities. The admission system of the National College Entrance Examination (NCEE) is an AA system that allocates college seats by region. The college expansion provides an exogenous change in educational opportunity. This study uses the variations in the expansions across regions and time to identify the effects of changes in the enrollment ratio on education investment. I have concluded that generally, increasing the college enrollment ratio increases education spending. Households with better education level, and higher household income level or urban Hukou[1] are more sensitive to changes in enrollment rate. In the long term, college expansion and AA widen the education and the income gap within each province.

Using Chinese data to investigate these topics has several advantages. (1) The signal of educational opportunities is clear. The college student's enrollment system is a centralized enrollment system. The Education Department announces the enrollment plan before the NCEE and the enrollment information after the NCEE of the year. Households can easily obtain the information on colleges quality[2] and the number of enrollments in each college. Compared with an independent enrollment system, the enrollment ratio is significantly

[1]Hukou is resident recording systems in China. Hukou officially identifies a person as a resident of an area. The two types of Hukou, namely, urban and rural, pertain to urban and rural population, respectively. Urban Hukou entitles its holders to better schools and welfare programs. (Hukou system, wikipedia)

[2]Four college ranks exist: top four years universities, other four years universities, regular three years universities, other regular three years universities and technical colleges. Higher rank colleges enroll students first. After the higher level finished their enrollment, the next level colleges begin their enrollment.

clearer for households. (2) The enrollment rule is simple. The sole factor that determines the enrollment decision is the students' rank in the examinations. A college sets a score threshold according to the score distribution of the students applying for college. The college enrolls students who garner scores that are higher than its score threshold. (3) The demographic and economic differences across regions are observable. NCEE allocates college quota by region, and not by race or other demographic characteristics. The Education Department provides each province with enrollment quotas. Students in different provinces have different enrollment ratios. Compared with the AA based on race, economic or social status, the endowment and talent differences in the AA based on regions are easier to capture. (4) College expansion was an exogenous shock to household behavior. The central government announced the college expansion plan in 1997. Thereafter, the enrollment ratio increased from about 30% to 75%. The rates of college expansion are different across regions, which lead to the differences in the changes of enrollment ratio. I use variations of the enrollment ratio of the different regions to study the effect of the education chance change on households' education investment.

2.2 Literature

This study follows the theoretical framework of Bodoh-Creed and Hickman (2015) and Cotton, Christopher, Hickman and Price (2015). Both studies use all-pay auctions to characterize the competition for university admissions. In an auction, a student's education investment is the bid. The payoff is the seat of colleges with different qualities. Bodoh-

Creed and Hickman (2015) develop a theoretical model that can track the college seats assignments with AA by matching a continuum set of university seats with a continuum of students. I present a simpler version of their model focuses on aspects of the market that is most readily testable using linear regression. In the current study, I use the continuum setting and the pure rank of the human capital to set the prize for the competition.

The current study contributes to the literature on the effects of educational opportunity and AA on household incentives to make human capital investment. Coate and Loury (1993) regard AA as a mandated equal assignment rate- that has different effects on minority workers' incentives in a job assignment market. Moro and Norman (2003) find that AA may increase the minority workers' incentive to invest in learning but diminish the others. Furstenburg (2003) builds an AA model of college admissions, and finds that AA admissions rule to enhance the academic quality of its class. Fu (2006) finds that AA encourages the non-minority to respond to the AA admissions more aggressively, which tends to widen the test score gap in terms of race. Ferman and Assuncao (2011) find that test scores among black students decreased because of an AA admissions quota in top universities in Brazil. Antonovics and Backes (2014) finds a modest effect in the abolition of AA at the University of the California system on the GPA of college-bound students. Hickman (2015) builds a structural empirical model and finds that AA increases the human capital level of the minority compared with color-blind admissions, representative quotas, and American-style preference-based AA. Cotton, Hickman, and Price(2015) conduct a field experiment to mimic essential aspects of competitive investment prior to the college market. The study finds that AA increases average human capital investment and examination performance for the majority of disadvantaged students targeted by the policy. Estevan, Gall, and Morin (2016) find that AA increased the targeted group's chance of entering colleges-but find limited evidence of behavioral reactions to the AA policy, in terms of test performance or application decisions. The study uses Chinese household survey data and NCEE enrollment information to estimate the effect of college enrollment ratio change on households' education investment. I have concluded that generally, increasing the college enrollment ratio decreases education spending.

As households or individuals have heterogeneous responses to changes in educational opportunity, AA changes the distribution of human capital level within the targeted group. For example, Bertrand, Hanna and Mullainathan (2010) uses the AA program for "lower-caste" groups in engineering colleges in India and finds that the lower-caste applicants obtain more college seats but caste-based targeting reduces the number of females entering college. Yeung (2013) finds that the college expansion in China will benefit households with higher levels of education to enter college and liminate or even reverse the gender gap in college attendance. The study finds households with better education level, and higher household income or urban Hukou are more sensitive to changes in enrollment rate. In the long term, college expansion and AA widen the education and the income gap within each province.

2.3 Background

NCEE is an annual academic examination. It is a prerequisite for entrance to all of the colleges and universities in China. A student's overall mark is a sum of the subject marks. The maximum possible mark[3] varies widely from different years and from different provinces. The students need to to choose their majors in NCEE. After the exam, colleges and universities enroll students separately in two majors. Students can take either a set of art and humanities subjects, or a set of science and engineering subjects, based on their choice of majors. Chinese, mathematics and a foreign language are mandatory subjects. The applicants of the science and engineering majors need to take a science-integrated test, and the applicants of the art and humanities majors need to take a humanities integrated test.

For each student, preparations for the NCEE start at least two years before the examination. Figure 1 shows the time line of NCEE each year. Local educational institution announces the enrollment plan in September before the year of the examination. Each student registers the exam and chooses a major in November. Students take the exam in June- and obtain their score by the end of the same month. They make their decision on which college they will attend in July. In other places, students list their university

[3] Figure 4 is examples of the score distribution. The first one is the score distribution of Yunnan in 2014. I labeled the score threshold to apply different level of universities. There are several hundreds of students get 0 each year, who give up the exam even they register it. The distribution is a relatively normal distribution if I drop the 0 score. Because the exam is not standard and the difficulty level is changing every year, the score distribution is changing every year. The number of students who took the exam is also changing every year because the population of the age cohort is changing.

or college preferences prior to the examination, after the examination, or after they find out their scores[4]. Applications are given to several tiers (including early admissions, key universities, regular universities, and technical colleges), each of which can include around 4 to 6 choices of institutions and programs. In some places, students list applications of separate tiers at different times (Zou, 2013). Usually, a student's choice of is made about two years before taking the NCEE. After making the choice, the students focus on studying the relevant set of subjects.

Admission quotas are distributed to each province. A university sets a fixed admission quota for each province, which must be approved by the Education Department of the central government[5]. After the approval, the Education Department will send the plan to the local education department of each province. The NCEE is administered uniformly within each province of China. In NCEE, test takers only compete with the students in the same province (Gu, 2012). The college seats are distributed unevenly across China and students are being discriminated during the admission process based on their geographic regions. Unequal admission schemes for different provinces and regions might intensify competition among examinees from provinces with fewer advanced educational resources(Gu, 2012). For example, in 2010, acceptance rate for students from Beijing, Shanghai, Shan-

[4]For example, in Shanghai, students list their applications for early admission, prior to the exam, but students in other provinces do so after they find out their scores.

[5]At the central government level, the Minister of Education and the National Development and Reform Commission (NDRC) jointly decides the total enrollment number of students every year, which then is subdivided into 31 provinces, municipalities, and autonomous regions, and these enrollment plans are implemented at the local government level (Gu, 2012).

dong and Henan who applied for universities of the first-ranking category were 20%, 18%, 7% and 4% respectively. High acceptance rates are likely to appear in the most and least developed cities and provinces.

The annual number of takers and the number of colleges to enroll in through the NCEE are shown in Figure 2. The 3 big turning points are shown: the first is in 1966, the cultural revolution, when the NCEE was stopped; the second is in 1976, the restoration of the NCEE; and the third is in 1998, when China began to expand the college rapidly after this year. After the restoration of NCEE, from 1978 to 1998, the scale of higher education kept increasing. The number of colleges increased from 598 to 1022, the number of new college students enrollment increased from 0.4 million to 1.1 million, and the number of college students increased from 0.9 million to 3.4 million (Li & Xing, 2010). In 1999, to support economic growth, the Chinese government announced the rapid college expansion policy[6] targeted at expanding tertiary education dramatically to reach an enrollment ratio of 15%[7]. The annual college enrollment increased from 1 million in 1998 to 6.3 million in 2009. In 1998, the ratio of people who enters the college among the age cohort was 9.8%; in 2007, after the rapid expansion, the number was 23% (Yeung, 2013).

[6] The official explanation of college expansions are: (1) the need for more talented personnel to sustain the rapid development of Chinese economy; (2) the public demand for higher education is increasing and the government has the obligation to meet their demand; (3) enrollment expansion can postpone employment of high school graduates and increase educational consumption, which is an important means to stimulate domestic consumption and promote growth in related industries; (4) enrollment expansion will reduce the pressure on high schools, discouraging test-oriented teaching and learning and promoting all-around education in elementary and secondary schools (Wan, 2006).

[7] In early 1999, the central government decided to increase the number of students admitted to tertiary education by 0.22 million. In June, the central government and the Ministry of Education suddenly made an announcement that a further 0.33 million new students will be admitted (Li & Xing, 2010).

In the mid-1990s, the central government decentralized the administration of colleges and universities. After the decentralization reform, the admission quota can be determined at the provincial level. The allocation of admission quota often is biased toward local students, even for the universities managed by the central government. Provinces with more high education resources have a greater capacity to expand after 1999, and individuals from these provinces are more likely to benefit from the expansion policy (Yeung 2013).

2.4 Theoretical Model

In this section, I present an all pay auction model of education investment decision to characterize households' education investment competition within each province. In an environment wherein households with different level of talent have heterogeneous responses to the change in college enrollment quota level. I show that the human capital level only affects the seat allocation and has no effects on external options. Households invest in human capital purely because they can gain economic benefits from the examination. I abstract from these features by focusing on education investment reactions to college seats allocation. I assume that one dimensional talent exists and that two groups have different talent distribution. Competition for college seats transpires among the students within the same group.

Assumption Two groups exist, with $i \in \{0, 1\}$. A continuum of heterogeneous students exists within the two groups, with μ population in group 0 and $1-\mu$ population in group

1. A household has talent type θ, which satisfies a distribution $F_i(\theta)$ in each group. This assumption implies different talent distributions of different group, such that $\theta\tilde{} F_i(\theta), \theta \in \left[\underline{\theta}, \overline{\theta}\right]$. I assume the following:

$$F_i(\theta, \lambda_i) = \begin{cases} \lambda_i e^{-\lambda_i \theta}, if\ \theta \geq 0 \\ 0, other \end{cases} \tag{2..1}$$

with $i \in [0, 1]$ and $\lambda_1 \geq \lambda_0$. In addition, I get the overall population distribution:

$$F_{all}(\theta) = (1 - \mu) F_1(\theta) + \mu F_0(\theta) \tag{2..2}$$

<u>Seat Allocation Rules</u> A continuum of college seats with $p = \left[\underline{p}, \overline{p}\right]$ exists. It denotes the quality of the college, which is drawn from distribution $F_P(p)$, such that:

$$F_P(p) = \begin{cases} 0, if\ p < 0 \\ p, if\ p \in [0, 1] \\ 1, if\ p > 1 \end{cases} \tag{2..3}$$

Two methods are used to allocate college seats that exist: quota (q) case and no-quota (w) case. Without quota, households in the two groups compete in the same auction. Only the human capital rank determines the college seat allocation. I use the case without quota as the bench mark case. In the quota case, I assume that the competition is found within each group. The human capital rank and the group determine college seat allocation.

In the no-quota case, households in the two groups compete in the same examination. Students in the two groups have the same chance of entering a university with the same human capital level. In the no-quota case, the college assigns students to seats in accordance with pure rank order. The mechanism determines the quantile rank of s_i within the overall human capital distribution and then matches student i to a seat at the corresponding quantile rank. For example, the 50th percentile student, regardless of group, matches the 50th percentile college seats. Formally, student j from group i receives the seat assignment, such that:

$$P^w\left(s^i_j|s_{-j}\right) = G^w_1(s_j|s_{-j}) = G^w_0(s_j|s_{-j}) \tag{2..4}$$

Under the no-quota rule, student j's seat depends on his human capital level relative to all other students in both groups. Thus, household education investment behavior only depends on household's talent θ and not on the group.

In the quota case, households compete within each group. College allocates different quota to each province and allocates the seats within each group by rank ordering. The mechanism still allocates μ seats to group 0 and $1-\mu$ seats in group 1. However, it provides a greater fraction of high order seats to group 0 and more low order seats to group 1 than the benchmark case. Formally, student j from group i receives the following seat assignment:

$$P^q_i\left(s^i_j|s^i_{-j}\right) = G^q_i(s_j|s^i_{-j}) \tag{2..5}$$

Under the quota rule, student j's seat depends on the human capital level relative to the other students in the same group. In other words, competition for college seats is found within each group. Thus, a household's education investment behavior depends on the household's talent θ and the group.

I assume that the seat given to 0 has a distribution $F_0(p) = \mu p^\tau$; and the seat give to 1 has a distribution $F_1(p) = p - \mu p^\tau$. τ represents the policy bias on each group. A bigger τ means that the more good seats allocated to 1. If $\tau = 1$, then households in group 0 and 1 have the same opportunity to enter any college, which is exactly same with the case without quota. If $\tau > 1$, then 0 is the province with a higher portion of students ending up with a low p than group 1. Otherwise, 0 is the province with higher portion of students ending up with high p than group 1. Figure 7 presents an example of seats allocation with $\tau > 1$. The first figure is the case without quota, and the college seats of each quality level are allocated according to the population distribution. The second figure shows the case with quota. If $\tau > 1$, then group 1 has a higher portion of students ending up with low p than group 0.

Household Decisions Talent type θ determines the marginal cost of human capital investment. I define the human capital level s. To gain human capital s, students need to spend $C(\theta, s)$, such that $C(\theta, s) = \theta s$. A payoff function $P_j^i(s)$ is determined by the rank of the seat within each group. I obtain the human capital distribution level in the equilibrium, which is $G_j^i(s)$, where j is the quota rule, and i is the province. When students go to a college seat p, they obtain a utility $U(p, s)$, with $U(p, s) = p^\alpha s^{1-\alpha}$ and $\alpha \in [0, 1]$. Taken the rule, $F_P(p), F_0(\theta)$ and $F_1(\theta)$ as given, household i's decision is:

$$\pi^i_j(s,\theta) = \max_{s_j} U\left[P^x_i(s^i_j|s^i_{-j}), s_j\right] - C(\theta, s_j) \tag{2..6}$$

In the no-quota case, all the individuals join in the single tournament. In the case with quota, two tournaments are held. Individuals join in the tournament within their own province. The tournament is an all pay auction process, in which s is the bid and $\pi^i_j(s,\theta)$ is the payoff.

<u>Model Predictions</u> In this section, I study the qualitative prediction on households' education investment choices and college seat allocation.

Proposition 8 *Given the utility function, in the final equilibrium, $s^*(\theta)$ is continuous and strict decreasing with θ.(see Appendix for proof)*

The proposition provides useful insight into how competition shapes incentives. It states that, given a fixed set of college seats, the competition causes the most able students to invest more on their human capital output, and the least talented students to decrease their human capital output. When the marginal cost is equal to the marginal benefit, students get the optimal human capital investment level. Students with higher levels of talent have a lower marginal cost to invest on human capital, for the same human capital level. The most able student has a highest optimal human capital level and invests more.

Proposition 9 *In the all pay auction game, all of the individual's payoffs are equal to the cost in the equilibrium. (see Appendix for proof)*

In the equilibrium, human capital level strictly increases on θ. The only stable equilibrium allocation is that each household's net payoff is zero. Otherwise, at least one household can increase the human capital level and gain net benefits from deviations from the equilibrium.

Proposition 10 *If $\tau > 1$, then every one in group 0 will take more s than no-quota case. If $\tau < \frac{\lambda_0}{\lambda_1}$, then everyone in group 0 will take less s than the no-quota case. If $\tau \in \left[\frac{\lambda_0}{\lambda_1}, 1\right], \exists\theta^* \in [0, \infty]$, then anyone in group 0 with θ larger than θ^*incentive will take more s than the case no-quota case. Anyone in group 0 with θ smaller than θ^*incentive will take less s than the no-quota case*

Figure 8 presents an example of the investment distribution in group 0. If $\tau > 1$, then everyone in group 0 will invest more on human capital than the no-quota case. This statement implies that the reverse AA will increase disadvantage people's investment incentive. If $\tau < \frac{\lambda_0}{\lambda_1}$, the everyone in group 0 will invest less on human capital than the no-quota case. The statement implies that the high level AA will encourage disadvantage people's investment incentive. If $\tau \in \left[\frac{\lambda_0}{\lambda_1}, 1\right]$, then $\exists\theta^* \in [0, \infty]$,anyone one in group 0 with θ larger than θ^*incentive will invest more human capital than the no-quota case. Anyone one in group 0 with θ smaller than θ^*incentive will take less s than the no-quota case.

2.5 Empirical Strategy

In this section, I present our empirical strategy to estimate the impact of college enrollment rate on household educational investment. The objective is to devise a strategy that controls for potential spurious correlation between the treatment and the outcome variables.

I start by regressing household i's education investment proxies in province j at time t (Y_{ijt}) on college enrollment rate in province j at time t (Q_{jt}). The model essentially identifies the average effect of enrollment rate by comparing the households in different regions and different periods. In particular, the first model estimated is expressed as:

$$Y_{ijt} = \beta_0 + \beta_1 Q_{jt} + \delta X_{it} + \gamma Z_{jt} + D_{pro} + D_{year} + D_{pro} \times D_{rural} + \varepsilon_{ijt} \qquad (2..7)$$

where Y_{ijt} is the education investment variables for individual or household i at time t. In this study, I use education spending, study spent studying, and whether the child dropped out of school before graduating from high school, to evaluate the education investment level; X_{it} is a vector the individual and household characteristics variables (including child's age, gender, Hukou, household income, asset, household size, highest education level and other household characteristics); Z_{jt} is a vector of province characteristics (including provincial teacher-student rate of different school levels, and provincial economic variables). D_{pro} includes provincial dummies, which control the provincial level fixed effects; D_{year} includes year dummies, which control the time fixed effects. D_{rural} is the rural household dummy.

To estimate the different reactions of people with different endowment levels, I add the intersection between the educational opportunity variables Q_{jt} and the endowment level variables E_{it} to check the heterogeneity of the reactions. The results in the following equation:

$$\begin{aligned} Y_{ijt} &= \beta_0 + \beta_1 Q_{jt} + \beta_2 Q_{jt} \times E_{it} \\ &\quad + \delta X_{it} + \gamma Z_{jt} + D_{pro} + D_{year} + D_{pro} \times D_{rural} + \varepsilon_{ijt} \end{aligned} \tag{2..8}$$

where E_{it} is the endowment level proxy. In this study, I investigate whether the household has rural Hukou, whether the household has member who graduated from a high school, and has a household income, and whether the household's income level is lower the median income level as the endowment level proxies. The inclusion of these large sets of controls helps in eliminating potential confounding effects that might lead to inconsistent OLS estimates. The center of interest includes the coefficient β_1 and β_2, where Q_{it} is measured by the college enrollment rate in the province at year t. .For example, I use the rural household dummy as the endowment proxy. In this specification, if β_1 is positive, then an average household will spend more on education if there is a better education chance if a better educational opportunity is presented; if β_1 is positive and β_2 is negative, then rural households are less sensitive to changes in educational opportunities.

<u>Instrumental Variables</u> To the extent that some economic and social changes affect the college enrollment system and households' education investment, one might be concerned that the OLS estimates of the model would be biased. In practice, though, even conditional

on the large set of individual, household and regional observed characteristics, a simple comparison of households with different educational change will not necessarily lead to consistent estimates of the effect of interest. As hinted at in the introduction, local economic shock and local education policies - just to quote a few- are all likely to affect both enrollment ratio and households' education investment. For example, the governor of a province, thought highly of education, may implement a series of education promoting policies. If these policies can increase local college enrollment scale and encourage household education investment at the same time, simple OLS estimates of enrollment rate on households' education investment are likely to lead to upward biased estimates of the effect of interest.

To deal with this problem. I propose a set of instrumental variables that will only influence education investment indirectly through their effect on the college enrollment rate. The proposed instrument set is based on the college enrollment mechanism in the Chinese college enrollment system where a college allocates more seats to the students in their provinces and the neighboring provinces than to those in other region. Rapid college expansion happened in 2000 through four methods: (1) building larger universities by merging colleges and universities in the same region; (2) building colleges by upgrading low level schools; (3) building new colleges; and (4) increasing the enrollment quota in each of the universities and colleges. Usually the first three methods create new universities and colleges. However, merging of colleges decreases the number of colleges but increases enrollment quota by return to scale.

Since the 1990s, the number of college in China has increased from about 1,000 to 2,400. New colleges increase the number of enrollments. The number of new colleges and the number of college enrollment capture the supply of local college seats, which is affected by local economic and social factors. In addition, this changes spillover effects on neighboring provinces by increasing the number of college enrollment in the neighboring provinces. Therefore, the number of colleges and the number of newly enrolled students in a province positively correlate the number of college enrollment in neighboring provinces through NCEE.

In the main specification, I control the number of local college and the number of newly enrolled college students in the province, which capture the common trend and local effect of college expansion. I use the number of new colleges and the number of college enrollment in the neighboring provinces as the instrument variables of enrollment rate, which capture the neighboring provinces' college expansion. Neighboring provinces' college expansion is determined by the national level common trends and neighboring provinces' local level economic influence. Controlled by the common trends in the model, they can be considered to affect local college enrollment ratio without directly influencing households' education investment.

I propose two sets of instruments. The first set of instruments is the number of colleges in neighboring provinces, the number of new colleges in neighboring provinces in the previous year, the number of new colleges in neighboring provinces in the three year and the number of new colleges in neighboring province in the last five year. The number of colleges captures the college emergence trends of the neighboring provinces. The last three instrument variables capture the college building trends in neighboring provinces. Emergence and the development of new colleges will increase the number of enrollments. The second set of instrumental variable is the number of college enrollment in neighboring provinces, which directly capture the college expansion in neighboring provinces. The Data section below provides a complete description of the construction of these instrumental variables. The main empirical strategy is to estimate of the equation presented earlier through instrumental variables, where enrollment ratio is estimated through the following first stage regression:

$$Q_{jt} = \alpha_0 + \theta Z_{jt} + \alpha_1 X_{jt} + D_{year} + \epsilon_{jt} \tag{2..9}$$

where Z_{jt} is a vector of instrumental variables that are excluded from the first stage. The set of variables X_{jt} are comprised of the number of college enrollment number and college number at the provincial level. I estimate the results for all children aged 6 to 18.

Specifically, I consider two neighboring provinces: A and B. Colleges in province A increase the enrollment scale by 100. Students in province A obtain 75 new seats. Whereas the students in province B get 25 new seats. The college expansion in province A increases the number of enrollments through the NCEE in both provinces. However, the increase in household's education spending and college expansion in province A can be attributed to the trends such as the increasing education return in province A. The unobserved trend in province A causes the overestimation of the effects of the increasing the enrollment ratio. The variables of college expansion in province B captures the common trends of both provinces and the local trend of province B. The college expansion trends in province A captures the local trend in province A and has spillover effects on the enrollment ratio in province B. The validity of the instruments relies on two assumptions: first, the college expansion of province -i must be correlated with neighbor i's enrollment number; second, these variables must be uncorrelated with province i's preference on education.

2.6 Data

The data have three parts: provincial level local education information, province level college enrollment ratio information, and a household survey information.

Provincial level education and economic information are taken from China Education Yearbooks, the Statistics Yearbooks of every province and the Education Yearbooks of some provinces. The China Education Yearbook (series) is compiled by the Ministry of Education and provides detailed statistics on education. The Education Yearbooks of some provinces are compiled by the Department of Education of each province and reports the information of examinations and school information. The Statistics Yearbooks of each province are compiled by the Statistics Bureau of each province, which reports each provinces' economic and social conditions. These yearbooks report GDP, population, the number of local college students and teachers, and middle and primary school information since the 1950s. The number of local colleges, the number of new college students, and all the high school information since the 1990s are taken. China Education Examination Yearbook (series) is compiled by the Ministry of Education Examination Center and provides detailed information on examination information about NCEE in all provinces. The China Education Examination Yearbook (from 1997 to 2013) and China Education Yearbook (from 1988 to 2012) of each year report the enrollment information for NCEE in some provinces. In addition, some of the education yearbooks of each province and provincial yearbooks also report the NCEE information of the year. I find the NCEE enrollment information (mainly the number of examination takers and the number of enrolled students) after 1990 in these resources.

NCEE enrollment information after 2005 is taken from China National Knowledge Infrastructure[8]. I obtain each provinces' annual NCEE summary reports compiled by each province's education department. This information contains the number of students that took the examination for the four majors, the number of students enrolled in different level universities, the score distribution, and the score threshold for the university level, the enrollment information of each university in each province, and the details of each university's enrollment plan and enrollment results. I collected the NCEE enrollment information before 2005 from the China Education Examination Yearbook and the Education Yearbooks of some provinces. I found college enrollment information in the 1980s from a book called the Composition of NCEE, which was published in 1988.

The household and individual data are taken from the Chinese Household Income Project Series (CHIPS). The data were collected through a series of questionnaire-based

[8]This online database contains information from journals, important newspaper, yearbooks, and thesis. The website is http://www.cnki.net/.

interviews conducted in rural and urban areas in 1988, 1995, 1999, 2002, 2007 and 2008. Individual respondents, reported on their economic status, employment, level of education, sources of income, household composition, and household expenditures. CHIPS can represent the national population. For surveys of urban local households and rural-urban migrant households, a total of the same nine provinces[9] were selected. The rural household survey also covered the nine provinces. CHIPS consists of three parts: the Urban Household Survey, the Rural Household Survey and the Migrant Household Survey. For example, the 2008 survey has 5,000, 8,000, and 5,000 households in the migration, rural, and urban samples, respectively. However, the 2007 survey includes an additional 5,000 rural households and 10,000 urban households. The 1999 survey only has urban households. In this study, I use individuals aged 6 to 18 as the study samples to study education investment decisions.

Using the province and year dummies, I match the CHIPS data with the provincial level education and economic data set. The main outcome variables of interest are reported education spending and household income. To construct my instrumental variables, I define the neighboring provinces as the provinces next to the province and the provinces without boundaries but are found in the same regions. I use the sum of the number of neighboring provinces' enrollment and the number of colleges as the instrument variables.

Summary Statistics Table 1 shows the provincial level education resource information of 31 provinces in China. The data set includes local school, economic, and population informa-

[9] They are Shanghai, Jiangsu, Zhejiang, and Guangdong from eastern China; Anhui, Henan, and Hubei from central China; Chongqing and Sichuan from western China.

tion since 1980. Figure 9 presents the average enrollment rate of different school levels in China. The average enrollment ratio of elementary schools is higher than 99% after 1986[10]. With the elementary school enrollment ratio being almost 100% in all the provinces, I use the number of elementary students as the population of different age cohorts. According to the plan of the Chinese government, except for some extremely poor regions, the majority of the population in China was covered by the nine year compulsory education programs in 1995

Table 2 describes the information of NCEE enrollment. It contains the number of students who took the examination for the two big majors, the number of students enrolled in different level universities, the score distribution and the score threshold at the university level, and the enrollment information of each university in each province. I collect the information from the local education departments of each province, which announces the NCEE information of the province each year, and some NCEE service websites, which provide the details of each university's enrollment plan and enrollment results.

Table 3 presents the information of the instrumental variables. I find that the number of colleges decreased from 96 to 91 between 1995 and 2002. During this period, by merging small colleges into large ones, China has built 431 new universities.[11] The number of college has increased from 1,071 in 1999 to 2,800 in 2015. At the same time, 678 three-year universities were converted to four-year universities. The average number of new students

[10] In1986, the Chinese central government started the nine year compulsory education program.

[11] From the report of Ministry of Education of the People's Republic of China.http://www.moe.gov.cn/publicfiles/htmlfiles/moe/moe_1680/201005/xxgk_88440.html

in neighboring provinces has increased from 167,000 to 1,475,800.

Figure 11 shows that the college enrollment ratio in the most advanced provinces was about 80% before 1998. After the expansion, it reached almost 100% in these provinces. In the most populous provinces, the enrollment ratio increased from less than 10% to about 40%. The average college enrollment ratio increased from 30% to 60% during the rapid college expansion period around 2000. After that period, the increasing rate slowed down. Enrollment ratios of some provinces declined after that period, because the number of test takers increased significantly in these years. Figures 13 and 14 present the enrollment ratio of middle schools and high schools used as proxies for educational change. The high school enrollment ratio did not increase as quickly as the college enrollment ratio during the college expansion period.

Table 4 reports the information of households' and individuals' characteristics. The main education investment proxies are education spending per child, the time children spent studying after school, and whether the child dropped out of school before graduating from high school. The 1988 rural survey does not have education spending information. The average education spending is about 6,000 Yuan. The average household income is 119,754 Yuan. Study duration information is only available in the data after 2000. The average number of hours spent studying is 555 hours. I also considered whether the child worked before 16 years old as another proxy for the reaction to the change in quota. Information on study performance is available after 2000, which control the students' ability. About 42% of children have a good school performance.

2.7 Main Results

In the regression, I keep the observations with college enrollment ratio information and education spending information. I only keep the individuals aged 6 to 20. I use household education spending share and income as the dependent variables. I control for a long list of individual and household attributes in X_{it}, partly because most of the demographic variables are collected at the household level, and partly because human capital investment decisions may be made by the household as a whole, instead of by each individual separately. Within X_{it} , the key variables are age, gender, Hukou, number of family members by age group (0-7 years, 7-16 years, 17- 20 years, 21-59 years, and 60+ years), household income, and other household characteristics. In the regression, I also controlled for the year dummies, provincial dummies, and provincial rural area dummies. The last three variables capture the regional characteristics and time trends. For the controls, when the control variable is missing, I define a dummy based on whether the variable is missing and redefine the missing value to 0. Thereafter, I put both variables in the control list.

Table 5 reports the main results of OLS regression. The coefficient β_1 is positive when I control for other variables. I find that the coefficient of college enrollment ratio is about 0.09. One consistent finding across Columns 3 to 7 is that the advantaged households who have a better education level, and urban Hukou or higher income level, are more sensitive to the increase in educational opportunities. The coefficients of the number of age groups show that households with more children spend more on education. I use the highest education level of the household as the proxy for the household education level. I find that households with better educational levels spend more on education. The coefficient of college enrollment ratio is about 0.09. As college enrollment ratio increases, households with high school degrees or with higher income level are more likely to increase their educational spending share. I also attempt the regression on education spending per child. The main coefficient is similar to the coefficient of the education spending share regression. The coefficient β_1 is positive when I control for other variables. Columns 3 to 7 shows that the advantaged households with better education level or higher income level, or those that are urban households, are more sensitive to the increase in educational opportunities..

Table 6 reports the main results of 2SLS regression. OLS regression tends to overestimate the effect of enrollment ratio. The 2SLS regression results show that the effects of enrollment ratio on household education spending share are around 0.08. As the enrollment ratio increases by 10%, education spending share increases by 0.8%. In addition, in Columns 5 to 7, similar to the OLS regression results, advantaged households with better education level, urban household, or higher income level, are more sensitive to the increase of education opportunity. A thorough analysis using instrumental variables begins with a demonstration of the strength of the instrumental variables proposed. Table 7 describes the results from the first stage of 2SLS regression from the equation, where the dependent variable is the province's NCEE enrollment rate in that year. The coefficients of instrument variables are significant. Reducing the number of colleges increases enrollment ratio. Merging of universities reduces the number of universities but increases the number of enrollments. AS more college students enrolled in colleges in neighboring provinces, the number of enrollments through the NCEE increases. In the regression, the information on the number of local colleges and the number of college enrollment control for the common trends of college expansion.

On average, households' education investment increases as they gain a better chance of entering college. College expansion and AA encourage the targeted group' to improve their human capital level. Households with different endowment and talent levels have heterogeneous responses to the change. Advantaged households with better education level or higher income level or those that is urban households, are more sensitive to the increase in the educational opportunity. College expansion and AA may widen the human capital disparity within each group.

2.8 Robustness Check

The following robustness checks to ensure that the reported effects of the change in education opportunities are not driven by sample selection, or variable construction.

Other Reactions to Enrollment Change To address the concern stating that the overall education investment does not increase and that households may spend more money but less time in education. I attempt other educational spending proxies, that is- study duration and the hour in which parents take care of their children, which are shown in Columns 1 to 6. The coefficient of the enrollment ratio is similar to the main results. The regression on the number of hours spent studying after school shows that the 10% enrollment ratio will increase monthly study hour from 0.7 to 1 and monthly hour of taking care of children from 0.12 to 0.17 on average. Similar to the main results, advantaged households are more sensitive to changes. On average, rural households only spend 0.2 hours more on study as the 10% enrollment ratio increases. Households invest more money and time on children's education as college enrollment ratio increases.

I also checked for other household reactions to changes in enrollment. I regress on the dummy on whether the student drops out of high school and the number of minutes spending on entertainment per week. The results are shown in Columns 7 to 12 of Table 8. The results show that high enrollment ratio makes students less likely to drop out of school and spend less time on entertainment. Similar to the main results, urban households or households with at least one member who graduated from high school responded more to the change in enrollment ratio.

These robustness checks are consistent with the main results. Households will increase time and money spent on education, as they gain a better chance of entering college. Similar to the main results, advantaged groups are more sensitive to the increase in enrollment ratio.

Heterogenous Responses of Different Groups Additional robustness checks consider different age groups. The first group includes children aged 16 to 19, which are the age of high school students. The second group includes children aged 12 to 15, which are the age of middle school students. The third group includes children aged 7 to 11, which are the age of primary school students. The fourth group includes children aged below 7 years. Results in Column 1 to 5 of Table 8, show that the reactions to changes in enrollment ratio do not vary significantly across the first three groups. The coefficients are about 0.4 for the households with children younger than 7 years old and about 0.8 to 0.9 for the households with children older than 7 years old. The influence is similar for households with primary school, middle school, and high school students. Chinese households prepare for their children's NCEE early on. Middle school students compete to enter the best high schools, which usually have significantly high college enrollment ratio. Primary school students compete to enter the best middle school. The competition starts even before primary school. Many children begin their training in piano, art[12], mathematics, and English at about 3 to 4 years old. The signal effects of NCEE enrollment ratio affect all of the age groups who are going to take the examination.

In addressing concerns about potential sample selection, Table 7 conditions the analysis sample on gender and household type difference. Column 1 and 2 of Table 7 shows the gender differences. The coefficient of enrollment ratio is 0.09 for the male sample and 0.06 for the female sample, which means that college expansion will widen the gender gap on human capital investment. In addition, NCEE gives "bonus points" to a minority. Columns 3 and 4 of Table 7 show that a minority family will reduce investment on education investment

[12] The NCEE give extra credit to students with special skill level in sports, music and art. For example, NCEE will give 10 bonus points to the students with the qualification for the grade nine in piano-playing.

as enrollment ratio increases.

Household Reactions on Other Education Chance I check for other education chance proxies on education spending. I regress on local middle school, high school, and local college enrollment ratios, which are defined by the number of students enrolled in the school level divided by the number of students who graduated from lower-level local schools. For example, high school enrollment ratio is defined as the number of students enrolled in the local high school level divided by the number of students who graduated from the local middle school. I find that the college enrollment ratio has the largest effects among the school levels. Other school level enrollment ratios have smaller or insignificant effects on education spending share. With the competition for high schools, middle schools, and primary schools are among students within the same county or city, the provincial-level enrollment information of these school does not have significant effects on households' education investment decisions.

Table 10 reports the effect of the enrollment ratio of differently-ranked colleges on education spending. Rank 1 colleges include the top 112 colleges or universities belong to the 211 projects[13]. The coefficient of rank 1 college enrollment ratio is 0.06, and that of four year college enrollment ratio is 0.04. The enrollment ratio of rank 1 colleges has a greater effect than other four-years colleges but less than the average. The enrollment ratio of three years colleges has greater effects on household education investment because three

[13] The 211 Project is a strategic cross-century project formulated by the Chinese government for the implementation of the strategy of invigorating the country through science, technology and education.

year college accept more students than others.

2.9 Mechanism

Change of Education Return The return to college education has increased after the 1990s. College expansion leads to a higher portion of the population receiving better education, which potentially changes the return to college education. According to the signaling function of education, college expansion may reduce the return to college education. I check for the change of the return to education at the different periods. Table 11 reports the results: the coefficients on the return to the college education are about 0.3 before 1998, 0.6 between 1998 and 2004, and 0.9 after 2004. The coefficients on the return of a high school degree are about 0.3 before 1998, 0.4 between 1998 and 2004, and 0.5 after 2004. By considering at the education gain of the different periods, I find that the college education gain increased from 0.3 before 1998 to 0.9 after 2004. In addition, I regress on rural and urban household separately. The coefficient are about 0.4 for rural household and 0.8 for urban household. The college education gain in urban area is larger than the one in rural areas. The return to college education increased after the college expansion. The result may be caused by the rapid economic growth and transition of the economic structure.

Effects of Public Education Spending Public education spending may substitute or complement private education spending. Table 12 reports the influence of public education investment on education spending. I regress the teacher student ratio of different school level and the intersection with the enrollment ratio. The coefficient of high school teacher student ratio shows that high school teacher student ratio reduces household education spending, which suggests that the public education on high school spending may substitute private education spending. However, public spending on primary schools shows a reverse direction. It shows the complementary effects of the public education of primary school on private education spending. The coefficients of the intersection between the enrollment ratio and teacher-student ratio are only significant in the middle school teacher-student ratio, which means that a middle school teacher-student ratio that is higher may reduce the effects of the college expansion on private education investment. Higher public education spending may crowd out private education spending.

Signal Effects of NCEE Enrollment Information The enrollment information cannot change the number of students takeing the NCEE. Another test is the signal effect of the enrollment information. I regress the number of students who took the NCEE and the number of students who chose two of the big majors-the arts and human sciences major, and the engineering and sciences major. Columns 1 to 3 of Table 14 show that NCEE enrollment ratios has no effects on the number of NCEE takers. However, the enrollment ratio of different major may change students' choice of major. In particular, higher enrollment ratio in the arts and human sciences major and lower enrollment ratio in the engineering and sciences major will encourage more students to choose the art and human sciences major. In addition, because most of the choices in major are made in the first or second year of high school, which is one or two year before the exam, the major enrollment ratio in two years before the exam will have greater effects than the ratio in the current year or the previous year. The enrollment ratio information cannot affect the number of students taking the NCEE but can affect their choices in major.

Factors Affecting Enrollment Ratio The population of age cohort determines the number of NCEE takers. As more students graduate from high school, more students take the NCEE and more students are enrolled through NCEE. In summary, the overall enrollment ratio decreases as more students take the NCEE. As more graduates graduate from high school, the college enrollment ratio decreases.

The number of local college enrollment number has greater effects on local enrollment ratio than the number of student enrollment in neighboring province. I add the number of local college enrollment and the number of student enrollment in neighboring provinces in the regression. Columns 4 to 6 of Table 13 report the results and show that high college enrollment in local and neighboring provinces increases the number of students taking and enrolled through the NCEE. The number of local college enrollment have much bigger effects than the number of enrollment in neighboring province.

Long Term Effects on Individual Income and Education I estimate the long term effects on individual income level and education level. Using CHNS data, I select the observations with income and education information after turning 18 years old, and household information before turning 18 years old. I control for individual characteristics variables after turning 18 years old, such as age, age square and gender; household characteristics before age 18, with household income, asset level, number of elders, number of children, number of laborers, highest education year level, and rural Hukou dummy. The results are shown in Table 15.

The effect of high enrollment ratio in education level is shown in Columns 6 to 10 of Table 15. Columns 6 and 7 shows that the coefficient of enrollment ratio is about 0.056, which means that 1% increase in enrollment ratio will increase people's individual education level by 0.056% after provincial fixed effects are controlled for. Columns 8 to 10 shows that the average effects are weakened by the heterogeneous effects. Columns 8 and 9 show that high college enrollment has greater effects on increasing the education level of rural students or the students coming from households with high school degree holder. Column 10 shows that a high number of college enrollment increases the education level of the households. But the effect is less for the households with higher income level.

The effect of high enrollment ratio on individual income is shown in Columns 1 to 5 of Table 15. Columns 1 and 2 show that on average, a high number of college enrollment increases annual individual income. The coefficient of enrollment ratio without provincial fixed effects controlled for is higher than the one where the fixed effects are controlled for. Regional inequality widens the income disparity caused by enrollment quota differences. When provincial fixed effects are controlled for, the coefficient of enrollment ratio is approximately 3, which means that a 1% increase in enrollment ratio will increase people's individual income by 3%. Columns 3 to 5 reveals that the coefficients of the intersection between the enrollment rate and endowment level or household variables are insignificant, which indicates limited heterogeneous effects.

People in provinces with more college seats tend to gain a higher income level because college expansion increases individual income. NCEE college enrollment rules widen the regional income gap. College expansion has greater effects on advantaged households, which enlarges the human capital and income disparity within each province.

2.10 Conclusion

The NCEE enrollment system and the rapid college expansion in China provide an excellent opportunity to deepen our understanding of the manner in which households with heterogeneous talent and endowments compete for different levels of college seats. The NCEE enrollment system has the same seat allocation rule as AA, which is based on the region rather than on race or other demographic characteristics. In addition, enrollment history and enrollment rate are publicly known to all households and students. The rapid college expansion that happened after 1998 enables a natural experiment to investigate to the effect of the increase in college enrollment rate on household education investment. I use the number of colleges, the number of new colleges, and the number of college enrollment in neighboring provinces as instrument variables for the enrollment ratio of a province. I use the education spending share, time spent studying after school, and dropping out before finishing high school as the proxies of household education investment. I find a large, positive, and significant effects of high enrollment ratio that encourage household education investment. I also find that the advantaged households with higher income levels, better education levels, and urban Hukou are more sensitive to the increase in the number of enrollment.

The encouraging effects of the increase in enrollment that are found in the current study imply that an AA policy and college expansion encourage targeted households' education investment. However, the heterogeneous response of households shows that the policy encourages the advantaged households more than others. The finding implies advantaged households in the targeted group gain higher human capital level and benefit more from the policy. The reverse AA which gives more college seats to more advantaged provinces, widens human capital disparity among regions and the AA will increase the human capital disparity within regions.

Interestingly, AA no only affects the households' education investment, but it could also create a number of socioeconomic implications including social mobility, economic inequality, and human capital distribution among group. Assessing the effect of AA on these socio-economic issues will be a promising direction for future research.

CHAPTER 3.

HOW DOES ONLINE PIRACY AFFECT FILM REVENUE IN CHINA?

The impact of online piracy on genuine products sales is under debate, because researchers cannot find representing proxies to evaluate piracy levels. In Chapter 3, I estimate the impact of online piracy on movies' box office performances in China from 2006 to 2013 using a unique dataset that reports the pirating data for 1,039 wide-release movies from several file-sharing websites. Using these piracy-level proxies and Chinese box office data, I estimate that movie piracy caused substantial box office losses, that the substitution elasticity of the consumption of pirated movies on consumption in theaters was small, and that government anti-piracy policies reduced box office losses, but only in the short term. I estimate that the average revenue loss caused by piracy was about 30 percent.

3.1 Introduction

The media industries believe that copyright infringement causes billions of dollars in losses. the Motion Picture Association of America (MPAA) estimated that piracy costs the U.S. movie industry some $20.5 billion per year in 2011. The industries, they contend, fallaciously assume that every person who pirates a work would have otherwise purchased it at full price (Lee, 2006). How much does online piracy affect the sale of genuine products? The difficulties of observing piracy behavior, the sample bias caused by special study groups and the correlation between product sale and the intensity of piracy make this a hard question to answer.

In this article, I evaluate the substitution effects of the pirated movies on their genuine counterparts, using the difference between a movie's theatrical release date and the date on which the pirated version appears as the proxy for piracy duration (piracy lag), and the number of searches for movies to pirate on search engines as the proxy for piracy intensity (piracy movie search amount). Both proxies could represent the piracy level of the whole population. The setting I consider is one in which online movies provide consumers a low-cost, low-quality alternative to the movies at theaters, and theaters adjust the ticket price according to the demand change caused by this alternative. The later the pirated movie is available, the more consumers lose patience waiting for the free movie. In my estimates, I find that in China, pirated movies caused box office losses of 65 percent, and daily revenue loss of about 70 percent. The paper also finds that only a small portion of Chinese Internet users are affected by piracy variation, which suggests that the substitution elasticity of pirated movies on theater movies is very small in China.

To structure my empirical analysis, I develop a partial equilibrium model with two types of movie supplies with different quality, prices, and frames. Theaters adjust the price of movies according to quality and piracy supply. Because of preference heterogeneity, consumers with a high willingness to pay purchase the tickets for high-quality theater movies. By assuming a waiting cost, given the ticket price, the later a pirated movie comes out, the more consumers choose to watch the movie in theaters. The model makes the following predictions: First, piracy resources reduce theater attendance, but increase the overall movie audience. Second, if the majority of the population cannot access theater movies, even if the piracy level is high in this market, the substitution elasticity of a given movie's piracy level on its box office is low.

To evaluate the impact of pirated movies and to test the prediction of the theoretical framework, I have constructed, based on Internet sources, a new dataset with thousands of movies of the following data: overall box office revenues weekly boxoffice revnenues, and daily box office information in each theater of a theater chain; prices, movie characteristics, global release schedules, the number of searches for movies on search engines, pirated movie versions, and available time on piracy websites in China. The dataset consists of 1,039 movies that were in wide release from 2006 to 2013. Data on piracy resources are rarely available to collect, especially since most of the piracy websites created before 2010 were shut down in a series of anti-piracy actions, and the survivors have strict membership registration systems and are open only to registered members. I found the seven websites[1] with pirated movies uploaded before 2006 and purchased memberships in them in order to collect information about their piracy activities. I use the information to construct piracy proxies.

I use two empirical approaches to estimate the impact of piracy. First, I use reduced form methods to get the elasticity of box office revenue in relation to the level of piracy. To deal with the potential endogeneity caused by the correlation between movie quality and supply gap, I use the government's anti-piracy policies which shut down the main movie websites and cut the supply of pirated movies in the short term as the natural experiment to

[1] The Chinese government shut most of them down in anti-piracy movements campaigns in 2014 and 2015.

estimate the revenue loss caused by piracy in the DID regressions. The number of searches for the name of a movie on search engines, which represents piracy intensity, is positively correlated to the movie's quality and its box office revenue. I use the number of searches conducted before the free piracted movie is available as the instrument variable of piracy intensity to get an unbiased estimation. Secondly, I use the theoretical model to make a structural estimation of the parameters. I use the estimated model parameters to make some predictions and find that the average box office could increase by about 30 percent if the piracy supply were permanently cut off.

3.2 Literature Review

A large theoretical literature (Novos &Waldman, 1984; Johnson, 1985; Takeyama, 1997; Yoon, 2007; Belleflamme & Peitz, 2014; Bae & Choi, 2006) has argued that the availability of a pirated good reduces a firm's profits. The degree to which piracy affects social welfare, however is still uncertain. In a static model, piracy is harmful to firms, but benefits consumers in the short term (Belleflamme and Peitz, 2010). Many studies find that piracy increases consumer welfare by providing products' information (Peitz & Waelbroeck, 2006; Gopal, Bhattacharjee & Sanders, 2006), reducing consumption cost (Ahn & Yoon, 2009) and exhibiting nextwork effects (Belleflamme, 2003; Belleflamme & Peitz, 2014). In the long term, changes in profits typically decrease firms' incentives to provide high-quality products (Novos & Waldman, 1984; Bae & Choi, 2006), but do encourage more varieties ((Johnson, 1985). In this article, without considering the long-run supply effects, the static model shows that the availability of a pirated good reduces the firm's profits, but increases overall welfare by increasing the number of consumers who are able to watch the movie.

The difficulties of observing piracy activities present the most serious obstacles to estimating losses due to piracy. Piracy websites make great efforts to hide piracy resources from governments, which are required by copyright laws to shut down the websites for these piracy activities. Most piracy websites have some mechanisms to avoid these problems. For example, some file-sharing websites have strict membership systems. Only members can access their piracy resources, and it's hard to become a member. Legal penalties for individual copyright infringement also give the users of pirated products incentives to hide their piracy behavior. From 2003 to 2006, for example, the Recording Industry Association of America sued more than 20,000 music fans for file sharing (Lambrick, 2009).

Given the difficulty of observing the illegal activities, empirical studies have used a small group's piracy level or other indirect methods to evaluate overall piracy levels. Oberholzer-Gee and Strumpf (2007) monitor an online service to develop product-specific measures of downloading activity over time. They use instrument variables (such as file size or German school holidays) in order to deal with a potential positive correlation due to unobserved heterogeneity; they find no displacement of music sales by piracy.

Some studies use individual-level survey data to ask whether persons who also engage in more unpaid consumption engage in more or less paid consumption. Rob and Waldfogel (2004) conduct a survey of university students and find that each album download reduces purchases by about 20 percent. Some studies examine whether products that are more often downloaded tend to be purchased more or less. Milot (2014) investigates the concept of lost sales at the box office related to the unauthorized downloading of "Cam" copies (clandestine video recordings of films made during projections in theaters) of widely released movies at a popular BitTorrent website and finds that the unauthorized downloading of Cam movies has no important effects on the box office sales of individual movies. The low quality of Cam movies means that they are not a good substitute for high-quality movies in theaters. In my article, I estimate the box office loss caused by TS versions and DVD versions with better quality and enough piracy supply variation.

There are concerns regarding the generalizability of the results and selection bias caused by the sample. Surveys or website tracking can cover only a very small segment of the population. Without complete knowledge of the distribution of products in the population through piracy as well as legal consumption, the piracy level estimated by the methods above can hardly represent the total piracy intensity. In addition, most estimations have an unresolved simultaneity problem since, in the data, the people who download a lot of music files also tend to be heavy purchasers of music (Liebowitz, 2008). The biased sample underestimates the true totals. In addition, most studies evaluate the short-term substitution effects of piracy products and genuine products. A large portion of consumers are not reacting to piracy variation and only choose piracy products in the short term, possibly switching to legal products in the long term.

The time difference between the two supplies is a proxy for piracy level without a sample bias problem. The digital piracy products supply to the global market at same time.

Once the first pirated resource comes out on one website, other websites copy and share it. Ma, Montgomery, Singh, and Smith(2011) use the availability of pre-release movies as a piracy level proxy to estimate movies' piracy loss. Their results show that pre-release piracy reduce box office revenue by 8 percent. The current article uses proxies similar to those used in the two studies above. In Ma et al. (2011), only a very small portion of movies in the US market have a pre-release version. In contrast, I use the time difference between the two supplies with enough variations to evaluate the problem. The supply method solves the generalizability problem, but can't show the intensity of piracy behavior. This article uses the number of searches for free movies on search engines as the piracy intensity proxy.

The article contributes to a growing literature that attempts to estimate the impact of piracy resources on the sale of legal products. A distinguishing feature of my approach to evaluating piracy loss is the use of a structural model to quantitatively study the roles played by the mechanism in explaining the observed effect. An advantage of this approach is that it is likely to improve external validity: an estimated model can be used to inform quantitative predictions about the impact of free piracy resources on legal products in different industries and countries.

3.3 Background

In this section I discuss some of the essential background features of the Chinese motion picture market, piracy websites, and the data collected in order to analyze how pirated movies change the sales of legal products.

China, the second largest motion picture market in the world, generated $6.78 billion in box office revenues in 2016. It has about six hundred million Internet users and a 44.1 percent internet penetration rate. Thousands of websites and numerous peer-to-peer file-sharing networks, FTP services, and free video websites make copyrighted works available for free to Internet users in China (Priest, 2006). The majority of the population, especially the segment living in small cities and rural areas, has no theaters nearby. Most Internet users in China turn to free online movies as the first choice for movie viewing, and their piracy behavior could hardly be "relatively large" enough to be punished by law, which is considered an important reason for the high piracy level in China (Priest, 2006). There are strict censorship systems[2] and protectionism in the Chinese film industry. All screenplays in China must be approved by the State Administration of Radio, Film, and Television (SARFT). Foreign producers can enter the Chinese market only by importing their movies as quota movies,[3] or as movies co-produced with local producers.[4] Foreign movies could be imported through buyout,[5] but this only allows the producers to sell the copyright to the Chinese distributor; they cannot participate in distribution. SARFT also blocks

[2] The Chinese State Administration of Press, Publication, Radio, Film and Television (SARFT) is responsible for censoring any materials that offend the sensibilities of the Chinese government or Chinese cultural standards. In 2001, the SARFT issued mandatory guidelines for film content that highlights 31 categories of prohibited content ; including violence, pornography, and anything else that may "incite ethnic discrimination or undermine social stability."

[3] The quota was 20 until 2013, when it was raised to 34.

[4] There are strict requirements on movie content and investment share. The key requirements are that one or more Chinese production entities accredited by the SARFT must be participants; at least one third of cast members must be from the mainland; and the story must have enough Chinese elements.

[5] A buyout movie is a foreign movie acquired by a Chinese local distributor at a fixed price to be released in China (Chinafilmbiz, 2012).

foreign movies during the peak season to increase the market share of domestic movies' (McCutchan, 2013). Censorship and policies of protectionism cause release delays which reduce the time lag between piracy supply and theater supply and at times make pirated movies available earlier than theatrically released movies. In China, the absence of legal DVD sales and on-demand Internet streaming media, such as Netflix, means that the box office is almost the only revenue resource for producers and distributors, which makes it is easier to capture the direct piracy loss in this market.

The level of copyright infringement in China is widely regarded as one of the worst in the world. A 2009 survey by the EntGroup[6] shows about 98 percent of Chinese Internet users have used the Internet to access movies and that free movies from the Internet are the first choice of about 80 percent of users. The piracy resource comes mainly from foreign piracy websites and some theaters in China. The piracy websites upload the movies or links to the movies on their webpages as soon as they get access to them. The websites also check the quality and label the version of these movies. Different versions of pirated movies are marked as Cam, TS, BD, DVDscr, R5, HD, or Blueray, depending on piracy method and quality. The Cam version, which appears online after the first preview or premiere of the film and has very low quality, has no substitution effects on theaterically released movies (Milot, 2014). The TS version is a copy shot in an empty cinema or from the projection booth with a professional camera mounted on a tripod and is directly connected to the sound source. The TS version's quality is much better than the Cam version and clear

[6] EntGroup, Inc. is a consulting firm that specializes in the Chinese movie industry.

enough to watch. It appears online several days to weeks after the first preview or premiere of the film. I define “DVD version” as any DVDscr, BD, R5, HD or Blue ray versions that come from DVD resources and are of higher quality than the TS version. This article estimates the box office loss caused by the TS and DVD versions. The Chinese government rarely makes efforts to punish individuals’ piracy activities, but has launched occasional campaigns against infringers in response to pressure from foreign and domestic copyright owners (Priest, 2006). Its typical action is to shut down major piracy websites and prohibit illegal DVD sales on the street. When these campaigns end, however, new piracy websites, come out to replace the old ones. Most of the current piracy websites were founded after 2010. The survivors of these campaigns usually have strict membership systems and allow only members to access piracy resources. I bought memberships in six websites and collected piracy information from them.

3.4 Data

The data for this study has three parts: box office data, piracy level data, and film characteristics data.. The box office dataset has three parts: the movies’ overall, weekly, and daily box office revenues at each theater. The movies’ overall box office data are from SARFT, which collects this data directly from the automatic ticketing system of each theater for the theaters that use this system.[7] These data pertain 1,039 movies in wide release from

[7] The theaters without an automatic system report their box office to SARFT every month. In order to prevent the hiding of box office sales, SARFT requires theaters to provide paper receipts with movie information to consumers. The real income received by the theater (rather than the ticket price) from discount tickets, group tickets, and theater VIP tickets are counted in the box office. The box office is the

2006 to 2012. Figure 26 reports summary statistics for the wide-release movies' box office revenue by year. It shows that from 2006 to 2013 the median box office increased. The daily box office data are from the Wanda Theater Chain the biggest theater chain in China with 15 percent of market share. The data are from December 2011 to June 2013.[8] Table 21 shows that the movies are allocated an average of 6.48 screens in each theater every day. The average daily audience is 225 people. The average price is 60 yuan.[9] Daily box office is 14,984 yuan. China's weekly box office data are from the Pacific Website and EntGroup.[10] 980 movies have weekly box office data. Hong Kong box office data is from Box Office Mojo. This company's international section covers weekly and historical box office information for Hong Kong.

Piracy-level data comprise the second part of the data. Figure 25 provides an example of these methods, using them to estimate the piracy level of the movie Hunger Games (2012).[11] This figure shows the opening delay, the date that the pirated version of the movie appeared, and the number of searches. There is a time lag between the US opening day and the Chinese opening day, which is called "opening lag." Between the two opening days,

---

money received by the theater chain, rather than the consumers' payment. The two values are different because some consumers buy tickets from ticketing websites that charge an extra fee (around 10 percent).

[8] Wanda posted each movie's daily box office data from its ticketing system. Because the Wanda ticketing system doesn't count price deal, the average price in this data is higher than the real average price.

[9] Usually, in Chinese theaters, the half-price ticket is about 25 to 40 yuan. The ordinary movie's price is about 60 to 70 yuan. The price of a new blockbuster is about 80 to 90 yuan. 3D and IMAX movie are more than 100 Yuan.

[10] I choose two resources, because, although Pacific was the first box office website to exist in China, it stopped providing box office data in 2010.

[11] The film's US release date was March 23, 2012. Its Chinese release date was June 14, 2012.

there are TS versions and DVD versions of the movie's available dates. The gap between the date the pirated version uploaded online and the movie's release date in China is the Δt, which is called "TS lag" and "Clear lag." The number of searches for "the hunger games download" on Baidu after the day the TS version became available can also be a proxy for piracy level. Searching after the TS available is a proxy for an attempt to pirate movies, which is called "name download search amount." The movies' opening date information comes from Mtime.[12] I collected the available piracy date from six Chinese online forums one file-sharing website,[13] and an international movie-sharing website.[14] On these websites, one can find movies with upload date and resource-type information. The earliest one is the version's date of availability.

Figure 29 shows the distribution of Δt. "Mainstream"[15] movies with very low quality become available to pirate two months after their release in theaters. The variable "the number of name download searches" records the number of times that the movie was downloaded, according to the Baidu index, a value that means the search amount per ten million users per day. I collected "movie name" and "movie name download" (both in Chinese) as

[12] Mtime, a China-based movie web portal, has dedicated itself to providing four categories of movie services: China's largest movie/TV database, China's top movie review and critics service, the only cinema and show time search engine in China, as well as the largest movie marketing and promotion services (Chinawhisper, 2012).

[13] The six forums are UUNiao, Feiniao, Shengchengjiayuan, the Third World, BTbbt, and Zhuzhu. The file_sharing website is Dygod.

[14] The Pirate Bay is a website that provides torrent files and magnet links to facilitate peer-to-peer file sharing using the Bit Torrent protocol (The Pirate Bay, 2013). It provides a movie piracy resources to Chinese Internet users. Even when the Chinese government blocks this website, China's Internet users are still able use the technique called "Fanqiang"—-breaching the Great Firewall of China —to reach the site.

[15] Mainstream" movies are usually produced for the Communist Party political propaganda.

the estimation of movie popularity and willingness to download. I use "film name download" between the date of TS availability and the movies' first showing month as a proxy for piracy level. In general, after the date of TS movie's availability, these attempts successfully find the piracy resources to download. Table 19 provides the summary of the variables generated from the Baidu index. On average, the search amount of "name download" is 35,533. Given the fact that China has 0.6 billion Internet users and Baidu has about 80 percent of the market share, the average download attempt is roughly about 1.5 million. In the estimation, omitted variables could cause a downward bias. I use the number of searches for "movie's name download" before the piracy resource is available as the instrument of the number of searches for "movie name download" after the piracy resource is available. The number of searches for "movie's name download" before the piracy resource is available can be viewed as the willingness to consume a pirated movie, which correlates to the piracy level, but does not affect box office directly. Table 19 shows that the average search number in the opening week is 11,993, which means that about 0.7 million people try to search for information about a movie during its opening week.

I obtained the film characteristics data from Mytime, which contains budget, director, actors, and actresses,[16] film length, film type, producer information, and distributor information. The Chinese public holiday system[17] consists of both solar calendar holidays

[16] The Chinese director level is ranked by Forbes Celebrity ranking. The foreign director rank is from the Celebrity Networth website. The ranking of the starring foreign actor and actress is from the Vulture websites. The authors define the director, actor and actress level dummies according to these ranks.

[17] I get public holiday dates from the Chinese central government's public holiday schedule, and then generate holiday dummies as season control.

and lunar calendar holidays. Month or week dummies are no longer enough to capture the movie season. If a movie is distributed in the period between two weeks before the season and the end of the holiday, I assume that the movie is distributed during the season and assign the season a dummy value of 1.

3.5 Model

In this section, I present a partial equilibrium model. The model serves two purposes. First, it provides an estimation framework that I use to quantitatively assess the substitution effect of movie piracy on theater movies. Second, it delivers qualitative predictions that I use to guide my empirical evaluation of the change in piracy supply lag. My interpretation is that delays in the piracy supply attract more people to watch movies released in theaters.

The economy consists of M consumers and a single theater. The theater and all the consumers make their decisions simultaneously. In my empirical application, there is theater supply and piracy supply, with time lag between the supply availability of the two. Consumers gain utility from watching earlier. Taking movie quality, supply lag, and demand as given, the theater chooses the optimal price for each movie. Taking the movie quality, supply lag and price as given, consumers choose the way to watch the movie or not to watch it. In my empirical application, I work with data (on qualities, supply lag, and box office revenues) that refer to movies. While my empirical setting considers 1,039 movies, for simplicity, the model is static.

Consumers have taste indexed by θ, with an exponential distribution: $f(\theta) = \begin{cases} \lambda e^{-\lambda\theta}, \theta \geq 0 \\ 0, otherwise \end{cases}$.

Only a portion of n consumers can access the movie in theaters, with $n \in [0, 1]$, which captures theater availability for the population. Movie j's ticket price is p_j; the quality is Q_j; and c is the opportunity cost to watch the movie. α is the discount effect to watch the movie online, with $\alpha < 1$, which means that the online version is always of lower quality than the theater version. Δt_j is the gap between the piracy availability date and the theatrical opening date. To capture the supply time difference, I assume that there are additional costs to waiting. If theaters distribute the movie earlier than the online resource does ($\Delta t_j > 0$), watching the movie in the theater has an additional benefit $\gamma \Delta t_j$; otherwise, watching the movie in the theater has a cost $-\gamma \Delta t_j$. $Q_j \in [\underline{Q}, \bar{Q}]$ and $\Delta t_j \in [\underline{\Delta t}, \bar{\Delta t}]$. To simplify the model, I assume $\bar{Q}\theta/\gamma > \bar{\Delta t}$, which captures the fact that almost all the movies have piracy resources within the few months after their release. The consumer i's utility [18]is:

$$u(\theta, Q_j, p_j, \Delta t_j) = \begin{cases} \theta Q_j + \gamma \Delta t_j - p_j - c, \text{if watched in theater} \\ \alpha \theta Q_j - c, \text{if watched online} \\ 0, \text{otherwise} \end{cases} . \quad (3..1)$$

Consumers maximize their utility by choosing their method for watching a movie. Consumer choices, given the taste distribution, are described in Figure 24. The people who watch the movie in the theater have type θ, such that: $\theta Q_j + \gamma \Delta t_j - p_j \geq \alpha \theta Q_j; \theta Q_j + \gamma \Delta t_j - p_j \geq c$. The people who watch the movie online have type θ, such that $\theta Q_j + \gamma \Delta t_j - p_j <$

[18]The utility setting is similar to Yoon (2002), Belleflamme (2003), Bae and Choi (2006) and Belleflamme and Peitz (2010). They have utility similar to $\theta q - p$. I have added some other costs to this model.

$\alpha\theta Q_j; \alpha\theta Q_j \geq c$. The lowest type to watch the movie in the theater is $\underline{\theta}(p_j) = \frac{p_j+c-\gamma\Delta t_j}{(1-\alpha)Q_j}$.

In China, the pricing strategy theaters employ is similar to monopoly pricing. Before its release, each movie's distributors negotiate with the theater chain, and determine the ticket price together. Usually, there is a unique price in one region. Theaters take α_j, Q_j, Δt_j and the consumers' reaction as given. The theater incurs no cost to release the movie[19]. Choose p_j to maximize movie j's profit:

$$\pi\left(p_j\right) = \max_{p_j} Mnp_j \int_{\underline{\theta}(p_j)}^{\infty} f\left(\theta\right) d\theta = \max_{p_j} Mnp_j e^{-\lambda\frac{p_j+c-\gamma\Delta t_j}{(1-\alpha)Q_j}}. \tag{3..2}$$

By the first-order condition, I get a movie's price: $p_j^* = \frac{(1-\alpha_j)Q_j}{\lambda}$. People who watch a movie in the theater have the type: $\theta \in \left[1 + \frac{c-\gamma\Delta t_j}{(1-\alpha)Q_j}, \infty\right]$. The box office revenue of a movie is:

$$\pi_j^{piracy} = Mn\frac{\left(1-\alpha\right)Q_j}{\lambda}e^{-1-\lambda\frac{c-\gamma\Delta t_j}{(1-\alpha)Q_j}}. \tag{3..3}$$

The number of theater consumers is: $N_{theater}^{piracy} = Mne^{-1-\lambda\frac{c-\gamma\Delta t_j}{(1-\alpha)Q_j}}$.The number of movie consumers is:$N_{all}^{piracy} = e^{-\lambda\frac{c}{\alpha Q_j}}$. I set a case in which only theater movies are available. The utility of i to watch movie j is given by:

$$u\left(\theta, Q_j, p_j, \Delta t_j\right) = \begin{cases} \theta Q_j + \gamma\bar{\Delta t} - p_j, \text{if watched in the theater} \\ c, \text{otherwise} \end{cases}. \tag{3..4}$$

[19]The main costs for Chinese theaters are the facility cost and the rent cost. These costs do not affect certain movies' pricing in the short term. In China, distributors don't charge theaters money for movie copies. There are no fixed costs for theaters to show movies. In addition, the marginal cost to show digital copies of movies is very small. To simplify the model, I assume that the marginal cost is 0.

In this case, I get that a movie's box office without piracy is $\pi_j^{no\_piracy} = Mn\frac{Q_j}{\lambda}e^{-1-\lambda\frac{c-\gamma\bar{\Delta}t}{Q_j}}$; a movie's price is $p_j^{no\_piracy} = \frac{Q_j}{\lambda}$; and the number of audience members for a movie without piracy is $N_{theater}^{nopiracy} = Mne^{-1-\lambda\frac{c-l*\bar{\Delta}t}{Q_j}}$.

Proposition 11 *The box office of theaters is strictly increasing in Δt_j and n.(See Appendix for proof)*

The later a pirated movie comes out the more people get tired of waiting for free movies and choose to watch it in a theater. So, a larger a pirated movies availability gap of pirated movies results in a higher box office. The more people who can access theaters means that more people are potential theater movie consumers.

Proposition 12 $N_{all}^* > nN_{all}^* > N_{theater}^{nopiracy} > N_{theater}^*$ *:When free movies are available, among the population who can access theaters, there are fewer people going to theaters and the price is lower. Based on the two facts above, the box office revenue is lower when free movies are available. Among the n population who can access theaters, more people watch movies when free movies are available. In addition, even if the population can access theaters, movie audiences (for both pirated and theater movies) are larger than without piracy.(See Appendix for proof)*

More people watch movies when free movies are available. Even if no piracy existed and everyone could access theater, not all movie consumers watch in the theater when free movies are available. In addition, there is $1 - n$ population that cannot access theaters. People's piracy behavior won't affect theaters' box office unless theaters are available to them. In reality, I can only observe $MN^{piracy}_{theater}$ and p^{piracy}_j. The real revenue loss is: $MN^{nopiracy}_{theater} p^{nopiracy}_j - MN^{piracy}_{theater} p^{piracy}_j$.

Proposition 13 *Among the n population that can access theaters, theaters' piracy loss on these people is strictly decreasing in Δt_j. Decreasing Δt_j decrease the number of people who watch movies in theaters. The change of Δt_j does not affect the overall number of audience members.(See Appendix for proof)*

Δt_j captures the differences of the supplies' availability time. The piracy loss variation caused by it is a short-term loss. Δt_j doesn't affect the overall number of audience members and just changes the proportion of people watching in theaters. Good movies can always attract a certain audience size, but the later the pirated movie comes out, the more the audience loses patience and goes to a theater.

3.6 Empirical Results

The theoretical prediction provides explanations for the reaction of theaters to changes in supply and demand and the impact of a changing piracy supply lag on box office revenues. I use both structural estimation and linear regressions to evaluate loss due to piracy. To relate the static model to my empirical setting, I take the simplest possible approach and assume that all players make their decisions simultaneously. This means that I do not take into account the dynamic price change and use only the average price in the first week as the proxy for a movie's price.

Structural Estimation In this section, I use a structural model to evaluate the parameters of the theoretical model and consider welfare implications based on the model.

I obtained the marginal substitution rate of the consumption of pirated movies on the consumption movies in theaters from the linear models. Both the demand piracy products and the demand for legal product are determined by movie quality, which means the linear model may have simultaneous problems. The structural model with the equations of both the quality effects on piracy demand and the quality effects on legal products' demand, rather than estimating the marginal effects in equilibrium, is estimating the two demands separately and evaluating the substitution rate of the two products without simultaneous bias.

In the linear model, I estimate the impact of the supply change of the piracy products on the demand for their theatrically released counterparts. The theory model, however, predicts that movies with different qualities may have different substitution effects between the two. The structural model takes heterogeneity into consideration. Based on the parameters of the structural model, I can check the substitution rate between the two products and the effects of the supply of piracy products on the demand for theaterically released movies based on differences in movie quality.

The linear model can obtain only the marginal effects of the supply of or demand for piracy products on the demand for their genuine counterpart, by fixing market size, the supply of legal products and other exogenous factors. The information from the linear model is not enough to enable a welfare analysis and policy experiment. In the structural model, with the assumptions and the parameters from the estimation, I can estimate the effects of the fixed factors on the demand of both the legal and piracy products, and on the substitution rate between the two products. From the estimation, I can make the welfare evaluation and policy experiment.

In the structural estimation, however, I need to make some assumptions to close the function forms. For example, in the structural model, I assume that the ticket price is a constant for each movie, which ignores theaters' reactions to the demand change and causes the estimation bias in the substitution rate. But in the linear regression, I am able to use the information from the daily box office data to control for these effects. These assumptions of the structural model could cause a different estimation bias from the linear estimation. Both the linear estimation and the structural estimation have their own disadvantages and advantages. I therefore show the results of both linear and structural models.

I collected data on the average ticket prices in the first week, which allow me to exploit theaters' reactions to changes in demand; I also collected movie-rating data from move-rating websites, which are used as a proxy for overall consumers. The following parameters c, γ, λ and α in the theoretical model are estimated in this part. In the data, Q_j is proxied by the rating on Mtime; p_j is proxied by the price in the opening week; Δt_j represented TS lag; and N_{all}^{piracy} is proxied by the number of people rating each movie on Douban. In the data, there are 793 movies with all the information. I get the following functions from the model: $N_{all}^{piracy} = Me^{-\lambda\frac{c}{\alpha Q_j}}; \pi_j^{piracy} = nP_j e^{-1-\lambda\frac{c-l*\Delta t_j}{(1-\alpha)Q_j}} M$ and $p_j = \frac{(1-\alpha_j)Q_j}{\lambda}$. Because the price proxy is not the real average price, $p_j^* = p_j^{\text{1st week}} + \varepsilon_1 = \frac{(1-\alpha_j)Q_j}{\lambda} + \varepsilon_1, \varepsilon_1$ is the measurement error.

Besides the four parameters mentioned above, there are also two things unknown M and n in the model. All three functions can be tranformed to linear functions, given the exponential taste distribution characters. In the linear regression, I can use year dummies to capture the effects of these unknown parameters and variables. I use the number of people rating on Douban as the proxy for N_{all}^{piracy}. I assume that N_{all}^{piracy} and N_{Douban} have the following relationship: $R_t * N_{all}^{piracy} = N_{Douban}$, where R_t is the ratio of Douban members among the overall movie consumers. Because both the number of active Douban members and the movie market were changing every year, R_t changes every year. By this assumption, I get: $\ln N_{Douban} - \ln R_t = \ln N_{all}^{piracy} + \varepsilon_2, \varepsilon_2$ is the measurement error of the number of audience members.

By the three equations, I get following linear functions: $\ln N_{Douban} = -\lambda\frac{c}{\alpha} \times \frac{1}{Q_j} + Dummies_t$ and $\ln \pi_j^{piracy} = \ln p_j - \frac{\lambda c}{(1-\alpha)} \times \frac{1}{Q_j} + \frac{\lambda\gamma}{(1-\alpha)} \times \frac{\Delta t_j}{Q_j} + \beta Dummies_t + \varepsilon_3$.The error term ε_3 is the unobserved factors captured by the unobserved demand shock. In the first equation, M and $\ln R_t$ are captured by $Dummies_t$ and the constant. In the second equation, M and n are also captured by $Dummies_t$ and the constant. From the functions above, I can estimate the four parameters c, γ,λ and α. I therefore get three linear equations with error terms:

$$p_j^* = \varphi * Q_j + \varepsilon_1 \tag{3..5}$$

$$\ln N_{Douban} = \kappa_1 \ * \frac{1}{Q_j} + \kappa_2 Dummies_t + \varepsilon_2 \tag{3..6}$$

$$\ln \pi_j^{piracy} = C + \beta_1 \ln p - \beta_2 * \frac{1}{Q_j} + \beta_3 * \frac{\Delta t_j}{Q_j} + \beta_4 Dummies_t + \varepsilon_3 \tag{3..7}$$

Assuming that all the error terms satisfy strict exogeneity, I use the OLS method to estimate the coefficients of the functions above. The first equation is estimated separately; the second and the third are estimated jointly. I estimate both models simultaneously, while accounting for the correlated errors, which can lead to efficient estimates of the coefficients and standard errors. The coefficients are: $\hat{\varphi} = 5.36; \hat{\kappa_1} = 9; \hat{\beta_2} = \frac{\lambda c}{(1-\alpha)} = 3.70$; and $\hat{\beta_3} = 0.0007$. Then, I get $c = 19.8; \gamma = 0.004; \lambda = 0.13$ and $\alpha = 0.3$.

I assume that the market size was $M \times n$=20,000,000 in 2013, which is proportional to the overall number of screens[20]. By $\pi_j^{no\_piracy} = Mn\frac{Q_j}{\lambda}e^{-1-\lambda\frac{c-\gamma\overline{\Delta t}}{Q_j}}$, I predict the box office with piracy, without piracy and the piracy products. Table 20 presents the the results. The estimate of piracy loss is about 30 percent, on average, and if a pirated movie is available before release, the average loss will be 32 percent In general, I fit the mean and standard deviation of box office revenue and price well. The estimated model is able to fit basic patterns, when the assumed market size is 20,000,000. In Figure 3, however, the distribution of real box office revenue is flatter than predicted. Table 20 presents the sample fit for the box office and price. The averages of the predicted box office and price

[20] The number of screens was about 18,000 in 2013; 13,000 in 2012; 9,200 in 2011; 6,200 in 2010; 4,700 in 2009; 4,000 in 2008; 3,500 in 2007; 3,000 in 2006.

are a little lower than those in the real data, and both standard errors of predicted value are lower.

Linear Regression Using two reduced-form specifications, I evaluate box office losses. The first specification is based on three levels of box office factors. The main specification is:

$$\ln \pi_{ijt} = \beta_0 + \beta_1 pirate\_level_{ijt} + \alpha X_i + \gamma Y_j + \delta Z_t + \varepsilon_{ijt}, \tag{3..8}$$

where π_{ijt} is movie i's box office revenue in region j at time t; $pirate\_level_{ijt}$ denotes movie i's piracy level before and during the movie's showing period; X_i denotes i's attributes, such as film types, producer and distributor information, season, home country, movie length, film rating on Mtime and Douban, director level, super star number, and the number of searches on Baidu for the movie's name; Z_t is the time characteristics; Y_j is region characteristics or theater dummies.

Δt as the proxy for piracy level The results of specification 1 are presented in Table 21. I use TS lag as the first piracy proxy. The first column controls only for year dummies. Without controlling for other variables, the coefficient of the piracy level is about 0.0003 and is insignificant. After controlling for the Douban rate, film types, producer and distributor information, and season, the coefficient jumps to 0.0004. After controlling for the detailed rating information on Mtime, I find that the coefficient increases to about 0.0005. This change is due to the fact that Δt has a negative correlation with movie quality, which is shown on the 9th column of Table 27. Given the positive correlation between movie quality and Δt, the estimated loss is the lower bound of estimations in this article. This results suggests that if the Δt increases by ten days, then the box office will decrease by 0.5 percent. Given the huge range of the opening lag, with an opening lag over 1000 days for some movies, the coefficient of Δt is very small. I regress one dummy whether the Δt is smaller than 0 on the box office. I find that if the Δt is smaller than 0, then the box office loss could be 32.2 percent of the total revenue. In addition, I also use clear lag and opening lag as the proxies of Δt, the results of which are presented in columns 5 and 6 of Table 21. The coefficient means that if the clear resource comes out 100 days earlier, the box office revenue will be lower by 4.7 percent. If the opening lag is 100 days bigger, the box office revenue will be 6 percent lower.

As I control more variables, the coefficient of Δt becomes smaller, which suggests that the omitted variables cause a downward bias. Both Chinese movie consumers and the SARFT prefer foreign movies that flatter China. Such movies may have an easier time entering the quota list and passing censorship, which leads to both a short release delay and high box office revenue. Daily data tend to reduce the endogeneity problems, because they estimate the box office difference caused by piracy within the same movie. Columns 8 and 9 of Table 21 present daily box office results. The available TS dummy's coefficient is -0.729. The results mean that TS resources will decrease daily box office by 72.9 percent. I also add the interactions between the number of days and piracy lag and find that as the number of days increases, the piracy loss declines.

The Chinese government has made some moves towards anti-piracy laws in order to shut down piracy websites. From August to October 2007, for example, in response to pressure from the US government, the Chinese government shut down 339 websites, and handled 1,001 piracy cases, which is two times more than in previous two years. These impacts reminded me to use DDD to do the estimation. There are three dimensions in this model: mainland or not, government control period or not, and high piracy level or not. A total of seven dummies are needed to estimate in the DDD regression.

$$\begin{aligned} \ln \pi_{ijt} = {} & \beta_0 + \beta_1 D_c \times D_{ml} \times D_{\Delta t<0} + \beta_2 D_{ml} + \beta_3 D_{\Delta t<0} \\ & + \beta_4 D_c + \beta_5 D_c \times D_{ml} + \beta_5 D_c \times D_{\Delta t<0} \\ & + \beta_5 D_l \times D_{\Delta t<0} + \alpha X_i + \gamma Y_j + \delta Z_t + \varepsilon_{ijt}. \end{aligned} \tag{3..9}$$

In the equation, D_c is the control period dummy; D_{ml} is the treated dummy; and $D_{\Delta t<0}$ is the piracy level dummy. Here, I use the availability of the TS version before the theatrical release as the piracy level. X_i denotes movie i's attributes, and Y_j denotes region j's attributes. $D_c \times D_{ml} \times D_{\Delta t<0}$ captures the effect of the movement.

The control group is the Hong Kong market, which has an independent legal system and a government that is relatively independent from mainland China. According to an industry association's research, Hong Kong is the region with the lowest piracy level in the world.[21] The Hong Kong motion picture market was not affected by these impacts. I use three periods as the treatment periods, the month after December 7, 2009 (when the

[21] Hong Kong's video piracy level was only about 20 percent in 2004 (China's was 93 percent). Because the government of Hong Kong has taken steps towards fighting piracy, the piracy level in Hong Kong has continued to drop over the last ten years.?

Chinese government closed more than 500 BT websites); the 100 days after July 14, 2006 (when the Chinese government took a 100-day movement against piracy); and one month after January 23, 2011 (when the Chinese government closed the P2P file-sharing websites). One problem with this Hong Kong box office data is that the movie characteristics data include only the 467 movies that have been shown in the Hong Kong market, which is less than half of the movies released during this period.

Table 23 presents DDD specifications results.[22] I define the two treatment periods to determine whether these anti-piracy movements have short- and long-term effects. The first period is from the beginning of the anti-piracy efforts to six months later. The government efforts did have some effect in reducing box office loss caused by piracy in the six-month period. To check whether the anti-piracy policies have long-term effects, I extend the control period to one year. The coefficient of $D_c \times D_{ml} \times D_{\Delta t<0}$ becomes insignificant in the one-year period regression. The insignificant results of the one-year regression indicate that, in the long run, when new websites and new technologies replace the old ones the piracy supply return. Without legal actions prohibiting individuals' downloading behavior, Internet users can always find new websites or new technologies to replace those that have been banned. In the first specification, the estimation of Δt is downward-biased due to omitted variables. In the DDD regression, the correlation between unobserved quality and Δt is captured by the region difference and period difference, which solves the endogenous problem. The coefficients suggest that the policies will decrease the piracy loss caused by TS availability

[22] A detailed explanation of the DDD results is shown in the appendix.

before theatrical release by about 65 percent in the half-year range. The estimated piracy loss is bigger than the first specification and proves that the results of the first specification are downward-biased.

<u>"Name download" search amount on Baidu</u> Next, I use the number of “name download" searches on Baidu as a proxy to evaluate piracy intensity. Omitted variables also affect this estimation. Unobserved qualities of the pirated movies may increase both piracy intensity and the number of audience members. An instrument is needed to solve this problem. In this article, I use the number of searches for “movie name download" before the piracy resource is available as the instrument of the number of searches for “movie name download" after the piracy resource is available. The number of searches for “movie name download" before the piracy resource is available can be viewed as the willingness to pirate a movie, which correlates to piracy level, but does not affect box office directly. The high willingness before TS availability could be transmitted to the period after TS availability, which is supported by the 5th and 6th columns of Table 26. After controlling for the number of "movie name download" searches before and after opening, which are proxies for popularity level, this concern is no longer a threat to the instrument. Table 22 presents the result using the number of “movie name download" searches as the piracy proxy. The result shows that the search download attempt increased by one in the first four weeks of showing. This correlates to a 10 percent drop in box office revenue.

From this result, we know that only a small portion[23] of searching is done by marginal consumers choosing between the free movie and the big screen. The remaining searches did not affect box office revenue but were conducted by Internet users who always choose free movies over theater viewing. This conclusion requires some assumptions. First, the marginal consumers have the same ability to search as other Internet users. If they need more attempts than others to get the free movie links, this result is under-estimated. Nevertheless, even if there are differences in searching abilities, the differences in searching for free movies between the two kinds of Internet users are small because it is easy to find the pirated movie once it is available. The piracy link spreads rapidly on these free movie websites once the resources come out. Second, these results ignore the social network effects of piracy. More people get the free movie from peer-to-peer sharing.

3.7 Discussion

The Welfare Implications Watching movies for free online is the only choice for most Internet users are able to access movies in China. In the model only the n population can access theaters. They react to a short-term piracy level variation, and I find that n is very small in China. Screen number, theater location, and ticket price are the major factors determining

[23] Baidu has 80 percent of the market share. There are 600,000,000 Internet users in China. The average number of searches is about 30,000, which means that for every 10,000,000 Baidu users, there are 30,000 attempts to search the free movies online. So, the average number of searches is about 30,000 times 50 (the market share times Internet users number, then divided by 10,000,000). The average number of searches # is therefore 1,500,000. The average box office revenue is about 50,000,000, which means that the average number of audience members is 1,000,000 given that the average price is about 50. This means that if the number of searches increases by 1,500,000, then the number of theater goers decreases by 100,000. Put another way, only about 100,000 people among the 1,500,000 Internet users are theater viewers.

the availability of theatrically released films. Before 2009[24], most people in China (especially the ones in small cities and rural areas) did not have access to movies in theaters. Even today, very few rural areas have theaters. The number of screens showing films in China has increased from about 1,000 in 2002 to 20,000 in 2014. At the same time, the annual box office increased from 0.95 billion yuan in 2002 to 21.7 billion yuan in 2014. The elasticity of box office revenue over the number of screens is 0.468, which suggests that about half of the movie's box office growth could be attributed to the growth in the number of screens. In addition, the ticket price is too high in comparision with other nations. The average price is 37 yuan (China Mainland Motion Picture Market report, 2013) which accounts for about 4 percent of the average monthly income of the urban population. This high price deters people from watching movies in the theaters.

In the model, even in the n population, piracy resources will increase the movie's overall number of audience members. Quotas and censorship limit the number of films that can be shown in theaters, by which n equals 0. Even if the movies cannot be shown on theater screens, Chinese Internet users can still find them online. To some extent taking the absence of legal DVD sales into consideration, policies such as censorship and quotas lead to a situation in which unpaid online movies are the only option Chinese Internet users have for watching certain movies. Given this circumstance, piracy provides consumers with more choices, in spite of the government's attempts at control. In China, there are incredibly long lists of banned movies. But all of them can be found on China's Internet. In this market,

[24] The Chinese government began to subsidize the building of theater chains in small cities in 2009.

which the government controls so tightly, piracy allows more people to watch movies and diversifies people's choices, which leads to welfare improvement. This model, however, has not considered the long-run supply change caused by piracy. In the long run, it may decrease producers' incentive to supply movies. As Zhang (2012) mentions, China's entire population may benefit from piracy; however, people attribute the relative lack of creative productivity in China to piracy. The overall welfare influence is thus still uncertain, as the previous literature concludes(Hinnosaar, 2002).

The Policy Implications Quota and censorship also constrain the motion picture market in China. The quantity and quality of the movies in the market are influenced by the government's preferences. The 9th column of Table 27 suggests that Δt is determined mainly by the opening lag, which is caused by censorship and protectionism. The elimination of "unsuitable content" through censorship increases the opening lag by 69 days, on average, according to the estimation in the 7th column of Table 30. Most movies with an opening lag are imported movies, which have higher average quality compared to domestic movies. Because of the quotas, most foreign movies are buyout movies,[25] which have very long opening delays. Good-quality foreign movies therefore usually have a long opening lag, which could indicate a positive relation between opening lag and the Douban rate quality indicator. The positive correlation between opening lag and movie quality also suggests that the main results of Δt are the low boundary of piracy loss. Government's protectionism

[25] Aside from the time loss caused by censorship, distributors also spend some time monitoring for movies' box office performances in foreign markets in order to determine which movies to purchase.

and censorship indirectly cause greater piracy loss.

Theaters' Reaction to Piracy By adding price and screen allocation into the daily regression, the coefficient of TS availability decreases from -0.729 to -0.172. Theaters drop the price and the number of screens[26] allocated to the movies when piracy resources are available. The difference between the two coefficients means that some of the loss of the box office sales is caused by theaters themselves because they allocate fewer screens to the movie when they know that the pirated movie is available online. Theaters' reactions to the demand change caused by piracy also contributes to the decrease in movies' box office revenue. The 13th and 14th columns of Table 27 present the theaters' reaction to piracy. As mentioned above, theaters decrease allocated screens and the price of the movies being pirated. This result may answer the question as to why most movies in China have very short opening periods compared with a market less vulnerable to piracy. As time passes, the movies being pirated more are less likely to entice consumers to spend their money at the theatre. In the first column of Table 27, I find that a high piracy level decreases a movie's opening week numbers. The short showing period prevents consumers who are insensitive to the latest movie news from watching these movies in theaters. Theaters' reactions caused by the limitation of the numbers of screens and market structure magnify piracy loss. Reacting quickly to change in demand and supply increase theaters' revenues and decrease the movies' overall box office revenues.

[26] For most theaters in China, people can book the ticket one or two days in advance.

Full Substitution in the Short Term In the second and third columns of Table 27, I regress the piracy level on the number of people rating the movie on Douban and Mtime, which are proxies for the overall number of audience members. I find that the piracy level does not affect the rating number on the two websites. The results indicate that the overall number of audience members is determined solely by movie quality and is not affected by piracy level. Internet users will always choose to watch a good movie, by whatever means are available. The way they choose to watch the movie, especially for the marginal consumers, depends mainly on the piracy resources available. The early free movies online pull just the marginal consumers out of the theater. In other words, Internet users will always switch to the free version online, especially if the movies are off of the big screen or not shown in theaters. Δt has no effects on the overall number of consumers, which is consistent with the model's conclusion.

3.8 Conclusion

This article has attempted to estimate the substitution effects of pirated movies- the free online movies available on peer-to-peer file-sharing websites- on theater movies. Using newly-constructed data collected from piracy websites in China as well as search engine data, I constructed two proxies: the time lag between piracy supply and legal supply, and the number of searches on the search engine Baidu, to evaluate the piracy level. I found a significant causal effect of piracy level on box office revenue in China: The results of the structural estimation show that the average piracy loss is 30 percent in this market and that pirated movies that become available before theatrical release reduce a movie's box office by 32 percent. OLS results imply that pirating before a day decreases the daily revenue by about 70 percent on average. The DID results imply that the anti-piracy actions that shut down the major websites could increase the box office of movies released in the period by about 70 percent, but the effects disappear in the long run. Protectionist and censorship policies cause major release delays in this market, which cause an earlier piracy supply relative to theater supply and lead to a higher piracy level.

The second aim of this article was to reveal the mechanism by which piracy activities affect box office revenue. Box office losses caused by piracy are substantial. Substution elasticity, however, is small. My instrumental variable results on the number of movie searches on Baidu, based on a plausibly exogenous instrument, find a small effects, and only a small portion of the free movie search attempts are made by marginal consumers who are choosing between watching big-screen movies and watching free movies. The rest of the piracy attempts are committed by the Internet users who cannot access theaters at all. The absence of legal products makes piracy products the sole choice for most of China's population. Besides the first-order loss caused by demand decreasing, theaters also adjust the price and screen allocation to the demand change; this behavior accounts for three quarters of the box office change.

A limitation of the present study is its focus, due to data constraints on the long-run supply change caused by piracy. Without considering the long-term supply effects, pirated movies increase social welfare. The government's protectionism and censorship policies reduce the quality and quantity of the supply of movies to theaters and indirectly cause piracy losses. In the static model, the producers' piracy losses will be transferred to consumers, thus affecting their welfare. Taking low theater coverage rates and policy restrictions into consideration, just in the short term, piracy diversifies consumers' choices, allows more people to watch movies, and improves overall welfare. Piracy may reduce social welfare, however, by reducing the producers' incentive to produce movies. Evaluating piracy's long-term welfare effects is a promising direction for future research.

REFERENCE

Antonovics, Kate, and Ben Backes. "The effect of banning affirmative action on human capital accumulation prior to college entry." IZA Journal of Labor Economics 3.1 (2014): 5.

Altonji, Joseph G., Erica Blom, and Costas Meghir. Heterogeneity in human capital investments: High school curriculum, college major, and careers. No. w17985. National Bureau of Economic Research, 2012.

Ahn, I., and Yoon, K. (2009). On the impact of digital music distribution. CESifo Economic Studies, 55(2), 306-325.

Baidu. (n.d.). In Wikipedia. Retrieved November 10, 2013, from https://en.wikipedia.org/wiki/Baidu

Battistin, E., De Nadai, M., and Padula, M. (2015). Roadblocks on the Road to Grandma's House: Fertility Consequences of Delayed Retirement.

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

Bauernschuster, S., and Schlotter, M. (2015). Public child care and mothers' labor supply-Evidence from two quasi-experiments. Journal of Public Economics, 123, 1-16.

Bae, S. H., and Choi, J. P. (2006). A model of piracy. Information Economics and Policy, 18(3), 303-320.

Belleflamme, P. (2003). Oligopolistic pricing of piratable information goods (No. ECON Discussion Papers (2003/87)). UCL.

Belleflamme, P., and Peitz, M. (2014). Digital piracy: an update (No. CORE Discussion

Paper (2014/19)). UCL.

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan. ”Affirmative action in education: Evidence from engineering college admissions in India.” Journal of Public Economics 94.1 (2010): 16-29.

Becker, G.S. (1974). A theory of social interactions. Journal of Political Economy 82, 1063-1093.

Blau, F. D., and Kahn, L. M. (2005). Changes in the labor supply behavior of married women: 1980-2000 (No. w11230). National Bureau of Economic Research.

Bonsang, E. (2007). How do Middle-Aged Children Allocate Time and Money Transfers to their Older Parents in Europe? Empirica, 34, 171--188.

Bodoh-Creed, Aaron, and Brent R. Hickman. 2015. ”Using Auction Theory to Study Human Capital Investment in Assortative Matching Markets: a Look at Affirmative Action in College Admissions.” working paper, University of Chicago.

Bodoh-Creed A, Hickman B. College Assignment as a Large Contest[R]. Mimeo, 2015.

Chen, F., and Liu, G. (2009). Population aging in China. In International handbook of population aging (pp. 157-172). Springer Netherlands.

Chen, F., Liu, G., and Mair, C. A. (2011). Intergenerational ties in context: Grandparents caring for grandchildren in China. Social Forces, sor012.

Chen, Y., Jin, G. Z., and Yue, Y. (2010). Peer migration in China (No. w15671). National Bureau of Economic Research.

China Hollywood Society. (2011). About Co-productions. Retrieved from http://www.chinahollywood.org/about-co-productions

Chinafilmbiz. (2012) How China's Movie Distribution System Works. Retrieved from https://chinafilmbiz.com/2012/11/07/

Chinawhisper. (2012). Top 15 most popular Chinese video websites, Retrieved from http://www.chinawhisper.com/top-15-most-popular-chinese-video-websites

Cigno, A. (2006). A constitutional theory of the family. Journal of Population Economics, 19(2), 259-283.

Coate, Stephen, and Glenn Loury. "Antidiscrimination enforcement and the problem of patronization." The American Economic Review 83.2 (1993): 92-98.

Cong, Z., and Silverstein, M. (2012). Caring for grandchildren and intergenerational support in rural China: a gendered extended family perspective. Ageing and Society, 32(03), 425-450.

Compton, J., and Pollak, R. A. (2014). Family proximity, child care, and women's labor force attachment. Journal of Urban Economics, 79, 72-90.

Cotton, Christopher, Brent R. Hickman, and Joseph P. Price. Affirmative action and human capital investment: Theory and evidence from a randomized field experiment. No. 1350. Queen's Economics Department Working Paper, 2015.

Douban. (n.d.). In Wikipedia. Retrieved November 1, 2013, from https://en.wikipedia.org/wiki/Douban

Estevan, Fernanda, Thomas Gall, and Louis-Philippe Morin. Redistribution without distortion: Evidence from an affirmative action program at a large Brazilian university. No. 2016˙07. University of San Paulo (FEA-USP), 2016.

Feng, Z., Liu, C., Guan, X., and Mor, V. (2012). China's rapidly aging population creates policy challenges in shaping a viable long-term care system. Health Affairs, 31(12), 2764-

2773.

Ferman, Bruno, and Juliano Assuncao. 2011. "Does Affirmative Action Enhance or UndercutInvestment Incentives? Evidence from Quotas in Brazilian Public Universities." Typescript, Massachusetts Institute of Technology Department of Economics.

Fu, Qiang. "A theory of affirmative action in college admissions." Economic Inquiry 44.3 (2006): 420-428.

Furstenberg, Eric. Affirmative Action, Incentives, and the Black-White Test Score Gap. Working paper, College of William and Mary, 2003.

Geurts, T., van Tilburg, T., Poortman, A. R., and Dysktra, P. A. (2015). Child care by grandparents: Changes between 1992 and 2006. Ageing and Society, 35(06), 1318-1334.

Gu, Jiafeng. "Spatial recruiting competition in Chinese higher education system." Higher Education 63.2 (2012): 165-185.

Gopal, R. D., Bhattacharjee, S., and Sanders, G. L. (2006). Do Artists Benefit from Online Music Sharing?*. The Journal of business, 79(3), 1503-1533.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn. Review of Economics and Statistics, 91: 437--456.

Krishna, Kala, Veronica Frisancho Robles, and Cemile Yavas. "Can Disadvantaged Students Catch Up? Disentangling Selection and Learning Effects." (2011).

Hamaaki, J., Hori, M., and Murata, K. (2014). Intergenerational Transfers and Asset Inequality in Japan: Empirical Evidence from New Survey Data. Asian Economic Journal, 28(1), 41-62.

Hank, K., and Buber, I. (2009). Grandparents caring for their grandchildren findings from the 2004 survey of health, ageing, and retirement in Europe.Journal of Family Issues, 30(1), 53-73.

Ho, C. (2015). Grandchild care, intergenerational transfers, and grandparents' labor supply. Review of Economics of the Household, 13(2), 359-384.

Hong Kong Trade Development Council. (2008). SARFT Reiterates Film Censor Criteria. Retrieved from http://info.hktdc.com/alert/cba-e0804c-2.htm

Hong Kong Trade Development Council. (2016). Film Entertainment Industry in Hong Kong. Retrieved from http://hong-kong-economy-research.hktdc.com/business-news/article/Hong-Kong-Industry-Profiles

Hsieh, H. L., Chou, S. Y., Liu, E., and Lien, H. M. (2015). Strengthening or Weakening? The Impact of Universal Health Insurance on Intergenerational Coresidence in Taiwan. Demography, 52(3), 883-904.

Hickman, Brent R. 2015. "Human Capital Investment and Affirmative Action: A Structural Policy Analysis of US College Admissions." Typescript, University of Chicago Department of Economics.

Hickman, Brent R. "Human Capital Investment and Affirmative Action: A Structural Policy Analysis of US College Admissions." Typescript, University of Chicago Department of Economics (2013).

Li, Hongbin, et al. "Does attending elite colleges pay in China?." Journal of Comparative Economics 40.1 (2012): 78-88.

Jellal, M., and Wolff, F. C. (2005). Free entry under uncertainty. Journal of Economics, 85(1), 39-63.

Johnson, R. W., and Sasso, A. T. L. (2006). The impact of elder care on women's labor supply. The Journal of Health Care Organization, Provision, and Financing, 43(3), 195-

210.

Johnson, N. E. (1985). Varieties of representation in eliciting and representing knowledge for IKBS. International Journal in Systems Research and Information Science, 1(2), 69-90.

Judd, K. L. (1998). Numerical methods in economics. MIT press.

Lambrick, J. (2009). Piracy, File Sharing... and Legal Fig Leaves. J. Int'l Com. L. and Tech., 4, 185.

Lee Tim. (2006). Texas-Size Sophistry. Retrieved from Technology Liberation Front website: https://techliberation.com/2006/10/01/texas-size-sophistry/

Liebowitz, S. J. (2008). Research Note-Testing File Sharing's Impact on Music Album Sales in Cities. Management Science, 54(4), 852-859.

Ma, L., Montgomery, A., Singh, P., and Smith, M. D. (2011). Pre-Release Movie Piracy and Box Office Sales: Estimates and Policy Implications. Unpublished working paper.

Keane, M. P., and Wolpin, K. I. (1997). The career decisions of young men.Journal of political Economy, 105(3), 473-522.

Kocherlakota, N. R. (1996). Implications of efficient risk sharing without commitment. The Review of Economic Studies, 63(4), 595-609.

Lockwood, L. M. (2014). Incidental bequests: Bequest motives and the choice to self-insure late-life risks (No. w20745). National Bureau of Economic Research.

Luo, Y., LaPierre, T. A., Hughes, M. E., and Waite, L. J. (2012). Grandparents providing care to grandchildren a population-based study of continuity and change. Journal of Family Issues, 33(9), 1143-1167.

McDonald, K. B., and Armstrong, E. M. (2001). De-Romanticizing Black Intergenerational

Support: The Questionable Expectations of Welfare Reform. Journal of Marriage and Family, 63(1), 213-223.

McGarry, K. (2000). Testing parental altruism: Implications of a dynamic model (No. w7593). National Bureau of Economic Research.

McCutchan. (2013). S. Government Allocation of Import Quota Slots to US Films in China's Cinematic Movie Market. Unpublished working paper.

Melenberg, B., and Zheng, J. (2012). Health Expectancy of the Chinese Elderly: Current Trends and Future Projection. Tilburg University.

Meng, Xin, Kailing Shen, and Sen Xue. "Economic reform, education expansion, and earnings inequality for urban males in China, 1988--2009." Journal of Comparative Economics 41.1 (2013): 227-244.

Milot, M. R. (2014). Testing the Lost Sale Concept in the Context of Unauthorized BitTorrent Downloads of CAM Copies of Theatrical Releases. Available at SSRN 2502931.

Moro, Andrea, and Peter Norman. "Affirmative action in a competitive economy." Journal of Public Economics 87.3 (2003): 567-594.

Motion Picture Association of America. (2010). 2010 Theatrical Market Statistics. Retrieved from https://wikileaks.org/sony/docs/03˙03

National College Entrance Examination, Wikipedia: The Free Encyclopedia. Wikimedia Foundation, Inc. 22 July 2004. Web. 10 Aug. 2016.

National Survey Research Center (2014). China Longitudinal Aging Social Survey. Beijing, China.

National Bureau of Statistics. (2016). National data. [Data file]. Retrieved from: http://data.stats.gov.cn/easyquery.htm?cn=C01

Novos, I. E., and Waldman, M. (1984). The effects of increased copyright protection: An analytic approach. The Journal of Political Economy, 236-246.

Oberholzer-Gee, F., and Strumpf, K. (2007). The effect of file sharing on record sales: An empirical analysis. Journal of political economy, 115(1), 1-42.

Peitz, M., and Waelbroeck, P. (2004). The effect of internet piracy on CD sales: Cross-section evidence. Unpublished working paper.

Peitz, M., and Waelbroeck, P. (2006). Piracy of digital products: A critical review of the theoretical literature. Information Economics and Policy, 18(4), 449-476.

Pirate Bay. (n.d.). In Wikipedia, Retrieved September 29, 2013. from https://en.wikipedia.org/wiki/The˙Pirate˙Ba.

Pirated movie release types. (n.d.). In Wikipedia. Retrieved November 3, 2013, from https://en.wikipedia.org/wiki/Pirated˙movie˙release˙types

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.

Posadas, J., and Vidal-Fernandez, M. (2012). Grandparent's child care and female labor force participation. IZA discussion paper no. 6398.

Priest, E. (2006). The future of music and film piracy in China. Berkeley Technology Law Journal, 795-871.

Schanzenbach, M. M., and Sitkoff, R. H. (2008). Bequest Puzzles in Economics: A Legal Perspective. In American Law and Economics Association Annual Meetings (p. 96).

Rob, R., and Waldfogel, J. (2004). Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students (No. w10874). National Bureau of Economic Research.

Rob, R., and Waldfogel, J. (2007). Piracy on the silver screen*. The Journal of Industrial Economics, 55(3), 379-395.

Shi, Li, and Chunbing Xing. ”China’s higher education expansion and its labor market consequences.” (2010).

Silverstein, M., Cong, Z., and Li, S. (2006). Intergenerational transfers and living arrangements of older people in rural China: Consequences for psychological well-being. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61(5), S256-S266.

Shi R, Xu H, Zhang B, and Yao J. (2015). Slowly, China Prepares to Raise Retirement Age. Retrieved from: http://english.caixin.com/100796734.html

Spear, S. E., and Srivastava, S. (1987). On repeated moral hazard with discounting. The Review of Economic Studies, 54(4), 599-617.

Sugawara, S., and Nakamura, J. (2014). Can formal elderly care stimulate female labor supply? The Japanese experience. Journal of the Japanese and International Economies, 34, 98-115.

Stark, O.,and Zhang, J. (2002). Counter-compensatory inter vivos transfers and parental altruism: Compatibility or orthogonality. Journal of Economic Behavior and Organization 47, 19--25.

State Administration of Radio, Film, and Television. (n.d.). In Wikipedia. Retrieved September 26, 2013. from https://en.wikipedia.org/wiki/SARFT

Takeyama, L. N. (1997). The Intertemporal Consequences of Unauthorized Reproduction of Intellectual Property 1. The Journal of Law and Economics,40(2), 511-522.

The Software Alliance. (2011). BSA Global Software Piracy Study 2011, Retrieved from http://globalstudy.bsa.org/2011/#

The World Bank. (2016). Real Interets Rate [Data file]. Retrieved from: http://data.worldbank.org/indicator/FR.INR.RINR

The World Bank. (2015). Fertility rate. [Data file]. Retrieved from: http://data.worldbank.org/indicator/SP.DYN.TFRT.IN

Trinh, T. (2006). China's pension system: Caught between mounting legacies and unfavourable demographics. Deutsche Bank research, February, 17

Wang, D. (2006). China's urban and rural old age security system: Challenges and options. China and World Economy, 14(1), 102-116.

Wang, Houxiong. "Research on the influence of college entrance examination policies on the fairness of higher education admissions opportunities in China." Chinese Education and Society 43.6 (2010): 15-35.

Wan, Yinmei. "Expansion of Chinese higher education since 1998: Its causes and outcomes." Asia Pacific Education Review 7.1 (2006): 19-32.

Yeung, Wei-Jun Jean. "Higher education expansion and social stratification in China." Chinese Sociological Review 45.4 (2013): 54-80.

Yoon, K. (2007). On the impact of digital music distribution. Information Economics and Policy, 18, 374-384.

Yu, S. (2007). China's summer blackout begins following Pirates release. Screen Daily. Retrieved from http://www.screendaily.com/chinas-summer-blackout-begins-following-pirates-release/4033106.article

Zhang, Yu. "History and Future of the National College Entrance Exam (NCEE) in China." National College Entrance Exam in China. Springer Singapore, 2016. 1-15.

Zou, Yihuan. Quality of Higher Education: Organizational and Educational Perspectives. Vol. 3. River Publishers, 2013.

Yu, Y. and Liu, S. (2010). Holding up half the sky? Are Chinese women given equal rights in political participation?, in IDAS Symposium: The rising Asia Pacific Region: Opportunities and challenges for cooperation, pp. 300-315.

Zentner, A. (2005). File sharing and international sales of copyrighted music: An empirical analysis with a panel of countries. The BE Journal of Economic Analysis and Policy, 5(1).

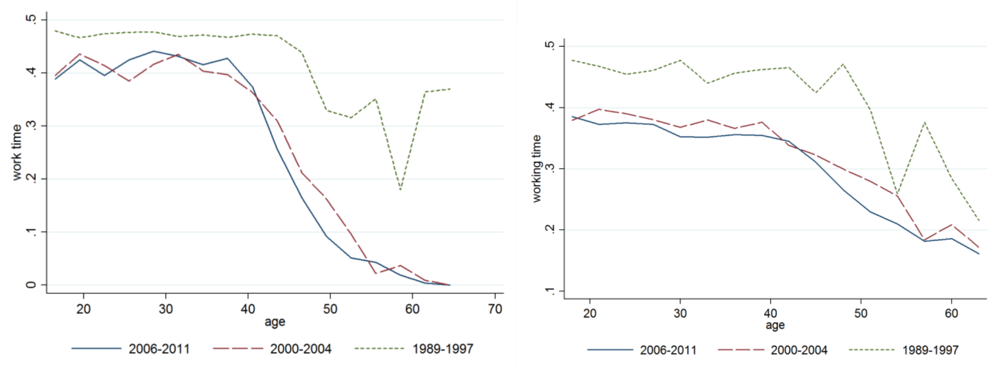
Zentner, A. (2006). Measuring the effect of file sharing on music purchases*.Journal of Law and Economics, 49(1), 63-90.

Zhang, Y., and Maclean, M. (2012). Rolling back of the state in child care? Evidence from urban China. International Journal of Sociology and Social Policy, 32(11/12), 664-681.

APPENDIX A

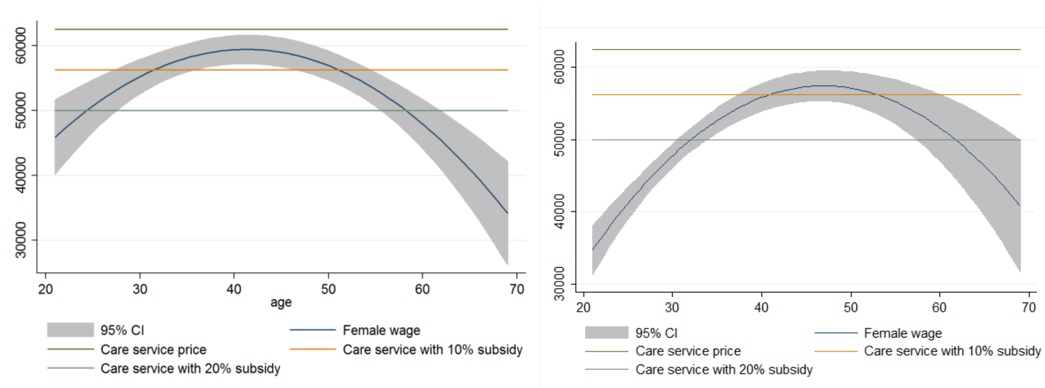
FIGURES AND TABLES

Figure 1 Female working time by age



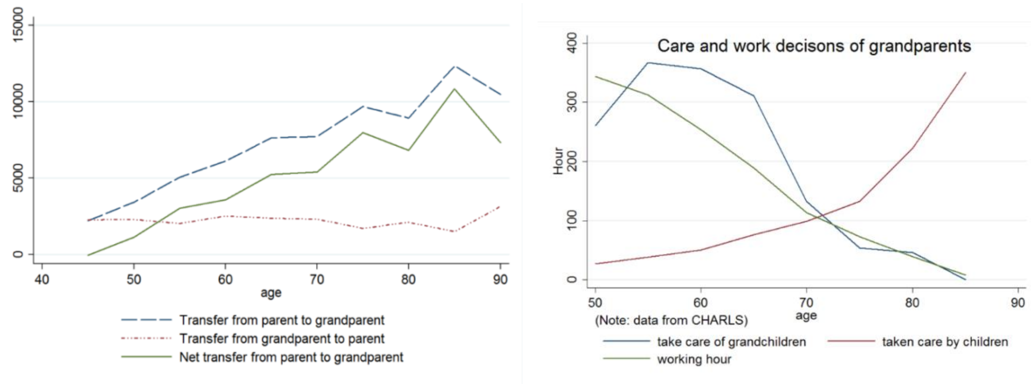
Note: Data from CHNS. The first graph is the female working time of urban area. The second one is the female working time of rural area. In the figures above, I normalize the overall time equal to 1. Assume each individual can have 14 hours to work at most each day. I define working time = (annual work month/12) × (daily work hour/14) × (weekly work day/7).

Figure 2 Wage structure by age cohort in simulation



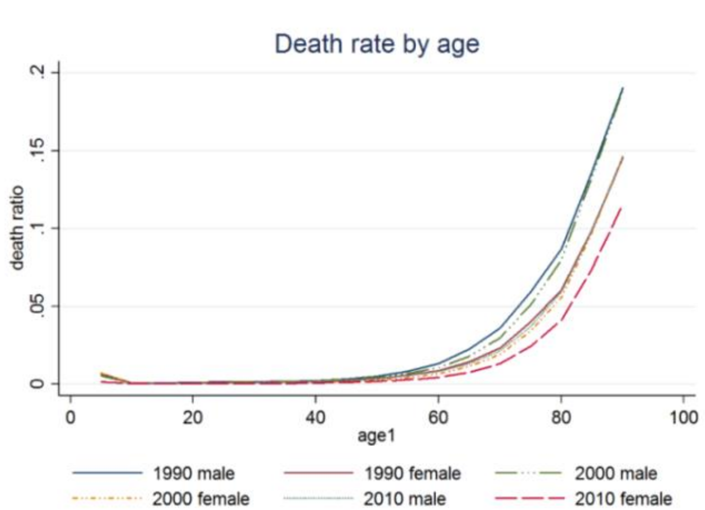
Note: The money unit is Yuan. The first graph represents the benchmark wage structure. I get the wage structure from CHNS data 2006-2011. The average wage rates of two households are always smaller than the care service. The second graph is the wage structure 2. In the wage setting, the wage growth rate by age is from the wage structure of 2000-2004.

Figure 3 Average income transfer and support by age



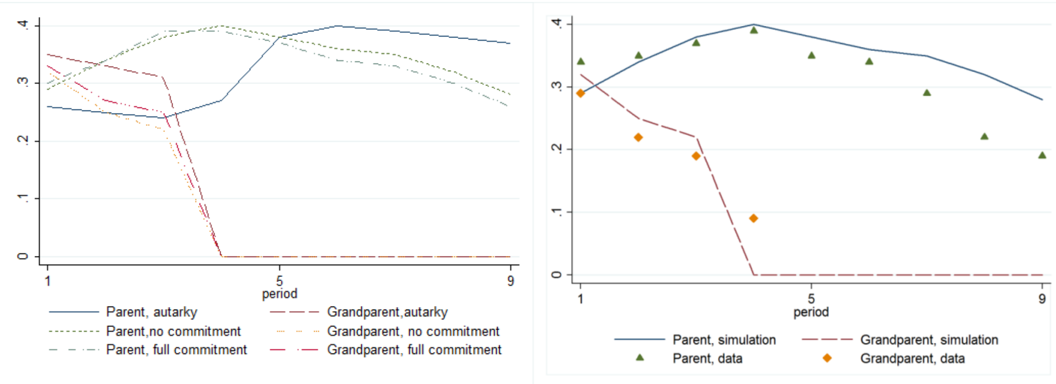
Note: Data from CHARLS 2008-2013. The first graph is the money transfer decisions. The second graph is the support decisions. The money unit is Yuan. Age is grandparents' age. Transfer is the sum value of the gift, regular monetary in-kind support, and non-regular monetary in-kind support. The care hour is the average hour to take care of grandchildren or taken care by children per year. The working hour is the average working hour of grandparents.

Figure 4 Average death rate by age



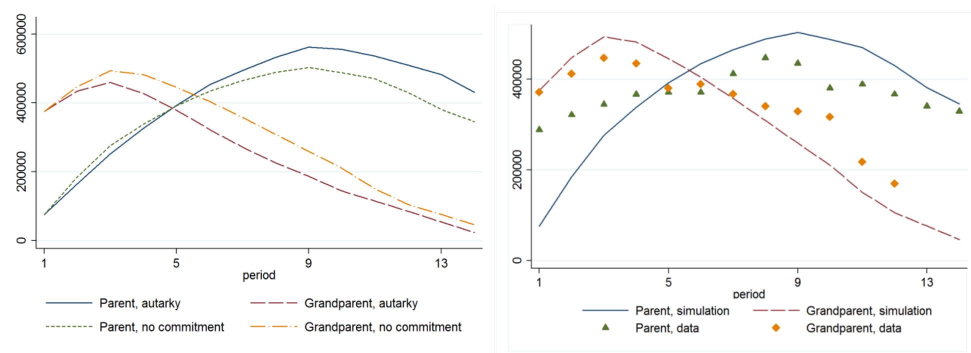
Note: The death rates are from the National Population and Reproductive Health Science Data Center of China. The value is the average death rate of the age cohort.

Figure 5 Simulation results: working hour by age



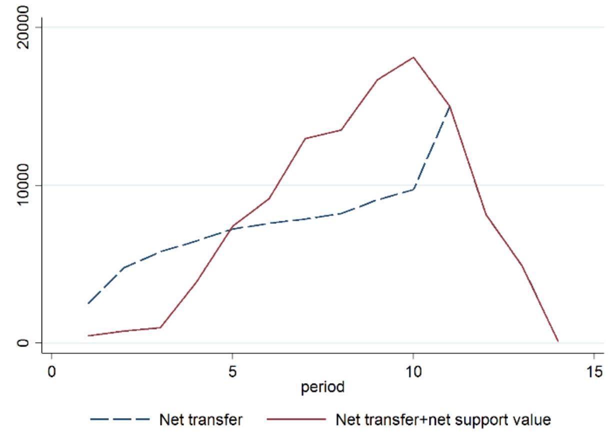
Note: Data from CHNS. Results of the first graph are from simulation. The parameters are following benchmark setting: care service subsidies are 0; inheritance tax is 0; mandatary retirement age is 10; the wage rate is given by wage structure 1 (shown in the first graphs of Figure 1); 42% are type 1 and 58% of households are type 2; grandparents' initial saving is 375,000; and I normalize the overall time to 1. The second graph compares the labor supply by age from simulation and data. The simulation result is using a contract with benchmark setting.

Figure 6 Simulation results: Average asset level by age



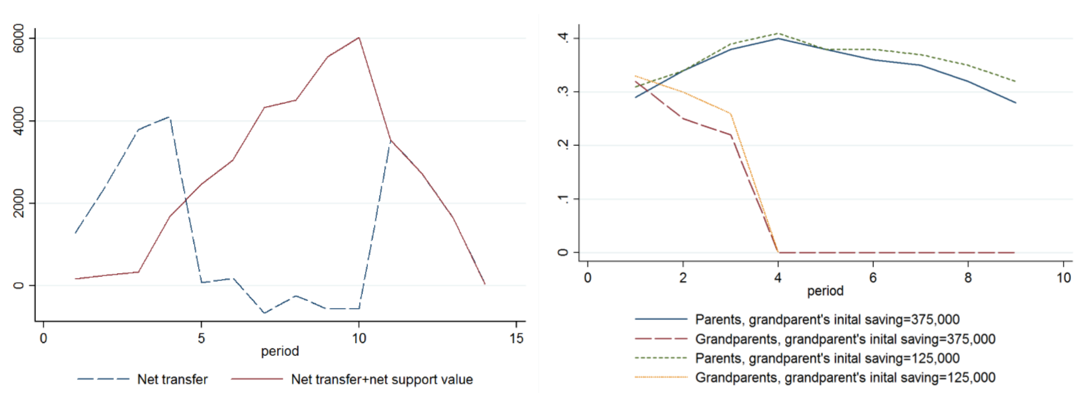
Note: Money unit is Yuan. The parameters are following benchmark setting: care service subsidies are 0; inheritance tax is 0; mandatary retirement age is 10; the wage rate is given by wage structure 1 (shown in the first graphs of Figure 1); 42% are type 1 and 58% of households are type 2; grandparents' initial saving is 375,000; and I normalize the overall time to 1. The first graph is the simulation results of the asset level by age. The second graph compares the asset level by age from simulation and data. The data is from CHNS 2006-2011. The simulation result is using a contract with benchmark setting.

Figure 7 Simulation results: Net transfer from parent to grandparent



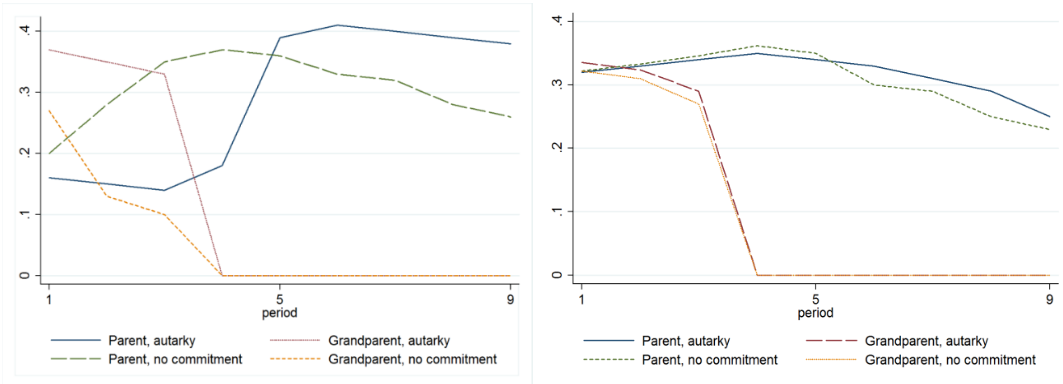
Note: The money unit is Yuan. Transfer and care support value is defined by net transfer plus labor support value. Labor support value is defined by the support hour time support opportunity cost. Money unit is Yuan. The parameters are following benchmark setting: care service subsidies are 0; inheritance tax is 0; mandatary retirement age is 10; the wage rates are given by wage structure 1 (shown in the first graphs of Figure 1); 42% are type 1 and 58% of households are type 2; grandparents' initial saving is 375,000; and I normalize the overall time to 1.

Figure 8 Experiment: grandparent initial saving is 125,000



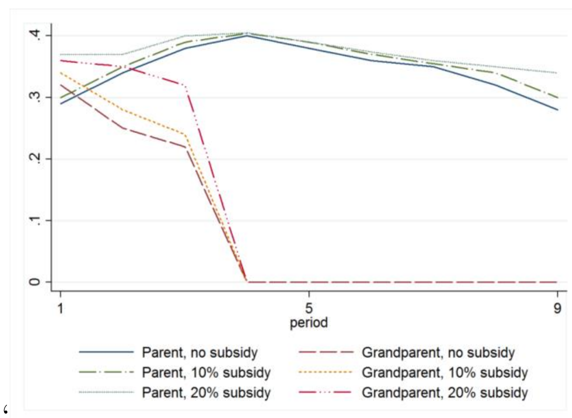
Note: The money unit is yuan. In this part, I have changed grandparents' initial savings from 375,000 in benchmark to 125,000. The first graph is transfer value information. The second graph is the labor supply information. Except grandparents' initial savings, the rest of parameters is following benchmark settings.

Figure 9 Simulation results: Labor supply of different types of households



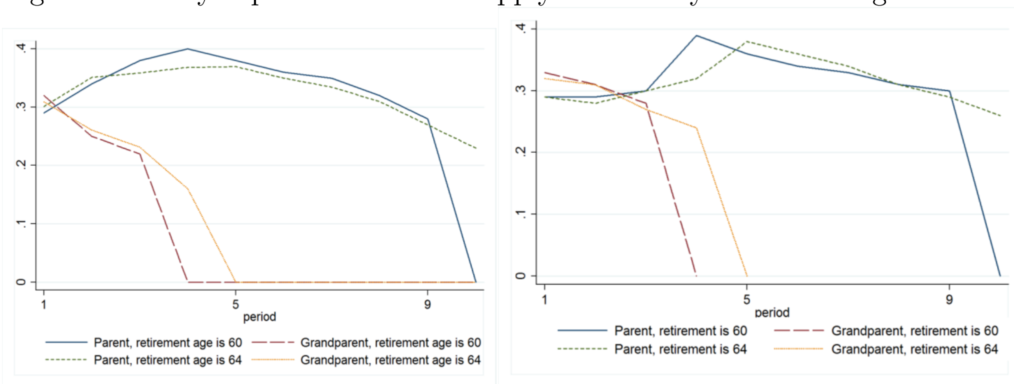
Note: I compare the labor supply of the households with different types of preference on child care and elder care. In Graph 1, both households are type1. In Graph 2, both households are type2. The money unit is Yuan. In the contract, 33.64% of the pair of parents and grandparents is both type 1 preference; and 17.64% are both type 2 preferences. Except the preference setting, the rest of parameters is following benchmark settings.

Figure 10 Policy experiment: Child/elder care subsidy and labor supply



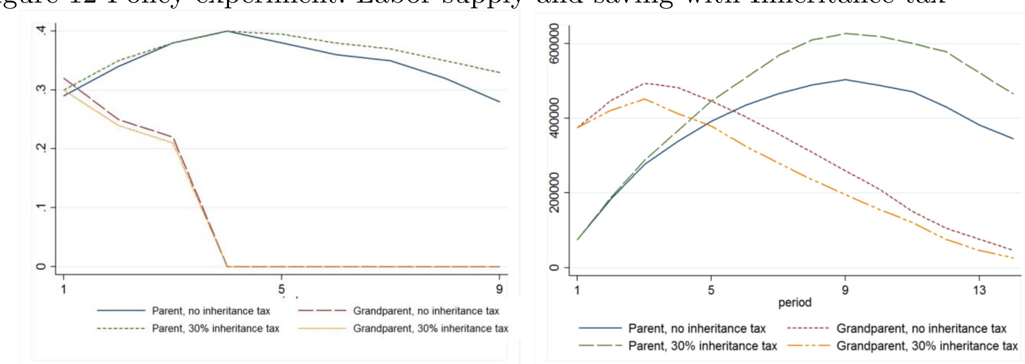
Note: Results from simulation. Both households are in a contract. There are care service subsidies on both child care and elder care service. Except for the car service subsidies setting, the rest of parameters is following benchmark settings.

Figure 11 Policy experiment: Labor supply after Delay retirement age



Note: Results from simulation. Both households are in a contract. In this part, I move the mandatory retirement age from 9 in benchmark to 10. Both households retire later for one period. In the first graph, except for the setting of mandatory retirement age, the rest of parameters is following benchmark settings. In the second graph, the wage rates are given by wage structure 2. Except for the setting of mandatory retirement age and wage rates, the rest of parameters is following benchmark settings.

Figure 12 Policy experiment: Labor supply and saving with Inheritance tax



Note: Results from simulation. Both households are in a contract. In this part, I move the mandatory retirement age from 9 in benchmark to 10. Both households retire later for one period. Except for the setting of mandatory retirement age, the rest of parameters is following benchmark settings. The first graph shows the result of the labor supply. The second graph shows the result of the savings.

Figure 13 timeline of NCEE

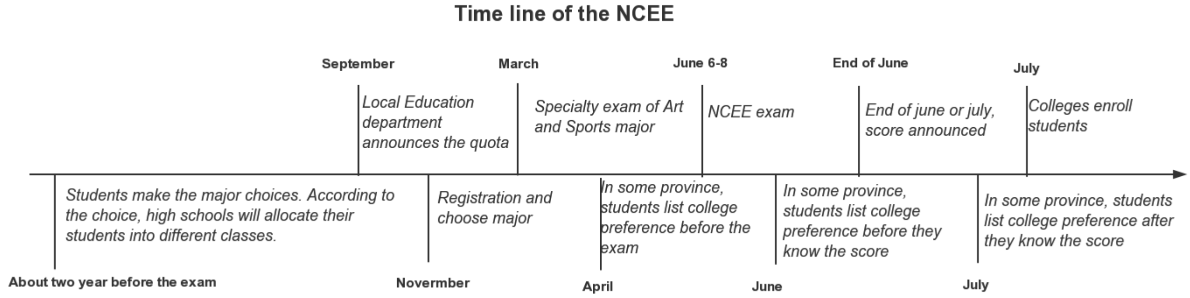
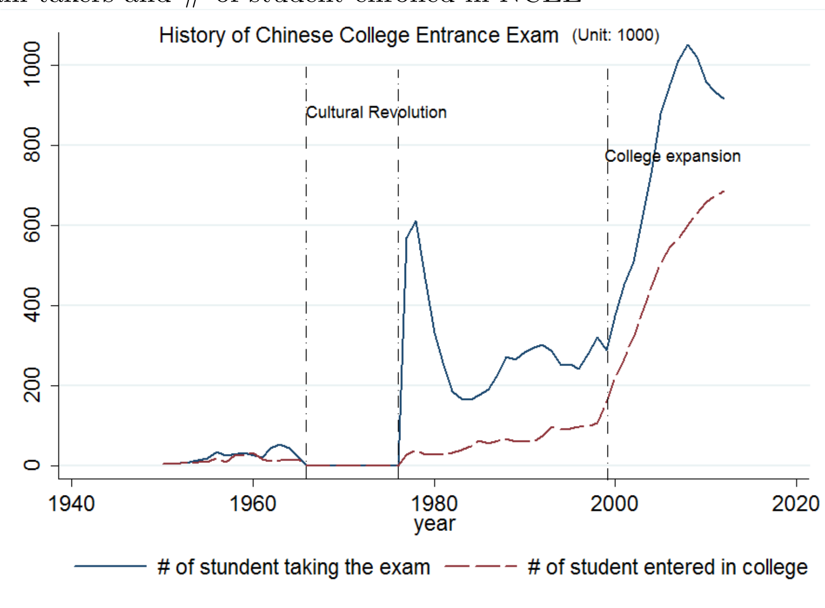
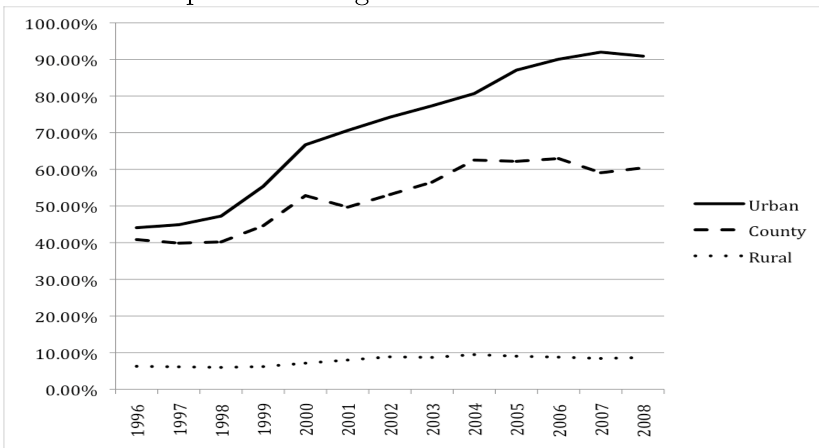


Figure 14 # of exam takers and # of student enrolled in NCEE



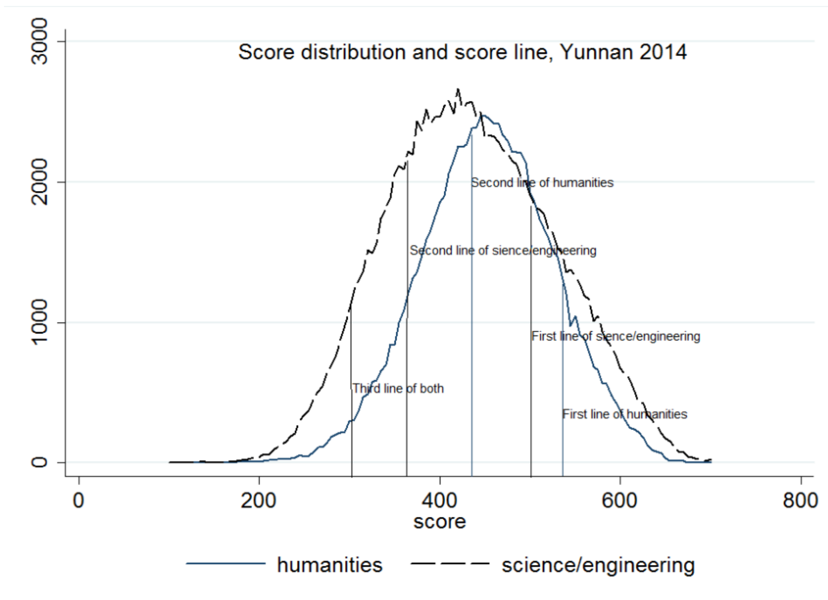
Note: Data from Chinese Educational Testing Yearbook 2010

Figure 15 Urban-Rural Disparities in High School Promotion Rate



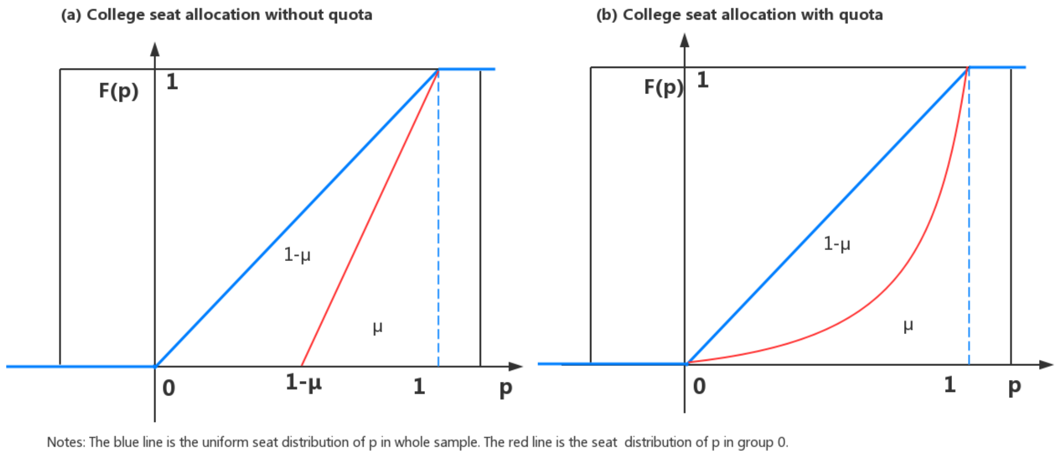
Note: data from china national statistical bureau

Figure 16 Score distribution of Yunnan in 2014



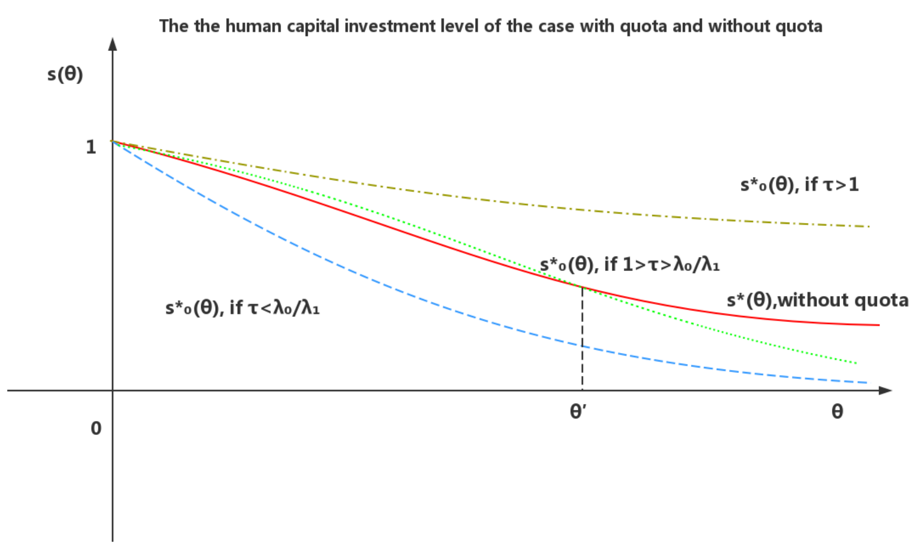
Note: Data from Yunnan Provincial Department of Education

Figure 17 College seat allocation example



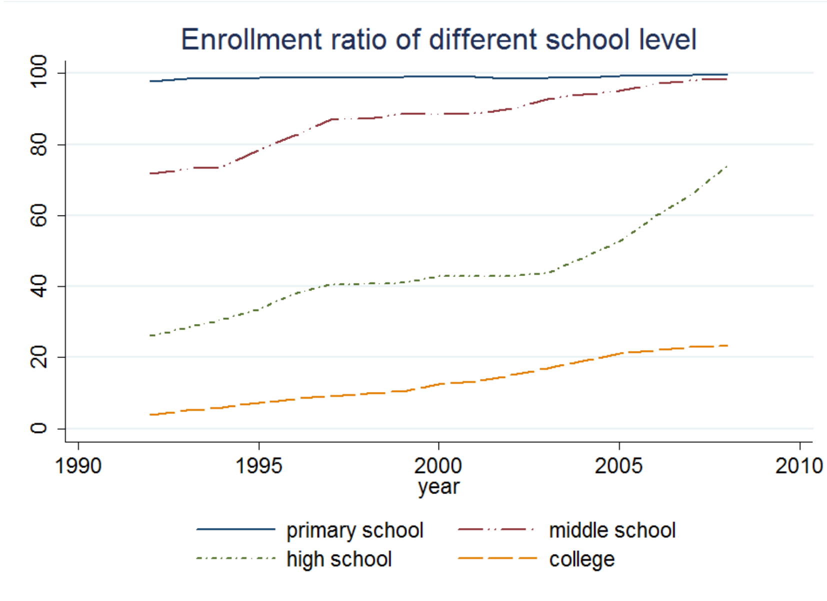
Note: the X axis is the talent and the Y axis is the cumulative distribution. The blue line is the uniform seat distribution of p in the whole sample. The red line is the seat distribution in group 0.

Figure 18 Example: Human capital investment level



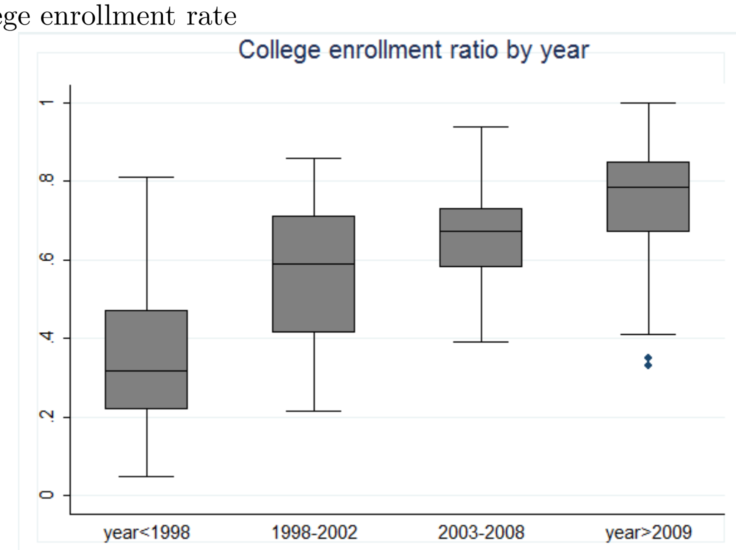
Note: the X axis is the talent and Y axis optimal human capital level.

Figure 19 Enrollment rate of different school level in China



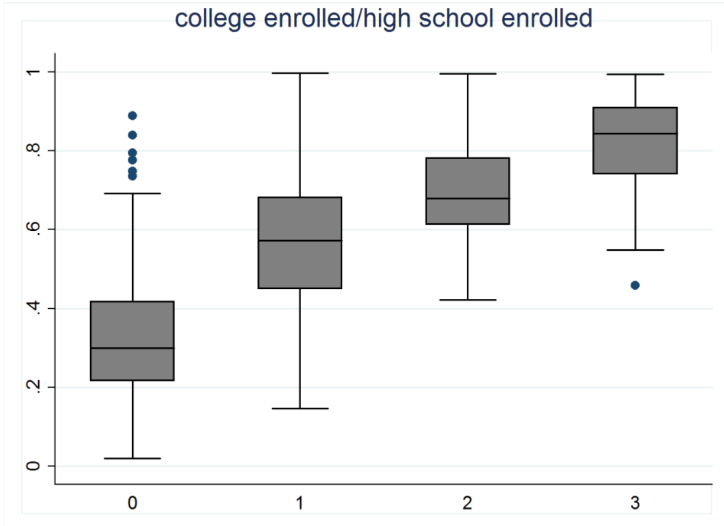
Note: In the figure above, the enrollment ratio means that the number of students enrolled in a certain level school to the number of people at the age to enter the school. In 1986, the Chinese began to provide nine-year compulsory education. After that, the primary school enrollment rate kept at about 100%, and the middle school enrollment rate increase from 70% to about 98%. The high school enrollment rate increased from 23% to about 75%, and college enrollment rate increase increased from 3% to 23%.

Figure 20 College enrollment rate



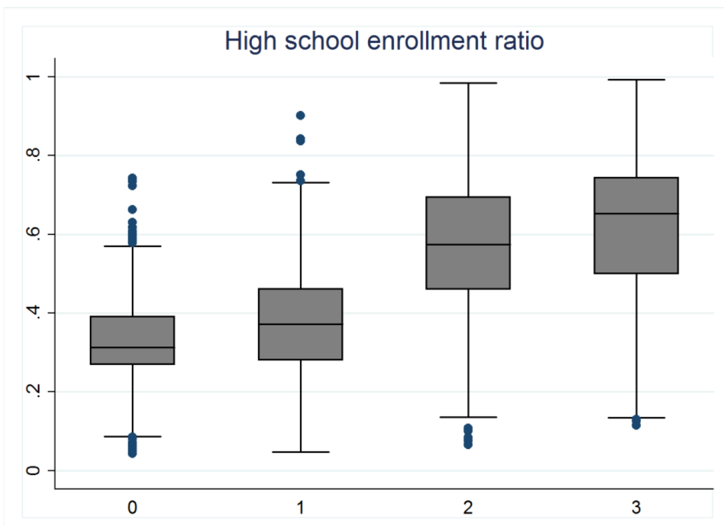
Note: Data from Chinese Education Yearbooks. College enrollment rate is defined by the number of students enrolled in college divided by the number of students that took the exam by year and province

Figure 21 College enrollment rate



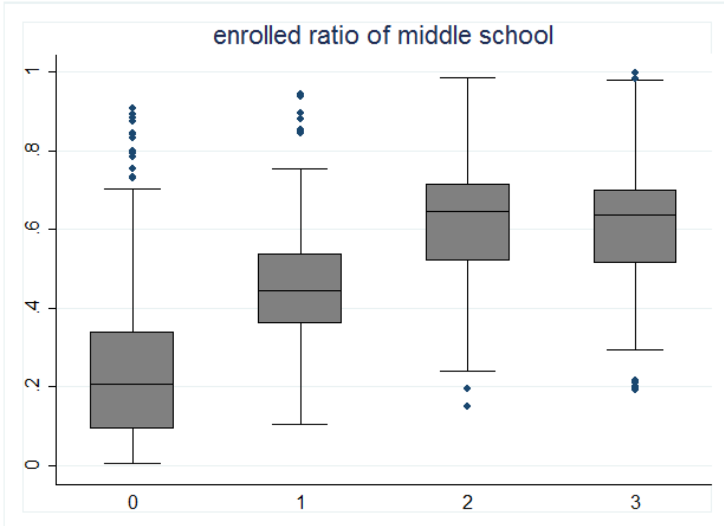
Note: Data from Chinese Educational Testing Yearbooks. College enrollment rate is defined by the number of students enrolled in college divided by the number of students graduated from high school. 0 is the sample before 1998; 1 is the sample between 1998 and 2002; 3 is the sample between 2003 and 2008; 4 is the sample after 2009.

Figure 22 High school enrollment ratio



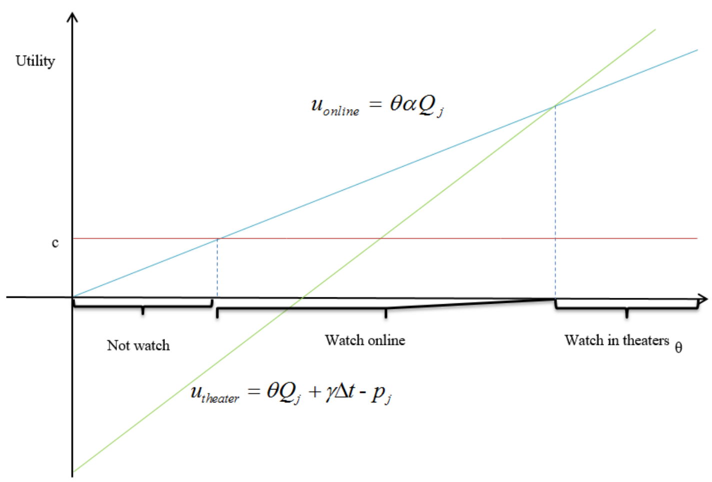
Note: Data from China Education Yearbook. High school enrollment rate means the number of students enrolled in high school divided by the number of students graduate from middle school) by year and province. 0 is the sample before 1998; 1 is the sample between 1998 and 2002; 3 is the sample between 2003 and 2008; 4 is the sample after 2009.

Figure 23 Middle school enrollment rate



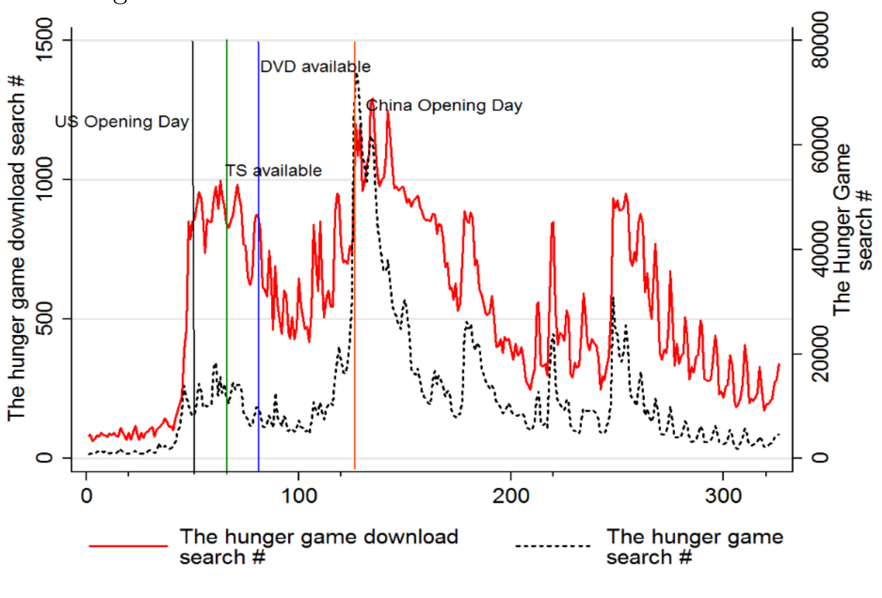
Note: Data from Chinese Education Yearbook. Middle school enrollment rate is defined by the number of student enrolled by middle school divided by the number of students graduated from primary school by year and province. Group 0 is the sample before 1998; group 1 is the sample between 1998 and 2002; group 2 is the sample between 2003 and 2008; group 3 is the sample after 20.

Figure 24 the Consumers' Choices



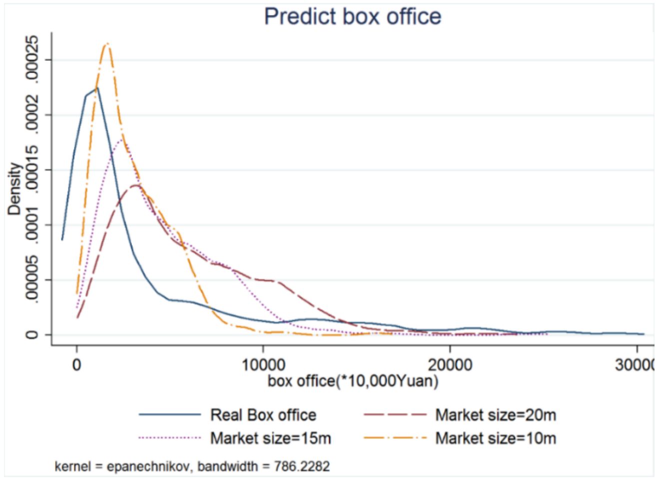
Notes: The y axis is the utility value. The x axis is the personal taste. u_{online} is the utility value of watching movie j online given different taste level θ. $u_{theater}$ is the utility of watching the movie j in the theater. Given the value of the two function if both utility values are smaller than 0, the consumer will not watch the movie. If watching movie online brings higher utility than watching in theaters and not watching the movie, consumer will choose to watch online; if watch movie in theater has higher utility than watching online and not watching the movie, the consumer will watch movie in theater.

Figure 25 The Hunger Games Baidu Search



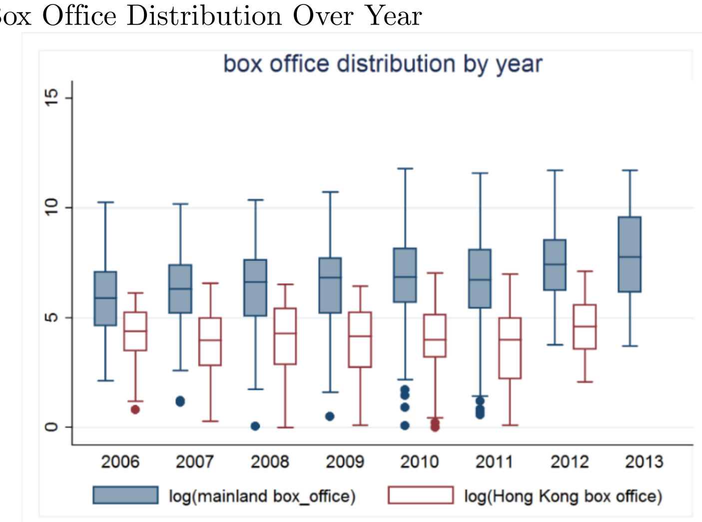
Notes: The data is from Baidu search data. The x axis is the day. The y axis is the search amount on Baidu. The number means the daily search amount per 10 million users. “The Hunger Games Download” search amount means the daily search amount of “The Hunger Games Download” (饥饿游戏下载) in Chinese. “The hunger game” search amount means the daily search amount of “The Hunger Games Download” (饥饿游戏) in Chinese. In this graph, the opening lag equals to the Chinese Opening day(the red line) minus the US opening day(the black line). TS/Clear lag equals the Chinese Opening day minus the TS(the green line)/ the DVD availability(the blue line) day. The Baidu search amount variable defined in the paper is the overall “The Hunger Games Download” search amount from the TS available day to the two weeks after China opening.

Figure 26 Predicted box office from estimation



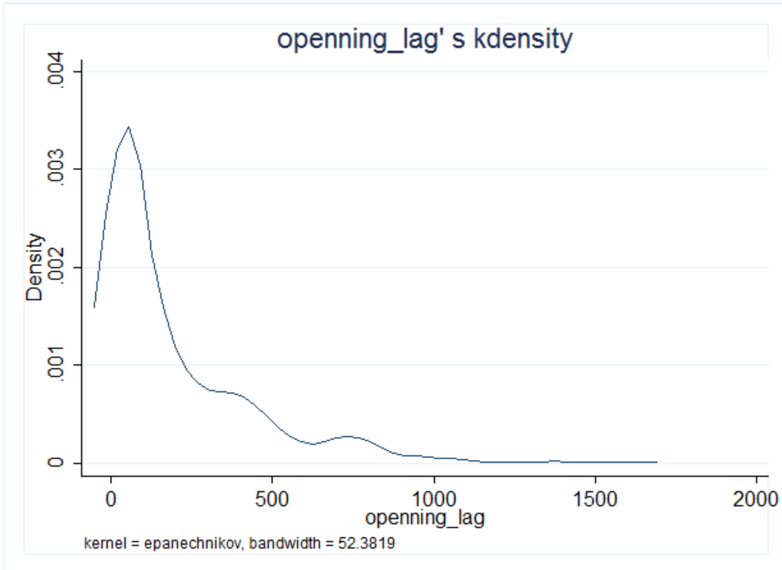
Notes: The data is from the real box office data and the estimation result. It used the subsample of the observations with weekly price information. The solid line is from the real box office data. The line of “market size=20m” is the predict box office assuming market size equaling to 20million. The line of “market size=10m” is the predict box office assuming market size equaling to10 million. The line of “market size=15m” is the predict box office assuming market size equaling to15 million.

Figure 27 The Box Office Distribution Over Year



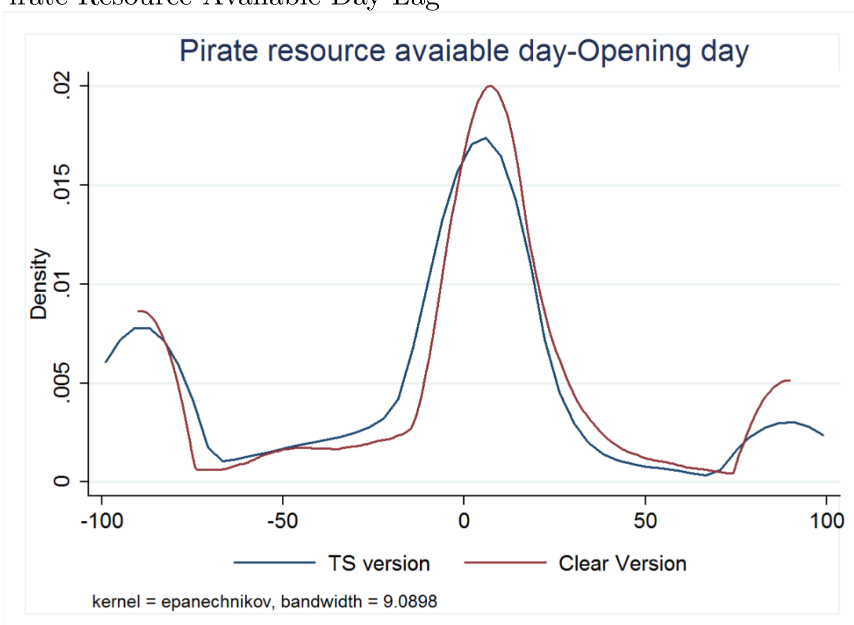
Notes: The mainland box office data is from the Chinese mainland box office data set. The mainland box office value is 10,000 Yuan (about 1,500~1,600 US Dollar). The box office data is from the Hong Kong box office data set. The Hong Kong box office value is 10,000 Hong Kong dollars. The upper line means the upper adjacent value. The lower line means the lower adjacent value. The upper hinge of the box means the 75th percentile value. The lower hinge of the box means the 25th percentile value. The middle line of the box means the median value.

Figure 28 Opening Lag Distribution



Notes: The data is from the movie characteristic data set. The sample is from the movies with positive opening lag (433 from 1040 movies). The x axis is valued in days.

Figure 29 Pirate Resource Available Day Lag



Notes: The data is from the movie characteristic data set. The x axis is valued in days. In the graph, "-100" means the value is smaller than 100 days; "100" means the value is larger than 100 days.

Table 1 Individuals and Household characteristics

Variable	Mean	SD	N	Variable	Mean	SD	N
CHNS Data							
working (age 24 to 60)	62%	0.49	11906	working time	0.34	0.13	8927
education year (age 24 to 60)	7.54	4.67	11906	annual HH income	35378	44957	10102
HH food and durable good consumption (age larger than 45)	20161	30270	20161	working income rate	33110	41474	8927
HH food and durable good consumption	31774	60717	53483	working income	11902	12467	8927
rural HH	80%	0.4	11906	HH net non-working income	27495	36770	31186
child care time	0.45	0.51	17128	Leisure	0.43	0.37	79613
house value (age larger than 45)	402506	724529	7967	Asset level (age larger than 45)	498564	1029226	7967
house value	364609	684929	21366	Asset level	455223	986722	21366
child taken care by outside service	19%	0.39	8,543	hour/day taken care by outside service	2.66	4.54	1,623
CHARLS Data							
education year	4.97	4.33	22768	rural HH	78%		22768
working	0.19	0.39	22768	eldest child's annual income	34908	39943	7142
HH food and durable good consumption	23131	66379	38177	total HH annual income	24618	69178	31186
HH consumption	30316	72578	31186	total HH annual working income	12197	48741	31186
asset level	423270	4470281	9368	total HH health spending	2275	8882	31186
house value	309035	3738449	9368	transfer from child last year	71%		22,228
money value from child last year	1537	3,477	22,228	transfer from parent last year	2%		22,228
child care hour last week	29	28	22,228	transfer to child last year	31%		22,228
money value from other relatives last year	459	1,962	22,228	transfer to child last year	26%		22,228
parent care hour last week	8.73	26.89	22,228	transfer to parent last year, age¡60	0.15	0.15	35,877
child care hour last week	31	30	22,228	transfer to parent last year, age¡60	0.11	0.12	8,877

Note: The money unit is Yuan. I define working as the people have any kind of income from farming, fishing, gardening, and business. I define migration as whether the people live in the same city or town. Retirement income is the sum value of pension and retirement subsidy. Consumption in CHNS data is all the spending on food and durable goods. Consumption in CHARLS data is all the spending on foods, durable good, clothes, traffic and other consumptions.

Table 2 estimated parameters

Parameter	Definition	Method	Mean	SE
η	Mean utility weight on leisure	Two stage GLS	0.88	(0.02)
ξ^{η}	The variance of utility weight on leisure	Two stage GLS	0.78	
α	Mean utility weight on child care	Two stage GLS	0.31	(0.01)
ξ^{α}	The variance of weight on child care	Two stage GLS	0.12	
γ	Mean utility weight on elder care (64-72)	Two stage GLS	0.06	(0.01)
ξ^{γ}	The variance of utility weight on elder care (64-72)	Two stage GLS	0.04	
γ	Mean utility weight on elder care (73-80)	Two stage GLS	0.20	(0.02)
ξ^{γ}	The variance of utility weight on elder care (73-80)	Two stage GLS	0.05	
γ	Mean utility weight on elder care (after 81)	Two stage GLS	0.52	(0.06)
ξ^{γ}	The variance of weight on elder care (after 81)	Two stage GLS	0.73	
ρ	Type 1 probability	GMM	0.42	
η_1	Type1 utility weight on leisure	GMM	0.75	
η_2	Type2 utility weight on leisure	GMM	0.97	
α_1	Type1 utility weight on child care	GMM	0.73	
α_2	Type2 utility weight on child care	GMM	0.01	
γ_1	Type1 utility weight on child care (64-72)	GMM	0.14	
γ_2	Type2 utility weight on child care (64-72)	GMM	0.00	
γ_1	Type1 utility weight on child care (73-80)	GMM	0.35	
γ_2	Type2 utility weight on child care (73-80)	GMM	0.09	
γ_1	Type1 utility weight on child care (after 81)	GMM	0.81	
γ_2	Type2 utility weight on child care (after 81)	GMM	0.31	
ι	Substitution rate of child/elder care from market and relatives	MLE	2.51	(0.79)
ϕ	Correlation between mother and grandmother's wage rate	Correlation	0.46	
β	Discount factor	Literature	0.97	
R	Real interests rate	From data	1.01	
w_{p0}	Initial wage rate of mother	OLS	55457	(21072)
w_{g0}	Initial wage rate of grandmother	OLS	59519	(22383)
π_1	Wage growth rate by age	OLS	0.056	(0.04)
π_2	Wage growth rate by age square	OLS	-0.0008	(0.00006)

Note: The data estimating income, child care and leisure parameters are from CHNS 2006-2011. Data estimating elder care parameters are from CHARLS 2008-2011. The money unit is Yuan. I use GMM methods to estimate the two type preferences. The interest rate is taken from World Bank, which is the average interest rate from 2006 to 2012.

Table 3 Actual vs. predicted choices and select measures

	Data	Contract	Autarky case
Parents' working time	0.36	0.34	0.33
Parent's working time (before 40)	0.37	0.36	0.27
Parent's working time (After 40)	0.35	0.34	0.38
Parents' leisure	0.43	0.44	0.37
Grandparents' working time	0.23	0.27	0.34
Grandparents' leisure	0.47	0.49	0.45
Child care from outside service	19%	17%	26%
Child care by grandparents	56%	62%	0
Elder care from parents	58%	63%	0
Elder care from outside service	7%	9%	100%
Transfer from parents last year	31%	34%	0
Transfer to parents last year	20%	13%	0
Grandparents' saving	423,270	378,515	318,080
Parents' saving	375,412	345,710	367,777

Note: The information of working, leisure, child care, and saving of data is from CHNS. Elder care information is taken from CHARLS. The prediction values are taken from simulation. Money unit is Yuan. The parameters are following benchmark setting: care service subsidies are 0; inheritance tax is 0; mandatary retirement age is 10; the wage rate is given by wage structure 1 (shown in the first graphs of Figure 1); 42% are type 1 and 58% of households are type 2; grandparents' initial saving is 375,000; and I normalize the overall time to 1. The second graph compares the labor supply by age from simulation and data. The simulation result is using a contract with benchmark setting.

Table 4 Province level education resource information

Stats	mean	sd	N
GDP per capital(Yuan)	2549.7	6324.85	1978
Population (*10000)	3427.91	2357.93	1888
Education Fund (*10000Yuan)	2671233	2665574	512
National education budget (*10000 Yuan)	1695296	1775728	512
Tuition income(*10000Yuan)	423943	475143.1	512
Local college #	48.91	30.58	864
New college students # (*10000)	10.02	11.02	864
College student # (*10000)	14.76	28.29	2054
College graduate # (*10000)	18.21	12.04	288
College teacher # (*10000)	1.29	1.66	2054
High school #	487.91	242.38	864
New high school student # (*10000)	17.4	15.23	864
High school student # (*10000)	48.87	43.42	864
Middle school #	566.1	1133.3	2054
New Middle school student # (*10000)	36.51	31.81	2054
Middle school student # (*10000)	145.9	146.32	2054
Middle school graduate # (*10000)	32.71	30.16	2054
Middle school teacher # (*10000)	10.02	13.05	2054
Primary school #	36600	8.11	2054
New Primary school student # (*10000)	51.55	33.87	2054
Primary school student # (*10000)	343.69	254.38	2054
Primary school graduate # (*10000)	69.94	88.89	2054
Primary school teacher # (*10000)	15.28	10.46	2054
High school graduate # (*10000)	14.49	13.53	864
High school teacher # (*10000)	1.29	1.66	2054

Note: Data from China's national statistical bureau. Money unit is Yuan.

Table 5 Enrollment information

Stats	mean	sd	N
NCEE enrolled # (*10000)	15.8	15.12	1215
NCEE taken # (*10000)	65.79	2189.39	1535
Humanities exam taker # (*10000)	2.33	2.86	1215
Humanities exam enrolled # (*10000)	5.22	4.94	1023
Science exam taker # (*10000)	4.22	3.87	1535
College # of the provinces	91.08	43.28	329235
New college # of the r province in last year	3.23	5.33	329235
New college # of the province in last 3 years	6.94	8.91	329235
New college # of the province in last 5 years	12.2	12.46	329235
Science exam enrolled # (*10000)	9.3	5.2	1535
Rank 1 college enrolled # (*10000)	18118.17	10943.54	863
Rank 2 college enrolled # (*10000)	30726.38	23048.6	863
Rank 3 college enrolled # (*10000)	20292.98	8606.91	863
Rank 4 college enrolled # (*10000)	70142.03	90726.96	863
New college # of neighbor province in last 3 years	34.47	34.68	329235
New college # of neighbor province in last 5 years	58.87	47.48	329235
New student # of neighbor province in last year (*10000)	70.49	67.08	329235
New college # of neighbor province in last year	16.94	21.92	329235

Note: The NCEE has four big majors- science and engineering, humanities, art and sports. Each student could only take one of them. All the four majors have three mandatory subjects: Chinese, Mathematics, and a foreign language—usually English. The students of the science and engineering major need to choose one to three from Physics, Chemistry, and Biology. The students of the humanities major need to choose one to three from History, Geography, and Political Education. There are 4 ranks of college: key universities, regular four year universities, regular three year universities, other regular three year universities and technical colleges. The higher ranking college enrolls students first. After the higher level finishes their enrollment, the next level college begins their enrollment. The university rank is also well known to everyone.

Table 6 Individual and Household Characteristics

Stats	mean	sd	N
Study hour after school	555.19	2401.83	22532
Child work before 16	0.14	0.35	48305
Leave School Before 18	0.26	0.44	68963
Leave School Before High School	0.35	0.48	95305
Good School Performance	0.42	0.49	48305
Bad School Performance	0.02	0.13	48305
Age	33.81	28.78	317919
Female	0.46	0.5	317919
Household income	119754	967336	148027
Rural household	0.22	0.42	317919
# of kids 0-6	0.23	0.51	83138
# of kids 6-12	0.32	0.6	83138
# of kids 12-15	0.2	0.45	83138
# of kids 15-18	0.61	0.87	83138
# of people 16-60	2.18	1.28	83138
# of people ¿60	0.35	1.1	83138
Education Spending	0.23	0.51	45466
Education year	7.16	4.37	315339
Communist Party Member	0.09	0.29	209216
Minority	33.81	28.78	315917

Note: The household and individual data is from Chinese Household Income Project. The data were collected through a series of questionnaire-based interviews conducted in rural and urban areas in 1988, 1995, 2002, 2007 and 2008. Individual respondents reported on their economic status, employment, level of education, sources of income, household composition, and household expenditures. Study hour is the annual study hour at home.

Table 7 Main results

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Education spending/household income				
College Enrollment Ratio(province level)	0.279***	0.0893***	0.0788***	0.0710***	0.0825***
	(0.00358)	(0.00970)	(0.0107)	(0.00998)	(0.0105)
Rural*College Enrollment Ratio			-0.0227**		
			(0.0103)		
College Enrollment Ratio*HH With High School Degree				0.0556***	
				(0.00855)	
Household income*College Enrollment Ratio					9.55e-07*
					(5.44e-07)
Household income					-4.24e-06***
					(3.21e-07)
age		0.00804***	0.00804***	0.00826***	0.00779***
		(0.00135)	(0.00135)	(0.00135)	(0.00132)
age square		-0.000244***	-0.000245***	-0.000254***	-0.000236***
		-5.06E-05	(5.06e-05)	(5.05e-05)	(4.96e-05)
female		-0.00183	-0.00183	-0.00182	-0.00166
		(0.00113)	(0.00113)	(0.00113)	(0.00111)
Rural HH		-0.0804***	-0.0754***	-0.0658***	-0.0371***
		(0.00759)	(0.0186)	(0.00777)	(0.00756)
HH asset		0**	0**	0**	0
		(0)	(0)	(0)	(0)
# of people age¡6		0.000958	0.000950	0.000647	0.000780
		(0.00123)	(0.00123)	(0.00123)	(0.00121)
# of people 5¡age¡12		0.0109***	0.0109***	0.0110***	0.0109***
		(0.000877)	(0.000877)	(0.000875)	(0.000860)
# of people 11¡age¡16		0.0117***	0.0117***	0.0119***	0.0122***
		(0.00101)	(0.00101)	(0.00101)	(0.000993)
# of people 15¡age¡18		0.00211***	0.00211***	0.00230***	0.00296***
		(0.000703)	(0.000703)	(0.000701)	(0.000692)
# of people 19¡age¡45		-0.00770***	-0.00770***	-0.00747***	-0.00598***
		(0.000959)	(0.000959)	(0.000958)	(0.000943)
# of people 45¡age¡61		-0.00482***	-0.00482***	-0.00489***	-0.00177
		(0.00109)	(0.00109)	(0.00109)	(0.00108)
# of people age¿60		-0.00454***	-0.00455***	-0.00447***	-0.00283***
		(0.00106)	(0.00106)	(0.00106)	(0.00105)
The highest education level in the HH		0.00225***	0.00225***	0.000665**	0.00263***
		(0.000216)	(0.000216)	(0.000279)	(0.000212)
Household with Communism Party Member		0.00267*	0.00269*	0.00236	0.00377**
		(0.00158)	(0.00158)	(0.00158)	(0.00155)
Household with minority		0.00425*	0.00427*	0.00332	0.00174
		(0.00230)	(0.00230)	(0.00230)	(0.00226)
Household with migrant		-0.0133	-0.0133	-0.0133	-0.0213
		(0.0138)	(0.0138)	(0.0138)	(0.0135)

Household with people working in state own firm		0.00533**	0.00532**	0.00457*	0.00480**
		(0.00245)	(0.00245)	(0.00245)	(0.00241)
Household with people working in foreign firm		-0.0220***	-0.0221***	-0.0235***	-0.0159***
		(0.00539)	(0.00539)	(0.00537)	(0.00528)
Household with people working in government		-0.00658***	-0.00657***	-0.00778***	-0.00148
		(0.00184)	(0.00184)	(0.00184)	(0.00181)
Year dummies		Control	Control	Control	Control
Province dummies		Control	Control	Control	Control
Province*rural dummies		Control	Control	Control	Control
Constant	-0.0260***	0.0798***	0.0746***	0.0794***	0.113***
	(0.00125)	(0.0192)	(0.0259)	(0.0193)	(0.0189)
Observations	28,169	28,169	28,169	28,169	28,169
R-squared	0.177	0.261	0.261	0.265	0.290

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The household and individual data is from Chinese Household Income Project.

Table 8 2SLS results

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household education spending/household income						
	OLS	2SLS					
Enrollment Ratio	0.0893***	0.0823***	0.0909*	0.413***	0.0854***	0.0544***	0.0823***
	(0.00970)	(0.0199)	(0.0496)	(0.0677)	(0.0116)	(0.0205)	(0.0212)
Rural*College Enrollment Ratio				-0.361***	-0.0203*		
				(0.0632)	(0.0112)		
College Enrollment Ratio*HH With High School Degree						0.0721***	
						(0.00909)	
Household income*College Enrollment Ratio							1.46e-06**
							(6.57e-07)
Control other variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.0798***	0.113***	0.106***	0.119***	0.185***	0.116***	0.0393*
	(0.0192)	(0.0230)	(0.0408)	(0.0229)	(0.0565)	(0.0226)	(0.0208)
Observations	28,169	28,169	28,169	28,169	28,169	28,169	28,169
R-squared	0.261	0.293	0.293	0.293	0.292	0.297	0.293
Instrumental variables		New college # of neighbor province	Enrolled student # of neighbor provinces' colleges	Both	Both	Both	Both
First stage t test		F(4, 28106)=26.45 Prob¿F=0.0000	F(1, 28108) =3.34 Prob¿F=0.0677	F(6, 28105) =20.36 Prob¿F=0.0000			
Tests of over identifying restrictions:		Sargan statistic: 25.614 Chi-sq(2) P-val =0.0000	Sargan statistic: 30.415 Chi-sq(3) P-val =0.0000	Sargan statistic: 34.512 Chi-sq(4) P-val= 0.0000			
Tests of endogeneity of: enrollment ratio		Endogeneity test : 0.791 Chi-sq(1) P-val = 0.3739	Endogeneity test: 0.019 Chi-sq(1) P-val= 0.8907	Endogeneity test : 1.945 Chi-sq(1) P-val = 0.1631			

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The instrumental variables "New college # of neighbor province" are College # of neighbor provinces, New college # of neighbor province in last year, New college # of neighbor province in last 3 year and New college # of neighbor province in last 5 year. The instrumental variables "enrolled student # of neighbor provinces' colleges" is enrolled student # of neighbor provinces' colleges.

Table 9 the first stage of 2SLS results

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Household education spending/household income			College enrollment Ratio		
College # of neighbor provinces	-0.000154***		-0.000127***	-0.000537***		-0.000579***
	(1.57e-05)		(1.75e-05)	(7.92e-06)		(8.83e-06)
New college # of neighbor province in last year	-0.00206***		-0.00219***	-0.0144***		-0.0142***
	(0.000427)		(0.000428)	(0.000216)		(0.000216)
New college # of neighbor province in last 3 year	0.00114***		0.00168***	0.00419***		0.00333***
	(0.000428)		(0.000455)	(0.000216)		(0.000230)
New college # of neighbor province in last 5 year	2.36e-05		-0.000138	0.00173***		0.00198***
	(0.000252)		(0.000256)	(0.000128)		(0.000129)
Enrolled student # of neighbor provinces' colleges		-0.000189*	-0.000596***		-0.00208***	0.000936***
		(0.000103)	(0.000171)		(5.95e-05)	(8.62e-05)
The college # of the province in last year	-7.59e-05		1.95e-05	-0.00502***		-0.00517***
	(0.000144)		(0.000147)	(7.30e-05)		(7.41e-05)
New college # of the province in last 1 year	-0.000622	-0.00247***	-0.00151	-0.0113***	-0.0251***	-0.00990***
	(0.000899)	(0.000868)	(0.000933)	(0.000454)	(0.000499)	(0.000471)
New college # of the province in last 3 year	0.00586***	0.00478***	0.00693***	0.0331***	0.0282***	0.0314***
	(0.00111)	(0.00105)	(0.00115)	(0.000559)	(0.000605)	(0.000579)
New college # of the province in last 5 year	-0.00515***	-0.00404***	-0.00577***	-0.0199***	-0.0145***	-0.0189***
	(0.000766)	(0.000717)	(0.000786)	(0.000387)	(0.000412)	(0.000397)
Enrolled student # of the province's colleges in last year	-0.0115***	-0.00340	-0.0124***	-0.0367***	0.0121***	-0.0353***
	(0.00264)	(0.00214)	(0.00265)	(0.00133)	(0.00123)	(0.00134)
Control other variables	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.275***	0.185***	0.269***	1.567***	0.860***	1.577***
	(0.0292)	(0.0186)	(0.0292)	(0.0147)	(0.0107)	(0.0147)
Observations	28,169	28,169	28,169	28,169	28,169	28,169
R-squared	0.291	0.291	0.292	0.920	0.897	0.921

Note: Standard errors in parentheses,*** p¡0.01, ** p¡0.05, * p¡0.1。 The control variables are same as the regression of the main results.

Table 10 Other education investment

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Study hour after school 2SLS			Hour taken care of children 2SLS		
College Enrollment Ratio	7.573***	7.975***	10.72***	1.201***	1.665**	-0.906
	(2.549)	(2.554)	(3.069)	(0.35)	(0.69)	(1.17)
College Enrollment Ratio*Rural		-5.416**			-0.374	
		(2.381)			(0.48)	
College Enrollment Ratio*HH With High School Degree			3.546*			3.110**
			(1.927)			(1.37)
Other variables	Control	Control	Control	control	control	control
Constant	-1.904	-2.485	-3.865	-25.25	-29.21	-64.82
	(2.941)	(2.951)	(3.147)	(85.94)	(88.04)	(177.70)
Observations	5,555	5,555	5,555	9,635	9,635	9,635
R-squared	0.127	0.128	0.128	0.196	0.159	0.178
	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	Quit school before graduate from high school Probit			Entertainment min per week 2SLS		
college enrollment ratio	-0.121	-3.749***	0.257	-5,755**	-14,184	-8,555**
	(0.26)	(0.44)	(0.26)	(2,927)	(8,724)	(4,039)
college enrollment ratio*rural hukou		5.125***			5,737	
		(0.47)			(5,540)	
college enrollment ratio*household with high school educated			-1.058***			807.3
			(0.18)			(759.9)
Other variables	Control	Control	Control	control	control	control
Constant	3.783***	5.665***	3.828***	19,821	-385,703	-239,862
	(0.52)	(0.55)	(0.53)	(321,038)	(521,544)	(415,278)
Observations	48,675	48,675	48,675	10,993	10,993	10,993
R-squared				0.347	0.248	0.285

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The sample is 7-19 years old. Study hour is the annual study hour at home. The dummy quit school before graduating from high school is the dummy that people quit or graduate before they get high school degree. Work dummy is defined by whether the people work or not. House work hour is the weekly house work hour. The control variables are same as main results regression.

Table 11 Robustness check on different age groups

	(1)	(2)	(3)	(4)	(5)
Age group	Age 7-19	age 16-19	age 12-15	age 7-11	age¡7
College enrollment ratio	0.0893***	0.0794***	0.0964***	0.0847***	0.0464***
	(0.00970)	(0.0174)	(0.0182)	(0.0148)	(0.0148)
Other variables	Control	Control	Control	Control	Control
Constant	0.0798***	0.637**	-0.0119	0.0702	-0.0332
	(0.0192)	(0.310)	(0.221)	(0.0490)	(0.0240)
Observations	28,169	9,690	8,967	9,512	8,079
R-squared	0.261	0.272	0.264	0.275	0.252

	(6)	(7)	(8)	(9)	(10)	(11)
Age group		Age 7-19		age 7-11	age 12-15	age 16-19
Middle school enrollment ratio	-0.0282***			-0.0338***		
	(0.00622)			(0.00944)		
High school enrollment ratio		0.0393***			0.0156	
		(0.00839)			(0.0153)	
Local college enrollment ratio			0.0293***			0.0250*
			(0.0108)			(0.0204)
Other variables	Control	Control	Control	Control	Control	Control
Constant	0.167***	0.136***	0.131***	0.158***	0.0248	0.643**
	(0.0179)	(0.0179)	(0.0192)	(0.0476)	(0.222)	(0.311)
Observations	28,169	28,169	28,169	9,512	8,967	9,690
R-squared	0.260	0.260	0.259	0.273	0.262	0.271

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The control variables are the same as the regression of the main results. Middle school enrollment ratio is the number of new middle school students divided by the number of primary school graduate student in the year. High school enrollment ratio is the number of new high school students divided by the number of middle school graduate students in the year. Local college enrollment ratio is the number of new college students divided by the number of high school graduate student number in the year. Middle school enrollment ratio is the number of new middle school students divided by the number of primary school graduate student in the year. High school enrollment ratio is the number of new high school students divided by the number of middle school graduate student in the year. Local college enrollment ratio is the number of new college students divided by the number of high school graduate students in the year.

Table 12 robust check on different type of households

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Education spending/household income					
	female	male	Minority	Not minority	HH with high school	HH without high school

					degree	degree
enrollment ratio	0.0600**	0.0939***	-0.138**	0.0471**	0.141***	0.0458
	(0.0256)	(0.0285)	(0.0672)	(0.0206)	(0.0330)	(0.0352)
Other variables	Control	Control	Control	Control	Control	Control
Constant	0.0591**	0.104**	0.113**	0.144***	-0.0686*	0.138***
	(0.0285)	(0.0474)	(0.0491)	(0.0237)	(0.0400)	(0.0289)
Observations	13,555	14,614	2,512	25,657	23,813	15,387
R-squared	0.290	0.300	0.318	0.296	0.249	0.216

	(7)	(8)	(9)	(10)	(11)
VARIABLES	Education spending/household income				
Income percentile	0-20 percentile	20-40 percentile	40-60 percentile	60-80 percentile	80-100 percentile
enrollment ratio	-0.0501**	-0.103	0.688***	0.148***	0.0336*
	(0.0216)	(0.0902)	(0.208)	(0.0321)	(0.0299)
Other variables	Control	Control	Control	Control	Control
Constant	0.221	0.422***	-0.174**	-0.0702**	0.0180
	(0.244)	(0.0729)	(0.0853)	(0.0324)	(0.0444)
Observations	30,296	15,853	27,706	42,710	42,551
R-squared	0.213	0.238	0.148	0.145	0.129

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The control variables are same as the regression of the main results. Minority is the households belonged to the minority races. High school degree household is the household where at least one person has a high school or higher degree level. No high school degree household is the household without anyone with at least high school degree.

Table 13 Other NCEE enrollment variables

	(1)	(2)	(3)
VARIABLES	Education spending/household income		
All college enrollment ratio	0.0901***		
	(0.00951)		
Rank 1 college enrollment ratio		0.0628**	
		(0.0256)	
4 years college enrollment ratio			0.0447**
			(0.0178)
Other variables	Control	Control	Control
Constant	0.111***	0.162***	0.169***
	(0.0189)	(0.0196)	(0.0183)
Observations	28,169	28,169	28,169
R-squared	0.290	0.288	0.288

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The control variables are same as the regression of the main results. Rank 1 colleges are the top 112 colleges or universities belong to the 211 Project. The 211 Project is a strategic cross-century project formulated by the Chinese government for the implementation of the strategy for invigorating the country through science, technology and education. The variable is defined by the number of students enrolled by 211 universities divided by the number of students that took the NCEE in the same year of the same province.

Table 14 Return to education

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Log(individual income)					
	All	Before 1998	1998-2004	After 2004	Rural	Urban
People with college education	0.675***	0.311***	0.632***	0.913***	0.439***	0.797***
	(0.0249)	(0.0557)	(0.0533)	(0.0358)	(0.0695)	(0.0260)
People with high school education	0.401***	0.338***	0.449***	0.460***	0.369***	0.484***
	(0.0161)	(0.0297)	(0.0324)	(0.0254)	(0.0248)	(0.0210)
People with middle school education	0.173***	0.255***	0.182***	0.200***	0.146***	0.255***
	(0.0130)	(0.0241)	(0.0265)	(0.0205)	(0.0168)	(0.0211)
age	0.0948***	0.133***	0.0844***	0.0722***	0.107***	0.0703***
	(0.00380)	(0.00665)	(0.00797)	(0.00653)	(0.00528)	(0.00494)
age square	-0.00115***	-0.00161***	-0.00104***	-0.000826***	-0.00135***	-0.000795***
	(4.58e-05)	(8.35e-05)	(9.73e-05)	(7.60e-05)	(6.38e-05)	(5.91e-05)
female	-0.246***	-0.173***	-0.246***	-0.220***	-0.247***	-0.252***
	(0.0104)	(0.0191)	(0.0212)	(0.0162)	(0.0147)	(0.0132)
rural hukou	-0.525***	-0.473***	-0.686***	-0.404***		
	(0.0124)	(0.0218)	(0.0252)	(0.0200)		
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes				Yes	Yes
Constant	7.844***	4.879***	6.994***	8.043***	7.142***	8.156***
	(0.0893)	(0.136)	(0.163)	(0.145)	(0.125)	(0.114)
Observations	48,590	13,475	13,279	21,836	30,441	18,149
R-squared	0.393	0.141	0.174	0.179	0.296	0.486

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The sample is people aged 20 to 60. I define rural people as the people in rural Hukou and urban people as the people in urban Hukou.

Table 15 the influence of local public education level

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Education spending/household income					
College enrollment ratio	0.0662***	0.0432*	0.0863***	0.210***	0.109***	0.0597*
	(0.0175)	(0.0244)	(0.0183)	(0.0429)	(0.0141)	(0.0354)
high school teacher student ratio	-0.447***	-0.530***				
	(0.0997)	(0.117)				
high school teacher student ratio*enrollment ratio		0.239				
		(0.178)				
middle school teacher student ratio			-0.0166	0.564***		
			(0.106)	(0.211)		
middle school teacher student ratio*enrollment ratio				-1.718***		
				(0.538)		
primary school teacher student ratio					0.659***	0.362
					(0.138)	(0.241)
primary school teacher student ratio*enrollment ratio						0.941
						(0.626)
Other variables	Control	Control	Control	Control	Control	Control
Age group	16-19 year old		12-15 year old		7-12 year old	
Constant	1.506***	1.606***	0.611*	0.766**	0.0664	0.0531
	(0.390)	(0.397)	(0.314)	(0.318)	(0.186)	(0.186)
Observations	9,690	9,690	8,967	8,967	11,664	11,664
R-squared	0.294	0.294	0.294	0.294	0.313	0.313

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The control variables are same as the regression of the main results. The education level is the provincial education level. For example high school teacher student ratio is defined by the number of high school teacher divided by number of high school student.

Table 16 factors affect enrollment ratio

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	# of student taken NCEE	# of student enrolled by NCEE	Enrollment ratio	# of student taken NCEE	# of student enrolled by NCEE	Enrollment ratio	# of student taken NCEE	# of student enrolled by NCEE	Enrollment ratio
high school graduate #	0.897***	0.524***	-0.00756***	0.911***	0.524***	-0.00742***	0.857***	0.492***	-0.00783***
	(0.0353)	(0.0163)	(0.000814)	(0.0342)	(0.0163)	(0.000810)	(0.0363)	(0.0174)	(0.000819)
local college enroll #	0.252***	0.398***	0.0134***	0.0296	0.381***	0.0111***	0.707***	0.483***	0.0254***
	(0.0488)	(0.0235)	(0.00113)	(0.0554)	(0.0267)	(0.00131)	(0.0864)	(0.0388)	(0.00193)
log(GDP)	-0.577**	0.0833	0.150***	-1.504***	0.0579	0.138***	-1.060***	0.210*	0.152***
	(0.282)	(0.118)	(0.00647)	(0.298)	(0.119)	(0.00693)	(0.277)	(0.119)	(0.00618)
log(population)	0.458	-0.790***	-0.0668***	1.165***	-0.750***	-0.0609***	0.542	-0.448**	-0.0359***
	(0.459)	(0.141)	(0.00988)	(0.440)	(0.144)	(0.00975)	(0.542)	(0.190)	(0.0118)
neighbor province enroll #				0.0568***	0.00420	0.000598***			
				(0.00746)	(0.00321)	(0.000175)			
high school employer #							-0.0913	-0.505***	0.00299
							(0.223)	(0.106)	(0.00503)
high school teacher #							-1.085**	0.614***	-0.0638***
							(0.477)	(0.226)	(0.0107)
Method	Random effects	Random effects	Random effects	Random effects	Random effects	Random effects	Random effects	Random effects	Random effects
Constant	1.068	3.555***	0.0629	-0.113	3.305***	0.0733	4.163	1.644	-0.0782
	(3.769)	(1.113)	(0.0810)	(3.553)	(1.129)	(0.0784)	(3.978)	(1.391)	(0.0868)
Observations	999	999	999	999	999	999	999	999	999
Number of province	30	30	30	30	30	30	30	30	30

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The regression is on the provincial level. Other variables contain GDP, population, local primary school teacher student ratio, local middle school teacher student ratio, and local high school teacher student ratio.

Table 17 Signal effect of enrollment information

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	# of student taken NCEE			# of NCEE taker choose art and human sciences major			# of NCEE taker choose engineering and sciences major		
NCEE enrollment ratio at t	-0.151	-0.136	-0.135						
	(1.121)	(1.120)	(1.123)						
NCEE enrollment ratio at t-1		-1.263	-1.261						
		(1.094)	(1.091)						
NCEE enrollment ratio at t-2			0.0146						
			(1.131)						
Enrollment ratio of art and human sciences at t				0.523**	0.611**	0.573**	0.140	0.127	0.112
				(0.238)	(0.251)	(0.244)	(0.235)	(0.238)	(0.236)
Enrollment ratio of engineering and sciences at t				-0.365	-0.416	-0.438	-0.0464	-0.0376	-0.0454
				(0.269)	(0.282)	(0.274)	(0.265)	(0.267)	(0.265)
Enrollment ratio of art and human sciences at t-1					-0.736***	-0.546**		-0.00675	0.0589
					(0.249)	(0.244)		(0.236)	(0.236)
Enrollment ratio of engineering and sciences at t-1					0.798***	0.653**		-0.0715	-0.119
					(0.280)	(0.274)		(0.266)	(0.265)
Enrollment ratio of art and human sciences at t-2						1.584***			0.891***
						(0.240)			(0.232)
Enrollment ratio of engineering and sciences at t-2						-0.951***			-0.800***
						(0.272)			(0.264)
Year dummies	control	control	control	control	control	control	control	control	control
Province dummies	control	control	Control	control	control	control	control	control	control
Other variable	control	control	Control	control	control	control	control	control	control
Constant	27.50***	28.51***	28.71***	19.40**	8.280***	6.979***	44.03***	44.07***	33.87***
	(5.807)	(5.873)	(5.886)	(7.938)	(0.887)	(0.906)	(7.830)	(7.825)	(4.115)
Observations	1,214	1,213	1,212	1,022	1,021	1,020	1,022	1,021	1,020
R-squared	0.931	0.931	0.931	0.670	0.640	0.660	0.788	0.788	0.791

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The regression is on the provincial level. Other variables contain GDP, population, local primary school teacher student ratio, local middle school teacher student ratio, and local high school teacher student ratio.

Table 18 Individual income, education and college enrollment ratio (2SLS regression)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Log(Education years)							
The college enrollment at age 18	3.765***	2.990***	2.103***	1.641***	0.0373	0.0564*	0.640	-0.405**
	(0.661)	(0.695)	(0.329)	(0.231)	(0.0310)	(0.0296)	(0.414)	(0.189)
The college enrollment at age 18*Rural Hukou			-0.596**				-0.764*	
			(0.271)				(0.449)	
The college enrollment at age 18*HH with high school degree				-0.146				0.500*
				(0.125)				(0.260)
Province dummies		Control	Control	Control		Control	Control	Control
Province dummies*Rural Hukou		Control	Control	Control		Control	Control	Control
Province characters	Control	Control	Control	Control	Control	Control	Control	Control
Individual characteristics	Control	Control	Control	Control	Control	Control	Control	Control
Household characteristics before age 18	Control	Control	Control	Control	Control	Control	Control	Control
Year dummies	Control	Control	Control	Control	Control	Control	Control	Control
Constant	0.143	0.617	0.634	0.921	2.707***	2.738***	2.315***	2.842***
	(1.529)	(1.542)	(0.883)	(0.874)	(0.0638)	(0.0671)	(0.259)	(0.102)
Observations	5,569	5,569	5,569	5,569	10,979	10,979	10,979	10,979
R-squared	0.221	0.202	0.203	0.202	0.157	0.062	-0.126	-0.551

Note: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The sample selected is the people have both income information after 18 years old and household characteristics before 18 years old. Individual characteristics variables contain age, age square and gender. Household characteristics before age 18 contain household income, asset level, number of elders, and number of children, number of labor, highest education year, and rural Hukou dummy. Province characters contain GDP, population, high school student teacher rate, middle school student teacher rate, and primary student teacher rate.

Table 19 The Summary of Box Office and Other Information, Study Sample

	Box office (*10,000Yuan)	Douban Rate	Rating # on Douban	Opening day lag	TS day lag	Clear version day lag	Hong Kong box office (*10000 HKD)
Mean	5700	5.84	36549	79.066	-75.658	-56.927	138
SD	(12872.44)	(1.51)	(57295.28)	(186.72)	(291.88)	(294.14)	(18778.29)
Cast	Director level			Actor and Actress #(Leading roles)			
	Hollywood Top	China Top10	China Top 30	China Top 20	China Top 30	Hollywood Top 10	World Top 30
Mean	2%	2%	3%	0.13	0.24	0.04	0.09
	Mtime rate	Impression rate	Performance rate	Director rate	Picture rate	Music rate	Mtime rating #
Mean	6.12	6.27	6.34	6.02	6.44	6.18	5327
SD	(1.66)	(1.69)	(1.71)	(1.75)	(1.72)	(1.77)	(9796.09)
	Daily Screen #	Daily Audience #	Daily occupancy rate	Daily Price	Daily box office(Yuan)	Weekly box office(*10,000 Yuan)	
Mean	6.48	224.78	17%	59.54	14984	3713	
SD	(5.98)	(412.77)	(0.18)	(26.13)	(31408.17)	(24504.86)	
Variable	Search # on Baidu of ”movie's name”				Search # on Baidu of ”movie's name download”		
Time	1 day after Opening	1 day before Opening	1 week after Opening	S1 week before Opening	from TS available to 35 days after opening		
Mean	713	490	11993	2290		35533	
SD	(4246.33)	(2406.45)	(71030.92)	(10065.14)		(243435.12)	
Film Type	Family	Romance	Adventure	Fantasy	Comedy	Horror	Drama
Ratio	8.34%	22.94%	19.05%	11.85%	27.68%	20.85%	43.79%
	Science Friction	Cartoons	Crime	War	Documentary	Thriller	Children
Ratio	7.77%	9.95%	8.15%	4.74%	0.85%	3.98%	0.95%
Film Type	Erotic	Martial	Stage	Western	Dance	Noir	Biography
Ratio	0.09%	0.95%	0.19%	0.57%	0.95%	0.19%	2.27%
Film Type	Main stream	3D	Imax	Import	Co-production	Musical	Action
Ratio	5.02%	9.10%	4.17%	13.08%	21.99%	1.33%	32.51%
Film Type	Sports	Costume					
Ratio	1.14%	1.90%					
Producer	China Producer						
	CFGC	SFGC	Bona	Enlight	Huayi	US big8	HK big10
Ratio	10.24%	4.36%	5.02%	2.27%	3.41%	11.56%	14.03%
Distributor	CFGC	Huaxia	Huayi	Bona	Enlight		
Ratio	35.73%	25.31%	3.22%	7.96%	3.98%		

Notes: The Chinese director level and the Chinese star rank are from Forbes Celebrity rank. The foreign director rank is from Celebrity Networth. The foreign starring actor and actress level rank is from VULTURE. The author defines the director, actor and actress level dummies by these ranks. The movie type data is from Mtime. Some movies are coproduced by more than two firms. Season is defined by the Chinese central government public holiday schedule. The film types are given by Mtime. The US top 10 studios in the authors data are 20th Century Fox Studios, Warner Bros Studios, Walt Disney Studios, Sony Pictures Studios, Universal Studios, Paramount Studios, New Line Cinema, Dreamworks Studios, MGM Studios, and Raleigh Studios. The top 10 Hong Kong firms are Media Asia Films, Emperor Motion Pictures, China Star Entertainment Group, World Wide Pictures, Mei Ah Films Production, Milkyway Image, Jet Tone Film production, Golden Harvest Films, Mandarin Films and Shaw Brothers.

Table 20 Estimation results

	Mean	SD
# of the observation in the sample	793	
Box office(*10,000Yuan)	6293.76	(13586.4)
Mtime rate	6.23	(1.51)
Price	34.94	(15.03)
# of Douban rating	39328	(58441.45)
$(1-\alpha)/\lambda$	5.36	(0.09)
$\lambda c/\alpha$	9	(0.51)
$\lambda c/(1-\alpha)$	3.7	(0.44)
$\lambda\gamma/(1-\alpha)$	0.0007	(0.0008)
c(Outside option)	19.8	
γ(Unit waiting cost)	0.004	
λ(Exponential distribution parameter)	0.13	
α (Quality discount)	0.3	
Market size in 2013(by assumption)	20,000,000	
Predict box office	5996.685	(3932.1)
Predict price	33.5	(8.13)
Predict box office loss	29.92%	(0.17)
Predict box office loss if piracy movie available before opening	30.23%	

Notes: Box office, Mtime rate, price and # of Douban rating are from the subsample of the overall box office data. Price in the estimation is the average price of the first opening week. Only 793 movies have the information. The market size is given by assumption. The market size of each year is proportional to the overall screen number. The screen # is about 18,000 in 2013; 13,000 in 2012; 9,200 in 2011; 6,200 in 2010; 4,700 in 2009; 4,000 in 2008; 3,500 in 2007; 3,000 in 2006.

Table 21 Main Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Log(box office)								
Sample	Overall box office data						Weekly data	Daily data	
TS lag	0.000332	0.000444*	0.000506**						
	(0.000320)	(0.000255)	(0.000222)						
TS available before opening				-0.322***					
				(0.10					

				6)					
clear day lag					-0.000474** (0.000205)				
Opening lag						-0.000606*** (0.000235)			
TS available before the week							-0.168*** (0.0484)		
TS available before the day								-0.729*** (0.0134)	-0.966*** (0.0155)
TS available before the day *day #									0.0580*** (0.0019)
Control Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Season Dummies		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Producer & Distributor Dummies		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Director & Cast level		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Film Type Dummies		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Movie Rating			Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control Week # Dummies							Yes		
Control Day # and Weekday Dummies								Yes	Yes
Observations	1,039	1,039	1,039	1,039	1,039	1,039	3,321	286,125	286,125
R-squared	0.691	0.692	0.69	0.692	0.691	0.692	0.590	0.336	0.338

Notes: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1.

Table 22 2SLS Regression of "Name Download" Search Amount on Baidu Index

	(1)	(2)	(3)
	Log(box office)		Log(Baidu download Search #)
	2nd stage of 2SLS	OLS	1st stage of 2SLS
Log(Baidu download search # between TS day and the 35 day of showing)	-0.104***	-0.00359	
	(0.0352)	(0.00995)	
Log(Baidu download search # 1 day before TS available)			0.00311***
			(0.000505)
Log(Baidu download search 14 day before TS available)			-0.00235***
			(0.000397)
Control "Movie's name" search #	Yes	Yes	
Control other variables	Yes	Yes	Yes
Observations	1,039	1,039	1,039
R-squared	0.643	0.678	0.262
F test	F(2, 929) = 24.56, Prob ¿ F = 0.0000		
Over identification test	Sargan N*R-sq test 61.390 Chi-sq(2) P-value = 0.0000		
Endogeneity test	Basmann test 0.035 Chi-sq(2) P-value = 0.9827		

Notes: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The authors use the same control variables as the main result. The data sample is the overall box office data set

Table 23 DDD with Hong Kong Market

VARIABLES	(1)	(2)
	Log(box office)	
Control period Definition	From the 1st day to 6 months later	From the 1st day to 12 months later
TS available before opening*control period*mainland	0.648*	-0.0133
	(0.358)	(0.323)
TS available before opening	0.553**	0.176
	(0.194)	(0.216)
TS available before opening*mainland	-1.036***	-0.798***
	(0.233)	(0.215)
TS available before opening*control period	-0.704*	0.232
	(0.343)	(0.285)
Mainland	1.255***	1.128***
	(0.235)	(0.248)
Control period	-0.00128	-0.385*
	(0.175)	(0.218)
Mainland* Control period	-0.113	0.230
	(0.176)	(0.234)
Control other variables	Yes	Yes
Observations	1,519	1,519
R-squared	0.632	0.630

Notes: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. The opening lag is given by a different market's opening lag. The author uses public holiday dummy to replace season control in specification1. The authors add interaction with the year and the mainland dummy in the control variables. The three columns definitions are different. The column 1 and 4's periods are from the first day of the action to 1 month later after the action. Column 2 and 5's periods are s from the first day of the action to 6 month after the action. Column 3 and 6's periods are from the first day of the action to 12 months after the action. In columns 4-6, the authors have controlled all the dummies and variables. We just show the variable of interaction of three dummies' result. To deal with heteroskedasticity, the authors cluster the regressions by regions*year variables. The authors assigned a value to give each year in each region. The authors use the same control variables as the main result. The data sample is the overall box office data set and the Hong Kong box office data set.

Table 24 Regression of the Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Open week #	Log(Mtime rate #)	Log(Douban rate #)	Log(box office)		TS lag	Log(box office)	
Method	OLS				2SLS	OLS		
Sample	Overall box office data							Foreign movie
TS lag	-0.000693**	-6.39e-05	-0.000185	0.00168**	-0.000510**	-0.000251		-0.00101***
	(0.000310)	(0.000227)	(0.000235)	(0.000743)	(0.000208)	(0.000488)		(0.000264)
# of Movies released in US in the global premiere week						7.529*		
						(4.443)		
TS day lag *douban rate				-0.000222*				
				(0.000119)				
Log(overall screen #)							0.468*	
							(0.278)	
Control other variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,039	1,039	1,039	1,039	1,039	1039	1,039	362
R-squared	0.506	0.690	0.736	0.713	0.690	0.215	0.712	0.726

	(9)	(10)	(11)	(12)	(13)	(14)	(15)
VARIABLES	TS lag	Opening lag	Log(box office)		Log(screen #)	Log(price)	Log(box office)
Method	OLS						
Sample	Overall box office data		Daily box office data				Hong Kong Data
Opening lag	-0.622***						
	(0.03)						
Douban rate	-2.877	9.690***					
	(3.70)	(6.735)					
Movie is cut or not		69.11***					
		(19.32)					
TS available before the day			-0.172***		-0.237***	-0.0866***	
			(0.0057)		(0.00402)	(0.00562)	
TS available day # before the day				-0.00097***			
				(1.42e-05)			
TS lag							-0.000251
							(0.000488)
Control price and screen #			Yes				
Control other variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,039	1,039	286,125	286,125	286,125	286,125	467
R-squared	0.468	0.384	0.88	0.812	0.342	0.136	0.530

Notes: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. In the regressions using the overall box office data set and Hong Kong box office data set, the author uses the same control variables as the main result. Hong Kong box office is valued in the Hong Kong dollar. The mainland box office is valued in Yuan. In the regressions using daily data sample, the authors add day and weekday dummies in other control variables. In addition, the authors use the dummy whether or not the day belongs to the season replaced the season dummies in the control variables. Foreign movie means that the movie has no producers in China participating in the production process.

Table 25 Distributor and Producer's Influence

VARIABLES	(1)	(2)	(3)	(4)	(5)		(6)	(7)
	Public holiday			Opening lag			Hottest season	Public holiday
Methods	Probit			OLS			Probit	
	All movies				Foreign movies		All movies	
Producer						**Produce nations**		
Englight	0.632**		1.038***	6.362	-81.03	USA	-0.419**	-0.120
	(0.264)		(0.397)	(46.30)	(164.6)		(0.214)	(0.201)
Huayi	0.499**		0.597	14.04	3.051	Hong Kong	-0.194	-0.181
	(0.218)		(0.680)	(85.43)	(161.6)		(0.211)	(0.179)
CFGC	0.303**		0.302*	-23.81	71.20	Taiwan	-0.226	-0.539**
	(0.135)		(0.159)	(19.67)	(65.65)		(0.285)	(0.256)
US big10	-0.121		-0.131	-32.65	-13.48	Korea	-1.261**	-0.987**
	(0.140)		(0.180)	(20.39)	(27.61)		(0.620)	(0.472)
HK big10	0.238**		0.369**	-17.14	85.54	Japan	0.115	-0.199
	(0.120)		(0.154)	(17.95)	(78.44)		(0.322)	(0.294)
Distributor								
CFGC		0.0556	0.125	-10.77	-82.17***	Russia	0.00684	-0.788
		(0.0895)	(0.117)	(13.79)	(26.63)		(0.577)	(0.565)
Huaxia		-0.170*	0.0214	39.34***	-5.184	France	0.232	0.310
		(0.100)	(0.129)	(14.62)	(24.95)		(0.340)	(0.291)
Huayi		0.557**	0.0320	-48.71	-89.18	England	0.318	-0.00981
		(0.222)	(0.702)	(88.32)	(170.5)		(0.204)	(0.176)
Control other type and cast variable			Yes	Yes	Yes		Yes	Yes
Control quality vairiables			Yes	Yes	Yes		Yes	Yes
Constant	-0.734***	-0.641***	-0.738*	-277.0***	-139.9	Constant	0.473	0.751
	(0.0567)	(0.0631)	(0.406)	(51.87)	(118.4)		(2.334)	(1.937)
Observations	1,039	1,039	1,039	1,039	432	Observations	1,039	1,039

Notes: Standard errors in parentheses, *** p¡0.01, ** p¡0.05, * p¡0.1. In the 6th and 7th columns, the authors' control variables don't contain season dummies here. Public season means whether movies are distributed in public holiday. Hottest season means that whether the movies are distributed in National Day, Spring Festival. The data sample is the overall box office data.

APPENDIX B

PROOF

The Optimal Choices of Contracts I have three cases here.

The Autarky Case

In autarky case, the only state variable is saving. I get the optimal savings by first order conditions:

$$s_{t+1}^{i} : \frac{1}{c_t^i} = \beta \left(1 - \varrho_{t+1}^{i}\right) \frac{\partial V_{pt+1}^{aut}(s_{t+1}^{p} + s_{t+1}^{g})}{\partial s_{t+1}^{p}} \qquad (FOC1)$$

The Full-commitment Case

Using the envelope theorem, I get:

$$G^t : \frac{\partial P^{ft}(s_t^p, s_t^g, G^t)}{\partial G^t} = -\lambda_t \qquad (ET1)$$

The first order conditions for the optimal contract problem are:

$$T_t^p \quad : \quad \frac{c_{tjz}^{g}}{c_{tjz}^{p}} = \lambda_t \qquad (FOC2)$$

$$G_{jz}^{t+1} \quad : \quad \frac{\partial P^{f,t+1}(s_{t+1jz}^{p}, s_{t+1jz}^{g}, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}} = -\lambda_t \qquad (FOC3)$$

The Non-commitment Case

The first order conditions for the optimal contract problem are:

$$
\begin{aligned}
T_t^p &: \quad \frac{c_{tjz}^g}{c_{tjz}^p} = \frac{\lambda + \varkappa_{jz}}{1+\omega_{jz}} && (FOC4) \\
G_{jz}^{t+1} &: \quad \frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}} = -\frac{\lambda + \varkappa_{jz} + \theta_{jz}}{1+\omega_{jz}+\mu_{jz}} && (FOC5)
\end{aligned}
$$

$$
\begin{aligned}
s_{t+1}^p &: \quad \frac{1+\omega_{jz}}{c_{tjz}^p} \\
&= \beta\left(1+\omega_{jz}\right)[\left(1-\varrho_{t+7}\right) \frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^p} + \varrho_{t+7}\frac{\partial V_{pt+1}^{aut}(s_{t+1jz}^p + s_{t+1jz}^g)}{\partial s_{t+1jz}^p}] \\
&\quad +\mu_{jz}\beta\left(1-\varrho_{t+7}\right)\left(\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^p} - \frac{\partial V_{pt+1}^{aut}(s_{t+1jz}^p)}{\partial s_{t+1jz}^p}\right) && (FOC6)
\end{aligned}
$$

and

$$
\begin{aligned}
s_{t+1}^g &: \quad \frac{\lambda+\varkappa_{jz}}{c_{tjz}^g} + \theta_{jz}\beta\left(1-\varrho_{t+7}\right)\frac{\partial V_{gt+1}^{aut}(s_{t+1jz}^g)}{\partial s_{t+1jz}^g} \\
&= \beta\left(1+\omega_{jz}\right)[\left(1-\varrho_{t+7}\right) \frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^g} + \varrho_{t+7}\frac{\partial V_{pt+1}^{aut}(s_{t+1jz}^p + s_{t+1jz}^g)}{\partial s_{t+1jz}^g}] \\
&\quad +\mu_{jz}\beta\left(1-\varrho_{t+7}\right)\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^g} && (FOC7)
\end{aligned}
$$

Using the envelope theorem, I get:

$$
\begin{aligned}
G^t &: \quad \frac{\partial P^t(s_t^p, s_t^g, G^t)}{\partial G^t} = -\lambda && (ET2) \\
s_t^p &: \quad \frac{\partial P^t(s_t^p, s_t^g, G^t)}{\partial s_t^p} \\
&= \sum_{j=1}^{J}\sum_{z=1}^{Z} R_t \pi_{tj}^p \pi_{tz}^g \left(\frac{1+\omega_{jz}}{c_{tjz}^p} - \omega_{jz}\frac{\partial v_{pt+1}^{aut}(s_t^p, \epsilon_{jt}^p)}{\partial s_t^p}\right) && (ET3) \\
s_t^g &: \quad \frac{\partial P^t(s_t^p, s_t^g, G^t)}{\partial s_t^g} \\
&= \sum_{j=1}^{J}\sum_{z=1}^{Z} R_t \pi_{tj}^p \pi_{tz}^g \left(\frac{\lambda+\varkappa_{jz}}{c_{tjz}^g} - \omega_{jz}\frac{\partial v_{gt}^{aut}(s_t^g, \epsilon_{zt}^g)}{\partial s_t^g}\right) && (ET4)
\end{aligned}
$$

Proof of Proposition 1 **Proof.** The Lagrangian of the contract problems is:

$$L=\sum_{j=1}^{J}\sum_{z=1}^{Z}\pi_{tj}^{p}\pi_{tz}^{g}\{[U_{pt}(c_t^p,l_t^p,K_t^p,N_t^p)+\beta\left(1-\varrho_{t+9}\right)P^{f,t+1}(s_{t+1}^p,s_{t+1}^g,G_{jz}^{f,t+1})$$

$$+\beta\varrho_{t+9}V_{pt+1}^{aut}(s_{t+1}^p+s_{t+1}^g)]+\lambda_t\left[\pi_{tj}^{p}\pi U_{gt}(c_t^g,l_t^g,K_t^g,N_t^g)+\beta\left(1-\varrho_{t+9}\right)G_{jz}^{t+1}-G^t\right]\}$$

The first order conditions for the optimal contract problem are:

$$\begin{aligned}
T_t^p &: \frac{c_{jz}^g}{c_{jz}^p}=\lambda_t\\
G_{jz}^{t+1} &: \frac{\partial P^{f,t+1}(s_{t+1}^p,s_{t+1}^g,G_{jz}^{t+1})}{\partial G_{jz}^{t+1}}=-\lambda_t
\end{aligned}$$

Using the envelope theorem, I get:

$$G^t:\frac{\partial P^{f,t}(s_t^p,s_t^g,G^t)}{\partial G^t}=-\lambda_t$$

By the three equation, starting from period 1, I get

$$\frac{\partial P^{f,t+1}(s_{t+1}^p,s_{t+1}^g,G_{jz}^{t+1})}{\partial G_{jz}^{t+1}}=\frac{\partial P^{f,t}(s_t^p,s_t^g,G^t)}{\partial G^t}$$

and

$$\frac{c_{tjz}^g}{c_{tjz}^p}=\lambda_t$$

for $\forall t,\epsilon_z^g,\epsilon_j^p$. The results mean in every state of every period, the consumption ratio equal to λ_t. ■

Proof of Proposition 2 **Proof. Two Incentive Constraints are not binding at the same time.**

By contradiction, I assume two ICs are binding which means:

$$g_t\left(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p, G^t\right) = v_{gt}^{aut}(s_t^g, \epsilon_z^g),$$
$$p_t\left(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p, G^t\right) = v_{pt}^{aut}(s_t^p, \epsilon_j^p).$$

Exist an allocation with zero transfer and support, and all the decisions are same to autarky case, satisfies all the constraints and available in the contract. In this allocation, we have $g\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right) = v_{gt}^{aut}(s_t^g, \epsilon_z^g)$, and $V_{pt+1}^{aut}(s_{t+1jz}^p + s_{t+1jz}^g) > V_{pt+1}^{aut}(s_{t+1jz}^p)$, such that:

$$U_{pt}(c_{tjz}^p, l_{tjz}^p, K_{tjz}^p, N_{tjz}^p) + \beta\varrho_{t+7}V_{pt+1}^{aut}(s_{t+1jz}^p + s_{t+1jz}^g)$$
$$+\beta\left(1-\varrho_{t+7}\right)V_{pt+1}^{aut}(s_{t+1jz}^p) > v_{pt}^{aut}(s_t^p, \epsilon_{jt}^p).$$

The new contract gives a higher utility to parent without hurting grandparent. Contradict to the old contract is optimal. Two incentive constraints cannot be binding at the same time.

Two Participation Constraints are not binding at the same time.

By the conclusions above, with

$$g_t\left(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p, G^t\right) \geq v_{gt}^{aut}(s_t^g, \epsilon_{zt}^g),$$
$$p_t\left(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p, G^t\right) \geq v_{pt}^{aut}(s_t^p, \epsilon_{jt}^p),$$

and by the conclusion that two incentive constraints cannot be binding at the same time. The two participation constraints cannot be binding at the same time. The two equations

$$\sum_{j=1}^{J}\sum_{z=1}^{Z}\pi_{tj}^p\pi_{tz}^g g_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right) \geq \sum_{j=1}^{J}\sum_{z=1}^{Z}\pi_{tj}^p\pi_{tz}^g v_{gt}^{aut}(s_t^g, \epsilon_{zt}^g),$$

and

$$\sum_{j=1}^{J}\sum_{z=1}^{Z}\pi_{tj}^{p}\pi_{tz}^{g}p_t\left(s_t^g, s_t^p, \epsilon_z^g, \epsilon_j^p, G^t\right) \geq \sum_{j=1}^{J}\sum_{z=1}^{Z}\pi_{tj}^{p}\pi_{tz}^{g}v_{pt}^{aut}(s_t^p, \epsilon_j^p),$$

have at least one inequality holding. It means at least one participation is not binding. Two participation constraints cannot be bind at the same time ■

Proof of Proposition 3 **Proof.** Define the spending on child care $\pi_{it}^K = \sum_j k_t^j w_t^k$ and elder care $\pi_{it}^N = \sum_j n_t^j w_t^n$, with j as the household member. The value of time spending by individual i is $\pi_{it}^L = w_{it}^l l_t^j$, with $w_{it}^l = w_t^i$ if individual i spends on working or the last unit of elder/child care is from i; $w_{it}^l = w_t^j$, if i provides child care, but the last unit is provided by j; $w_{it}^l = p$, if i provides child care but the last unit is provided by the market. I define the overall endowment spending on period t's utility is:

$$\begin{aligned} E_t^i = R_t s_t^i + \sum_h w_{ht}^l + \epsilon_{lt}^i - s_{t+1}^i - T_t^i \\ + \sum_h \left(n_h^i w_t^n + k_h^i w_t^k\right) - \sum_l \left(n_l^i w_t^n + k_l^i w_t^k\right), \end{aligned} \tag{1}$$

in which h is the household member belonging to the household i, l is the household member not belonging to the household i. The first part, $R_t s_t^i + \sum_h w_{ht}^l + \epsilon_{lt}^i$ is the household endowment before transfer and support, s_{t+1}^i is the saving for the next period. T_t^i is the net transfer from household i to the other household. $\sum_h \left(n_h^i w_t^n + k_h^i w_t^k\right) - \sum_l \left(n_l^i w_t^n + k_l^i w_t^k\right)$ is the net child care and elder care support from household i to the other household. The overall household consumption π_{ir}^C is the household's overall spending on food, clothes, traffic, durable goods, utility, fuel, entertainment, education, beauty, and other consumption goods. From the theoretical model, I get:

$$\begin{cases} \pi_{ir}^{C} = \frac{E_t^i}{1+\eta+\alpha_t+\gamma_t} \\ \pi_{it}^{K} = \frac{E_t^i \alpha_t}{1+\eta+\alpha_t+\gamma_t} \\ \pi_{it}^{N} = \frac{E_t^i \gamma_t}{1+\eta+\alpha_t+\gamma_t} \\ \pi_{it}^{L} = \frac{E_t^i \eta}{1+\eta+\alpha_t+\gamma_t} \end{cases} \tag{2}$$

Define $A_t = (1+\eta_t+\alpha_t+\gamma_t)+\lambda\left(1+\eta_{t+8}+\alpha_{t+8}+\gamma_{t+8}\right)$, $\varpi = \frac{\lambda+\varkappa_{jz}}{1+\omega_{jz}}$, and $\kappa_t = \min\left\{w_t^g, w_t^p, p_t^o\right\}$.

The hour of elder care from market is:

$$n_t^o = \begin{cases} \frac{\gamma_{t+8}\varpi E_t}{A_t p_t}, if\ p_t < w_t^p \\ \frac{\left(\gamma_{t+8}\varpi+\eta_t\right)E_t}{A_t p_t} - 1, if\ p_t \geq w_t^p \& \frac{\left(\gamma_{t+8}\varpi+\eta_t\right)E_t}{A_t w_t^p} < 1 \\ 0, otherwise \end{cases}$$

The hour of elder care from parent is:

$$n_t^p = \begin{cases} \frac{\gamma_{t+8}\varpi E_t}{A_t w_t^p}, if\ p_t \geq w_t^p \& \frac{\left(\gamma_{t+8}\varpi+\eta_t\right)E_t}{A_t w_t^p} < 1 \\ \max\{0, 1-\frac{\eta_t E_t}{A_t w_t^p}\}, if\ p_t \geq w_t^p \& \frac{\left(\gamma_{t+8}\varpi+\eta_t\right)E_t}{A_t w_t^p} < 1 \\ 0, otherwise \end{cases}$$

The hour of child care from market is:

$$k_t^o = \begin{cases} \frac{\alpha_t E_t}{A_t p_t}, if\ p_t < w_t^p,\ p_t < w_t^g \\ \frac{\left(\alpha_t+\varpi\eta_{t+8}\right)E_t}{A_t p_t} - 1, if\ \frac{\left(\alpha_t+\lambda\eta_{t+8}\right)E_t}{A_t w_t^g} \geq 1 \& w_t^p \geq p_t \geq w_t^g \& \frac{\left(\alpha_t+\varpi\eta_{t+8}+\eta_t\right)E_t}{A_t w_t^g} < 2 \\ \frac{\left(\alpha_t+\eta_t\right)E_t}{A_t p_t} - 1, if\ \frac{\left(\alpha_t+\eta_t\right)E_t}{A_t w_t^p} \geq 1 \& w_t^g \geq p_t \geq w_t^p \& \frac{\left(\alpha_t+\varpi\eta_{t+8}+\eta_t\right)E_t}{A_t w_t^p} < 2 \\ \frac{\left(\alpha_t+\varpi\eta_{t+8}+\eta_t\right)E_t}{A_t p_t} - 2, if\ \frac{\left(\alpha_t+\varpi\eta_{t+8}+\eta_t\right)E_t}{A_t w_t^g} \geq 2 \& p_t = \max\left\{p_t, w_t^g, w_t^p\right\} \\ \frac{E_t\left(\varpi\eta_{t+8}+\alpha_t\right)}{A_t p_t} - 1, if\ \frac{E_t}{A_t w_t^g}\left(\varpi\eta_{t+8}+\alpha_t\right) > 1 \& p_t < w_t^p \& p_t \geq w_t^g \\ 0, otherwise \end{cases}$$

The hour of child care from parent is:

$$k_t^p = \begin{cases} \frac{\alpha_t E_t}{A_t w_t^p}, if \ \frac{(\alpha_t+\eta_t)E_t}{A_t w_t^p} < 1 \& p_t \geq w_t^g \geq w_t^p \\ \max\left\{0, 1-\frac{\eta_t E_t}{A w_{tt}^p}\right\}, if \ \frac{(\alpha_t+\eta_t)E_t}{A_t w_t^p} \geq 1 \& p_t \geq w_t^g \geq w_t^p \\ \max\left\{0, \frac{(\alpha_t+\varpi\eta_{t+8})E_t}{A_t w_t^p} - 1\right\}, if \ \frac{(\alpha_t+\varpi\eta_t)E_t}{A_t w_t^g} \geq 1 \& p_t \geq w_t^p \geq w_t^g \& \frac{(\alpha_t+\varpi\eta_{t+8}+\eta_t)E_t}{A_t w_t^g} < 2 \\ 0, otherwise \end{cases}$$

The hour of child care from grandparent is:

$$k_t^g = \begin{cases} \frac{\alpha_t E_t}{A_t w_t^g}, if \ \frac{(\alpha_t+\varpi\eta_{t+7})E_t}{A_t w_t^g} < 1 \& p_t \geq w_t^p \geq w_t^g \\ 1-\frac{\varpi\eta_{t+7}E_t}{A_t w_t^g}, if \ \frac{(\alpha_t+\varpi\eta_{t+7})E_t}{A_t w_t^g} \geq 1 \& p_t \geq w_t^p \geq w_t^g \\ \frac{(\alpha_t+\eta_t)E_t}{A_t w_t^p} - 1, if \ \frac{(\alpha_t+\eta_t)E_t}{A_t w_t^g} \geq 1 \& p_t \geq w_t^g \geq w_t^p \& \frac{(\alpha_t+\varpi\eta_{t+7}+\eta_t)E_t}{A_t w_t^p} < 2 \\ 0, otherwise \end{cases}$$

The leisure of the parent is:

$$l_t^p = \begin{cases} 1, if \ \frac{\eta_t E_t}{A_t \kappa_t} \geq 1 \\ \frac{\eta_t E_t}{A_t \kappa_t}, otherwise \end{cases}$$

The leisure of the grandparent is:

$$l_t^g = \begin{cases} 1, if \ \frac{\varpi\eta_{t+8}E_t}{A_t \kappa_t} \geq 1 \\ \frac{\varpi\eta_{t+8}E_t}{A_t \kappa_t}, otherwise \end{cases}$$

If $p_t = \min\{w_t^p, w_t^g, p_t\}, n_t^p = k_t^p = k_t^g = 0$; if $w_t^p = \min\{w_t^p, w_t^g, p_t\}$ and $\gamma_{t+7} > 0$ ($\alpha_t > 0$), $n_t^p > 0$ ($k_t^p > 0$); if $w_t^g = \min\{w_t^p, w_t^g, p_t\}$ and $\gamma_{t+7} > 0$ ($\alpha_t > 0$), $n_t^p > 0$ ($k_t^p > 0$). ■

Lemma 1

Lemma 14 *With more overall saving, at least one household gets a higher ex-ante utility value. If* $s^p_{t+1jn} + s^g_{t+1jn} < s^p_{t+1jm} + s^g_{t+1jm}$*, then*

$$G^{t+1}_{jn} \geq G^{t+1}_{jm},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) \geq P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}),$$

cannot happen at the same time.

Proof. Toward contradiction, assume in an optimal contract, exists two states such that:

$$s^p_{t+1jn} + s^g_{t+1jn} < s^p_{t+1jm} + s^g_{t+1jm},$$

$$G^{t+1}_{jn} \geq G^{t+1}_{jm},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) \geq P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}).$$

The consumption choice and saving choice with saving s^p_{t+1jn} and s^g_{t+1jn} are still available at with saving s^p_{t+1jm} and s^g_{t+1jm}. Setting, $s^{p'}_{t+1jm} = s^p_{t+1jm} + s^g_{t+1jm} - s^g_{t+1jn}$, $s^{g'}_{t+1jm} = s^g_{t+1jn}$, $G^{'t+1}_{jm} = G^{t+1}_{jn}$, without changing any variables on time t, I have $P^{t+1}(s^{p'}_{t+1jn}, s^{g'}_{t+1jn}, G^{t+1'}_{jn}) > P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})$. The new PCs hold, because

$$G^{t+1}_{jn} = G^{t+1'}_{jm} \geq V^{aut}_{gt+1}(s^{g'}_{t+1jm}) = V^{aut}_{gt+1}(s^g_{t+1jn}).$$

Without changing current consumption, new ICs also hold. In the interior region that IC is binding, fixed ϵ^g_{zt}, as ϵ^p_{jt} increase or decrease, I get

$$p\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt}(s^p_t, \epsilon^p_{jt}),$$

and

$$p\left(s^g_t, s^p_t, \epsilon^g_{zt} + \epsilon, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt}(s^p_t, \epsilon^p_{jt} + \epsilon),$$

which means

$$\frac{1}{c^{paut}_{tjz}} = \frac{\partial v^{aut}_{pt}(s^p_t, \epsilon^p_{jt})}{\partial \epsilon^p_{jt}} = \frac{\partial p\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)}{\partial \epsilon^p_{jt}} = \frac{1}{c^p_{tjz}},$$

and

$$\frac{1}{c^{paut}_{tjz}} = \frac{\partial v^{aut}_{pt}(s^p_t, \epsilon^p_{jt})}{\partial s^p_t} = \frac{\partial p\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)}{\partial s^p_t} = \frac{1}{c^p_{tjz}}.$$

The marginal utility on saving of contract and autarky are equal, when IC is binding. It means when a IC with saving s^p_{t+1jn} and s^g_{t+1jn} is binding, the contract value are still equal to the autarky value as saving increase and new IC is binding. When a IC with saving s^p_{t+1jn} and s^g_{t+1jn} is not binding, contract value may increase less than autarky value. But when the two are equal, the contract value and autarky value will keep equal as saving increase. The new IC is still holding. All the new ICs in period t are holding, I get $P^{t+1}(s^{p'}_{t+1jn}, s^{g'}_{t+1jn}, G^{t+1'}_{jn}) \geq V^{aut}_{pt+1}(s^{p'}_{t+1jm})$. In addition, the new ICs are satisfied because the utility values at period t are unchanged and the utility values in period t+1 increased. The new allocation is feasible and will increase two households' utility. Contradict to the contract is optimal.

If $s^p_{t+1jn} + s^g_{t+1jn} < s^p_{t+1jm} + s^g_{t+1jm}$, I cannot get:

$$G^{t+1}_{jn} \leq G^{t+1}_{jm},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) \leq P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}),$$

at the same time. In an optimal contract, if more overall saving means at least one household is better off ex-ante. ■

<u>Lemma 2</u>

Lemma 15 *With a bigger income shock, at least one household gets a higher ex-post utility value. If $\epsilon^g_{mt} > \epsilon^g_{nt}$, then*

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

cannot happen at the same time. If $\epsilon^p_{mt} > \epsilon^p_{nt}$, then

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{mt}, G^t\right) \leq p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{nt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{mt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{nt}, G^t\right),$$

cannot happen at the same time.

Proof. Fixed ϵ^p_{jt}, for $\forall\ \epsilon^g_{mt} > \epsilon^g_{nt}$,

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

cannot happen at same time. If $\epsilon^g_{mt} > \epsilon^g_{nt}$, all allocation in state n is also available on state m. If, I have:

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

by changing contract choice in state m to the contract choice to state n, will increase at least one household's utility value, without hurting the other. In addition, the new ICp and PCs are unchanged and the new ICg holds, because

$$v^{aut}_{gt}(s^g_t, \epsilon^g_{nt}) \leq v^{aut}_{gt}(s^g_t, \epsilon^g_{mt}) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right).$$

The new allocation is feasible and will increase two households' utility. Contradiction to the contract is optimal.

Use the same method, I can prove that fixed ϵ^g_z, for $\forall\ \epsilon^p_{mt} > \epsilon^p_{nt}$,

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{mt}, G^t\right) \leq p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{nt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{mt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{nt}, G^t\right),$$

cannot happen at the same time. ■

Lemma 3

Lemma 16 *When one's incentive constraint is binding, the marginal utility of autarky value and contract values on saving are the same, and the marginal utility on saving and income shock are equal in both autarky case and contracts. If parent's incentive constraint is binding, then*

$$\frac{\partial v_{pt}^{aut}(s_t^p,\epsilon_{jt}^p)}{\partial s_t^p}=\frac{\partial v_{pt}^{aut}(s_t^p,\epsilon_{jt}^p)}{\partial \epsilon_{jt}^p}=\frac{\partial p_t\left(s_t^g,s_t^p,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)}{\partial \epsilon_{jt}^p}=\frac{\partial p_t\left(s_t^g,s_t^p,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)}{\partial s_t^p}.$$

If grandparent's incentive constraint is binding, then

$$\frac{\partial v_{gt}^{aut}(s_t^g,\epsilon_{zt}^g)}{\partial \epsilon_{zt}^g}=\frac{\partial v_{gt}^{aut}(s_t^g,\epsilon_{zt}^g)}{\partial s_t^g}=\frac{\partial g_t\left(s_t^g,s_t^p,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)}{\partial \epsilon_{zt}^g}=\frac{\partial g_t\left(s_t^g,s_t^p,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)}{\partial s_t^g}.$$

In the interior region that IC is binding, fixed ϵ_{zt}^g, as ϵ_{jt}^p increase or decrease,

$$p_t\left(s_t^g,s_t^p,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)=v_{pt}^{aut}(s_t^p,\epsilon_{jt}^p),$$

and

$$p_t\left(s_t^g,s_t^p+\epsilon,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)=v_{pt}^{aut}(s_t^p,\epsilon_{jt}^p+\epsilon).$$

I get

$$\frac{\partial v_{pt}^{aut}(s_t^p,\epsilon_{jt}^p)}{\partial \epsilon_{jt}^p}=\frac{\partial p_t\left(s_t^g,s_t^p,\epsilon_{zt}^g,\epsilon_{jt}^p,G^t\right)}{\partial \epsilon_{jt}^p},$$

then

$$\frac{1}{c_{tjn}^{paut}}=\frac{1}{c_{tjm}^{paut}}=\frac{1}{c_{tjm}^{p}}=\frac{1}{c_{tjmn}^{p}}.$$

The marginal utilities on saving of contract and autarky case are equal, when IC is binding. It means when a IC is binding, the contract value is still equal to the autarky value as saving increasing and new IC is binding. When we increase or decrease saving in the interior region that IC is binding,

$$p_t\left(s_t^g, s_t^p+\xi, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right) = v_{pt}^{aut}(s_t^p+\xi, \epsilon_{lt}^p),$$

with ξ is number small enough. In the interior region that IC is not binding,

$$p_t\left(s_t^g, s_t^p+\xi, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right) > v_{pt}^{aut}(s_t^p+\xi, \epsilon_{lt}^p).$$

A marginal value ϵ_{lt}^p between such that,

$$p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right) = v_{pt}^{aut}(s_t^p, \epsilon_{lt}^p).$$

ξ is number small enough, such that

$$p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{lt}^p-\xi, G^t\right) > v_{pt}^{aut}(s_t^p, \epsilon_{lt}^p-\xi).$$

I get

$$\begin{aligned} & p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right) - p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{lt}^p-\xi, G^t\right) \\ > \quad & v_{pt}^{aut}(s_t^p, \epsilon_{lt}^p) - v_{pt}^{aut}(s_t^p, \epsilon_{lt}^p-\xi), \end{aligned}$$

which means

$$\frac{\partial p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{lt}^p-\xi, G^t\right)}{\partial\left(\epsilon_{lt}^p-\xi\right)} < \frac{\partial v_{pt}^{aut}(s_t^p, \epsilon_{lt}^p-\xi)}{\partial\left(\epsilon_{lt}^p-\xi\right)}.$$

Because

$$\frac{\partial v_{pt}^{aut}(s_t^p, \epsilon_{jt}^p)}{\partial \epsilon_{jt}^p} = \frac{\partial v_{pt}^{aut}(s_t^p, \epsilon_{jt}^p)}{\partial s_t^p},$$

and

$$\frac{\partial p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)}{\partial \epsilon_{jt}^p} = \frac{\partial p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)}{\partial s_t^p},$$

as saving increase. As ϵ_{lt}^p increases by ξ, with

$$\frac{\partial p_t\left(s_t^g, s_t^p + \xi, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right)}{\partial s_t^p} = \frac{\partial v_{pt}^{aut}(s_t^p + \xi, \epsilon_{lt}^p)}{\partial s_t^p},$$

I get:

$$p_t\left(s_t^g, s_t^p + \xi, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right) = v_{pt}^{aut}(s_t^p + \xi, \epsilon_{lt}^p).$$

As ϵ_{lt}^p decrease by ξ, with

$$\frac{\partial p_t\left(s_t^g, s_t^p - \xi, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right)}{\partial s_t^p} < \frac{\partial v_{pt}^{aut}(s_t^p - \xi, \epsilon_{lt}^p)}{\partial s_t^p},$$

I get

$$p_t\left(s_t^g, s_t^p - \xi, \epsilon_{zt}^g, \epsilon_{lt}^p, G^t\right) > v_{pt}^{aut}(s_t^p - \xi, \epsilon_{lt}^p).$$

Increase or decrease saving and consumption by the same amount will keep the new IC holding.

Proof of Proposition 4 **Proof.** Toward contradiction assume $-\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}} < \frac{c_{tjz}^g}{c_{tjz}^p}$.

By the first order condition,

$$\frac{\lambda + \varkappa_{jz} + \theta_{jz}}{1 + \omega_{jz} + \mu_{jz}} = -\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}},$$

and

$$\frac{c_{tjz}^g}{c_{tjz}^p} = \frac{\lambda + \varkappa_{jz}}{1 + \omega_{jz}},$$

I get

$$\frac{\lambda + \varkappa_{jz} + \theta_{jz}}{1 + \omega_{jz} + \mu_{jz}} < \frac{\lambda + \varkappa_{jz}}{1 + \omega_{jz}}.$$

It means $\mu_{jz} > 0$, $\theta_{jz} = 0$. Parent's participation constraint is binding and grandparent's participation constraint is not binding, such that:

$$G_{jz}^{t+1} \geq V_{gt+1}^{aut}(s_{t+1jz}^{g}),$$

$$P^{t+1}(s_{t+1jz}^{p}, s_{t+1jz}^{g}, G_{jz}^{t+1}) \geq V_{pt+1}^{aut}(s_{t+1jz}^{p}).$$

In addition, all parent's incentive constraints at time t+1 are binding. Otherwise, each contract value is no less than autarky value. As one incentive constraint not binding, the ex-ante contract value $P^{t+1}(s_{t+1jz}^{p}, s_{t+1jz}^{g}, G_{jz}^{t+1})$ must be larger than autarky case, contradict to participation constraint binding. If parent's incentive constraint is binding, by Propostion 2, grandparent's parent's incentive constraint is not binding and with the Lagrangian multiplier equal to zero.

$$\frac{c_{t+1nm}^{g}}{c_{t+1nm}^{p}} = \frac{\lambda + \varkappa_{jz} + \theta_{jz}}{1 + \omega_{jz} + \mu_{jz} + \omega_{t+1nm}} \leq \frac{\lambda + \varkappa_{jz} + \theta_{jz}}{1 + \omega_{jz} + \mu_{jz}}.$$

Exist an affordable allocation to improve both households' utilities without violating all the constraints. $\exists\, \varepsilon$, ι and ζ, all are number small enough with $c_{t,jz}^{g*} = c_{t,jz}^{g} - \varepsilon$, $c_{t,jz}^{p*} = c_{t,jz}^{p} + \varepsilon$, $s_{t+1jz}^{g*} = s_{t+1jz}^{g} + \zeta$, $s_{t+1jz}^{p*} = s_{t+1jz}^{p} - \zeta$, $c_{t+1nm}^{g*} = c_{t+1nm}^{g} + \zeta$, and $c_{t+1nm}^{p*} = c_{t+1nm}^{p} - \zeta$ for any m and n. Leave other variables unchanged, such that:

$$\begin{aligned} & \sum_{m=1}^{J}\sum_{n=1}^{Z}\pi_{t+1n}^{p}\pi_{t+1m}^{g}\left(\ln\left(c_{t+1nm}^{g} + \zeta\right) - \ln\left(c_{t+1nm}^{g}\right)\right) \\ = \; & \ln\left(c_{t,jz}^{g}\right) - \ln\left(c_{t,jz}^{g} - \varepsilon\right). \end{aligned}$$

Because $\frac{\lambda+\varkappa_{jz}+\theta_{jz}}{1+\omega_{jz}+\mu_{jz}} > \frac{\lambda+\varkappa_{jz}}{1+\omega_{jz}}$,and ε and ζ are small enough,

$$\begin{aligned} & \frac{1+\omega_{jz}}{\lambda+\varkappa_{jz}}\frac{\varepsilon}{c^g_{t,jz}} \\ > \quad & \beta\left(1-\varrho_{t+9}\right)R_t\frac{1+\omega_{jz}+\mu_{jz}}{\lambda+\varkappa_{jz}+\theta_{jz}}\zeta\sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^g_{t+1nm}}. \end{aligned}$$

In addition, ε and ζ are small enough, I get:

$$\ln\left(c^p_{t,jz}+\varepsilon\right)-\ln\left(c^p_{t,jz}\right)=\frac{1+\omega_{jz}}{\lambda+\varkappa_{jz}}\frac{\varepsilon}{c^g_{t,jz}},$$

and

$$\begin{aligned} & \sum_{m=1}^{J}\sum_{n=1}^{Z}\pi^p_{t+1n}\pi^g_{t+1m}\left[\ln\left(c^p_{t+1nm}\right)-\ln\left(c^p_{t+1nm}-\zeta\right)\right] \\ & \qquad \leq \frac{1+\omega_{jz}+\mu_{jz}}{\lambda+\varkappa_{jz}+\theta_{jz}}\zeta\sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^g_{t+1nm}}. \end{aligned}$$

which means:

$$\begin{aligned} & \ln\left(c^p_{t,jz}+\varepsilon\right)-\ln\left(c^p_{t,jz}\right) \\ > \quad & \beta\left(1-\varrho_{t+9}\right)R_t\sum_{m=1}^{J}\sum_{n=1}^{Z}\pi^p_{t+1n}\pi^g_{t+1m}\left[\ln\left(c^p_{t+1nm}\right)-\ln\left(c^p_{t+1nm}-\zeta\right)\right]. \end{aligned}$$

Setting s^{p*}_{t+1jz} such that $P^{t+1}(s^{p*}_{t+1jz}, s^{g*}_{t+1jz}, G^{t+1*}_{jz}) \geq V^{aut}_{pt+1}(s^{p*}_{t+1jz})$. The new allocation satisfies all incentive constraints, because no one is worse off than the old allocation. $G^{t+1*}_{jz} > G^{t+1}_{jz} \geq V^{aut}_{gt+1}(s^{g*}_{t+1jz})$, because ζ is a number small enough. All the PCs are satisfied. By Lemma 3, increasing or decreasing saving and consumption by a same amount, the contract value and autarky value will keep equal as saving increase, and all the new ICs and PCs are still holding.

In addition, because the overall savings and overall consumption doesn't change, the new allocation is feasible. The new allocation gives parent higher utility without hurting grandparent. Contradiction to the old allocation is from optimal contract. $-\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}} < \frac{c^g_{tjz}}{c^p_{tjz}}$ cannot happen. Use the same method, I can prove $-\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}} > \frac{c^g_{tjz}}{c^p_{tjz}}$ cannot happen.

I then want to prove $\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}} = \frac{\sum_{m=1}^{J}\sum_{n=1}^{Z} \pi^p_{t+1n}\pi^g_{t+1m}/c^p_{t+1mn}}{\sum_{m=1}^{J}\sum_{n=1}^{Z} \pi^p_{t+1n}\pi^g_{t+1m}/c^g_{t+1mn}}$.

By the first order conditions, I get:

$$\begin{aligned} & \frac{1+\omega_{jz}}{c^p_{tjz}} \\ = \; & \beta\left(1+\omega_{jz}\right)\left[\left(1-\varrho_{t+7}\right)\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^p_{t+1jz}}\right. \\ & \left. + \varrho_{t+7}\frac{\partial V^{aut}_{pt+1}(s^p_{t+1jz}+s^g_{t+1jz})}{\partial s^p_{t+1jz}}\right]. \end{aligned} \tag{3}$$

and

$$\begin{aligned} & \frac{\lambda+\varkappa_{jz}}{c^g_{tjz}} \\ = \; & \beta\left(1+\omega_{jz}\right)\left[\left(1-\varrho_{t+7}\right)\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^g_{t+1jz}}\right. \\ & \left. + \varrho_{t+7}\frac{\partial V^{aut}_{pt+1}(s^p_{t+1jz}+s^g_{t+1jz})}{\partial s^g_{t+1jz}}\right]. \end{aligned} \tag{4}$$

By

$$\frac{1+\omega_{jz}}{c^p_{tjz}} = \frac{\lambda+\varkappa_{jz}}{c^g_{tjz}},$$

and

$$\frac{\partial V^{aut}_{pt+1}(s^p_{t+1jz}+s^g_{t+1jz})}{\partial s^p_{t+1jz}} = \frac{\partial V^{aut}_{pt+1}(s^p_{t+1jz}+s^g_{t+1jz})}{\partial s^g_{t+1jz}},$$

I get:

$$\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^p_{t+1jz}} = \frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^g_{t+1jz}}. \tag{5}$$

In addition, by Envelope theorem, I get:

$$\begin{aligned} & \frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^p_{t+1jz}} \\ = & \sum_{m=1}^{J}\sum_{n=1}^{Z}\pi^p_{t+1n}\pi^g_{t+1m}\left[\begin{array}{c} \frac{1}{c^p_{t+1mn}} \\ +\omega_{t+1mn}\left(\frac{1}{c^p_{t+1mn}} - \frac{\partial v^{aut}_{pt+1}(s^p_{t+1jz},\epsilon^p_{t+1j})}{\partial s^p_{t+1jz}}\right) \end{array}\right]. \end{aligned} \tag{6}$$

Two cases exist: $\omega_{t+1jz} = 0$ or $\omega_{t+1jz} > 0$. If $\omega_{t+1mn} > 0$, then

$$p_t\left(s^p_{t+1jz}, s^g_{t+1jz}, \epsilon^g_{zt}, \epsilon^p_{jt}, G^{t+1}_{jz}\right) = v^{aut}_{pt+1}(s^p_{t+1jz}, \epsilon^p_{jt+1}),$$

and

$$\frac{1}{c^p_{t+1mn}} = \frac{\partial v^{aut}_{pt+1}(s^p_{t+1jz}, \epsilon^p_{jt+1})}{\partial s^p_{t+1jz}}.$$

Both cases have

$$\omega_{t+1mn}\left(\frac{1}{c^p_{t+1mn}} - \frac{\partial v^{aut}_{pt+1}(s^p_{t+1jz}, \epsilon^p_{jt+1})}{\partial s^p_{t+1jz}}\right) = 0,$$

and

$$\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^p_{t+1jz}} = \sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^p_{t+1mn}}.$$

I get:

$$\begin{aligned} & \frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^g_{t+1jz}} \\ = & \sum_{m=1}^{J}\sum_{n=1}^{Z}\pi^p_{t+1n}\pi^g_{t+1m}\left[\begin{array}{c} \frac{\lambda+\varkappa_{jz}}{1+\omega_{jz}}\frac{1}{c^g_{t+1mn}} \\ +\chi_{t+1mn}\left(\frac{1}{c^g_{t+1mn}} - \frac{\partial v^{aut}_{gt+1}(s^g_{t+1jz},\epsilon^g_{zt+1})}{\partial s^g_{t+1jz}}\right) \end{array}\right]. \end{aligned} \tag{7}$$

Two cases exist: $\chi_{t+1jz} = 0$ or $\chi_{t+1jz} > 0$. If $\chi_{t+1mn} > 0$, then

$$g\left(s^p_{t+1jz}, s^g_{t+1jz}, \epsilon^g_z, \epsilon^p_j, G^{t+1}_{jz}\right) = v^{aut}_{gt+1}(s^g_{t+1jz}, \epsilon^g_{t+1z}),$$

and

$$\frac{1}{c^g_{t+1jz}} = \frac{\partial v^{aut}_{gt+1}(s^g_{t+1jz}, \epsilon^g_{zt+1})}{\partial s^g_{t+1jz}}.$$

Both cases have

$$\chi_{t+1mn}\left(\frac{1}{c^g_{t+1mn}} - \frac{\partial v^{aut}_{gt+1}(s^g_{t+1jz}, \epsilon^g_{zt+1})}{\partial s^g_{t+1jz}}\right) = 0,$$

and

$$\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial s^g_{t+1jz}} = \sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\lambda + \varkappa_{jz}}{1+\omega_{jz}}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^g_{t+1mn}}.$$

In addition, by

$$\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}} = \frac{\lambda + \varkappa_{jz}}{1+\omega_{jz}},$$

and

$$\frac{c^g_{tjz}}{c^p_{tjz}} = \frac{\lambda + \varkappa_{jz}}{1+\omega_{jz}},$$

I get:

$$\begin{aligned}\frac{c^g_{tjz}}{c^p_{tjz}} &= \frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}} \\ &= \sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^p_{t+1mn}} \Big/ \sum_{m=1}^{J}\sum_{n=1}^{Z}\frac{\pi^p_{t+1n}\pi^g_{t+1m}}{c^g_{t+1mn}}.\end{aligned} \tag{8}$$

■

Corollary 1

Corollary 17 *Before the last period, the Lagrangian multipliers associated with participation constraints always equal to 0. For* $\forall\ \epsilon^g_{zt}$ *and* ϵ^p_{jt}, $\frac{\lambda_t+\varkappa_{jz}+\theta_{jz}}{1+\omega_{jz}+\mu_{jz}} = \frac{\lambda_t+\varkappa_{jz}}{1+\omega_{jz}}$. *If* $t<T$, $\theta_{jz} = \mu_{jz} = 0$.

Proof. From Proposition 4, I get $\frac{c^g_{tjz}}{c^p_{tjz}} = -\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}}$, which means $\frac{\lambda+\varkappa_{jz}}{1+\omega_{jz}} = \frac{\lambda+\varkappa_{jz}+\theta_{jz}}{1+\omega_{jz}+\mu_{jz}}$.

By Lemma 1 at most one PC is binding before period T, if $\theta_{jz} > 0$, then $\mu_{jz} = 0$; if $\mu_{jz} > 0$, then $\theta_{jz} = 0$. Only when $\theta_{jz} = \mu_{jz} = 0$, I can get $\frac{\lambda+\varkappa_{jz}+\theta_{jz}}{1+\omega_{jz}+\mu_{jz}} = \frac{\lambda+\varkappa_{jz}}{1+\omega_{jz}}$. ■

Lemma 4

Lemma 18 *Fixed the consumption ratio, with a bigger income shock, both households will get bigger ex-post utility values. Fixed* $\frac{c^g_{tjm}}{c^p_{tjm}} = \frac{c^g_{tjn}}{c^p_{tjn}} = \lambda_t$, *if* $\epsilon^g_{mt} > \epsilon^g_{nt}$, *then*

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) > p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) > g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right),$$

for $\forall \epsilon^p_j$ *; if* $\epsilon^p_{kt} > \epsilon^p_{lt}$, *then*

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{kt}, G^t\right) > p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{lt}, G^t\right),$$

and

$$g_t\left(s^g_t, s^p_t, \epsilon^g_z, \epsilon^p_k, G^t\right) > g_t\left(s^g_t, s^p_t, \epsilon^g_z, \epsilon^p_l, G^t\right).$$

for $\forall\ \epsilon_z^g$.

Proof. Toward contradiction, assume $\epsilon_{mt}^g < \epsilon_{nt}^g$, and $p_t\left(s_t^g, s_t^p, \epsilon_{nt}^g, \epsilon_{jt}^p, G^t\right) \leq p_t\left(s_t^g, s_t^p, \epsilon_{mt}^g, \epsilon_{jt}^p, G^t\right)$.

By Lemma 2, I get $g_t\left(s_t^g, s_t^p, \epsilon_{nt}^g, \epsilon_{jt}^p, G^t\right) > g_t\left(s_t^g, s_t^p, \epsilon_{mt}^g, \epsilon_{jt}^p, G^t\right)$. By equation 5, I get:

$$\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^p} = \frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^g}.$$

Define with $\Omega_{tjz} = \left\{n_{tjz}^p, k_{tjz}^g, T_{tjz}^p, s_{t+1jz}^g, h_{tjz}^g, n_{tjz}^m, s_{t+1jz}^p, h_{tjz}^p, k_{tjz}^p, k_{tjz}^m\right\}$. By Proposition 4, the problem is equivalent to solveing:

$$\max_{\Omega_{tjz}} p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right) + \lambda g_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right).$$

Subject to promise constraints and participation constraints. If λ is fixed, as $s_{t+1jz}^p + s_{t+1jz}^g$ increase, by Lemma 2, both $P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})$ and G_{jz}^{t+1} will increase. In addition, with

$$\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^g} = \lambda \frac{\partial G_{jz}^{t+1}}{\partial s_{t+1jz}^g},$$

and

$$\frac{\partial P^{t+1}(s_{t+1jz}^p, s_{t+1jz}^g, G_{jz}^{t+1})}{\partial s_{t+1jz}^p} = \lambda \frac{\partial G_{jz}^{t+1}}{\partial s_{t+1jz}^p},$$

as λ fixed, I get:

$$\lambda \frac{\partial G_{jz}^{t+1}}{\partial s_{t+1jz}^g} = \frac{\partial G_{jz}^{t+1}}{\partial s_{t+1jz}^p}.$$

If $\frac{\partial P^{t+1}(s_{t+1jn}^p, s_{t+1jn}^g, G_{jz}^{t+1})}{\partial s_{t+1jn}^p} > \frac{\partial P^{t+1}(s_{t+1jm}^p, s_{t+1jm}^g, G_{jm}^{t+1})}{\partial s_{t+1jn}^p}$, I get

$$G_{jm}^{t+1} < G_{jn}^{t+1},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) < P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}).$$

There are two cases.

Case 1 when $\mathbf{s}^p_{t+1jn}$+$\mathbf{s}^g_{t+1jn}$≤ $\mathbf{s}^p_{t+1jm}$+$\mathbf{s}^g_{t+1jm}$

First, I get $\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jm}, G^{t+1}_{jz})}{\partial s^g_{t+1jn}} < \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^g_{t+1jn}}$.

If I have:

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jz})}{\partial s^g_{t+1jn}} \geq \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^g_{t+1jn}},$$

and

$$s^p_{t+1jn} + s^g_{t+1jn} \leq s^p_{t+1jm} + s^g_{t+1jm}.$$

I can get $\frac{\lambda}{c^g_{t,jm}} \leq \frac{\lambda}{c^g_{t,jn}}$.

By Proposition 4, and $\frac{\partial P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})}{\partial G^{t+1}_{jz}} = \lambda$, as $P^{t+1}(s^p_{t+1jz}, s^g_{t+1jz}, G^{t+1}_{jz})$ increase or decrease, G^{t+1}_{jz} should also increase or decrease. With $s^p_{t+1jn} + s^g_{t+1jn} \leq s^p_{t+1jm} + s^g_{t+1jm}$, I get:

$$G^{t+1}_{jm} \leq G^{t+1}_{jn},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) \leq P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}).$$

However with $s^p_{t+1jn} + s^g_{t+1jn} \leq s^p_{t+1jm} + s^g_{t+1jm}$, without changing the decisions variable on period t, changing the saving decisions on state m s^p_{t+1jm} and s^g_{t+1jm} equal to s^p_{t+1jn} and s^g_{t+1jn} and promised value G^{t+1}_{jm} equal to G^{t+1}_{jn}, will increase both households' utility level without violate the incentive and participation constraint. Contradiction to the contract is optimal on state m. By

$$c^g_{tjn} \leq c^g_{tjm},$$

and

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jm}, G^{t+1}_{jz})}{\partial s^g_{t+1jn}} \geq \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^g_{t+1jn}}.$$

I get

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \leq g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right),$$

which contradict to $g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) > g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right)$. So I get

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jm}, G^{t+1}_{jz})}{\partial s^g_{t+1jn}} < \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^g_{t+1jn}}.$$

Second, I get $\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})}{\partial s^p_{t+1jn}} > \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^p_{t+1jm}}$. If

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})}{\partial s^p_{t+1jn}} \leq \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^p_{t+1jm}},$$

and

$$s^p_{t+1jn} + s^g_{t+1jn} \leq s^p_{t+1jm} + s^g_{t+1jm},$$

by Equations 3 and 4, I get $\frac{1}{c^p_{t,jm}} \leq \frac{1}{c^p_{t,jn}}$.

If

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})}{\partial s^p_{t+1jn}} \leq \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial s^p_{t+1jm}},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) < P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}),$$

by Lemma 3 I get $G_{jn}^{t+1} < G_{jm}^{t+1}$.

To get $g_t\left(s_t^g, s_t^p, \epsilon_n^g, \epsilon_j^p, G^t\right) > g_t\left(s_t^g, s_t^p, \epsilon_m^g, \epsilon_j^p, G^t\right)$, I have $c_{tjn}^g > c_{tjm}^g$. However, by $c_{tjn}^p \leq c_{tjm}^p$ and $c_{tjn}^g > c_{tjm}^g$, I get $\frac{c_{tjm}^g}{c_{tjm}^p} < \frac{c_{tjn}^g}{c_{tjn}^p}$, which contradict to $\frac{c_{tjm}^g}{c_{tjm}^p} = \frac{c_{tjn}^g}{c_{tjn}^p} = \lambda$. I get

$$P^{t+1}(s_{t+1jn}^p, s_{t+1jn}^g, G_{jn}^{t+1}) \geq P^{t+1}(s_{t+1jm}^p, s_{t+1jm}^g, G_{jm}^{t+1}),$$

and

$$p_t\left(s_t^g, s_t^p, \epsilon_{nt}^g, \epsilon_{jt}^p, G^t\right) > p_t\left(s_t^g, s_t^p, \epsilon_{mt}^g, \epsilon_{jt}^p, G^t\right),$$

which contradict to

$$p_t\left(s_t^g, s_t^p, \epsilon_{nt}^g, \epsilon_{jt}^p, G^t\right) < p_t\left(s_t^g, s_t^p, \epsilon_{mt}^g, \epsilon_{jt}^p, G^t\right).$$

With the two equations, I get:

$$\begin{aligned} & \frac{\partial P^{t+1}(s_{t+1jn}^p, s_{t+1jn}^g, G_{jz}^{t+1})}{\partial s_{t+1jn}^g} - \frac{\partial P^{t+1}(s_{t+1jn}^p, s_{t+1jn}^g, G_{jn}^{t+1})}{\partial s_{t+1jn}^p} \\ < \quad & \frac{\partial P^{t+1}(s_{t+1jm}^p, s_{t+1jm}^g, G_{jz}^{t+1})}{\partial s_{t+1jn}^g} - \frac{\partial P^{t+1}(s_{t+1jm}^p, s_{t+1jm}^g, G_{jm}^{t+1})}{\partial s_{t+1jm}^p}. \end{aligned}$$

Contradiction to both sides equal to 0. In the optimal contract, $s_{t+1jn}^p + s_{t+1jn}^g > s_{t+1jm}^p + s_{t+1jm}^g$.

Case 2 When $\mathbf{s}_{t+1jn}^p + \mathbf{s}_{t+1jn}^g > \mathbf{s}_{t+1jm}^p + \mathbf{s}_{t+1jm}^g$

By assumption $s_{t+1jn}^p + s_{t+1jn}^g > s_{t+1jm}^p + s_{t+1jm}^g$, and lemma 1, I get

$$G_{jn}^{t+1} > G_{jm}^{t+1},$$

and

$$P^{t+1}(s_{t+1jn}^p, s_{t+1jn}^g, G_{jn}^{t+1}) > P^{t+1}(s_{t+1jm}^p, s_{t+1jm}^g, G_{jm}^{t+1}).$$

By Equations 3 and **??**, I get $c^g_{tjn} > c^g_{tjm}$ and $c^p_{tjn} > c^p_{tjm}$. Both households have higher utility on state m. Fixed $\frac{c^g_{tjl}}{c^p_{tjl}} = \lambda$, both $p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)$ and $g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)$ are strictly increasing on ϵ^g_{zt}. I get Use the same the method, I can prove that fixed $\frac{c^g_{tjl}}{c^p_{tjl}} = \lambda$, both $g_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)$ and $p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)$ are strictly increasing on ϵ^p_{jt}. ■

Proof of Proposition 5 **Proof.** I use three steps to prove the lemma.

Step 1 Define boundary for full risk sharing.

First, I define the upper bounds

I define the marginal value of income shock given income shock ϵ^p_{jt}, cause parent incentive constraint binding. $\exists$ *a* number $\epsilon^g_{kt} \in \left[\epsilon^g_{1t}, \epsilon^g_{Zt}\right]$ (ϵ^g_{kt} can be a shock or not) such that satisfies that:

$$\frac{c^g_{tjk}}{c^p_{tjk}} = \lambda,$$

and

$$p_t\left(s^g_t, s^p_t, \epsilon^g_k, \epsilon^p_j, G^t\right) = v^{aut}_{pt}(s^p_t, \epsilon^p_j).$$

Given ϵ^p_j, $\forall\ \epsilon^g_{zt} < \epsilon^g_{kt}$, such that the $\frac{c^g_{tjk}}{c^p_{tjk}} < \lambda; \forall\ \epsilon^g_{zt} > \epsilon^g_{kt}$, such that the $\frac{c^g_{tjk}}{c^p_{tjk}} \geq \lambda$.

For any ϵ^p_{jt}, the marginal value ϵ^g_{kt} only exists once at most. By Lemma 3, fixed $\frac{c^g_{tjl}}{c^p_{tjl}} = \lambda$, $p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)$ is strictly increasing on ϵ^g_{zt}. $v^{aut}_{pt+1}(s^p_t, \epsilon^p_{jt})$ is a constant, given ϵ^p_{jt}. Only one number satisfies $p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt+1}(s^p_t, \epsilon^p_{jt})$. For any ϵ^p_{jt}, the marginal value ϵ^g_{kt} only exists once at most.

Second, I define the lower bounds.

Define another marginal value $\exists \epsilon_{lt}^{g} \in \left[\epsilon_{1t}^{g}, \epsilon_{Zt}^{g}\right]$ (can be a shock or not) such that:

$$\frac{c_{tjl}^{g}}{c_{tjl}^{p}} = \lambda,$$

and

$$g_t\left(s_t^g, s_t^p, \epsilon_{lt}^g, \epsilon_{jt}^p, G^t\right) = v_{gt}^{aut}(s_t^g, \epsilon_{lt}^g).$$

Fixed ϵ_{jt}^{p}, if $\epsilon_{zt}^{g} > \epsilon_{lt}^{g}$, $\frac{c_{tjl}^{g}}{c_{tjl}^{p}} > \lambda$, and $g_t\left(s_t^g, s_t^p, \epsilon_{lt}^g, \epsilon_{jt}^p, G^t\right) = v_{gt}^{aut}(s_t^g, \epsilon_{lt}^g)$. Use the same method above, I can prove the uniqueness of the marginal value.

Step 2 within the boundaries, full risk sharing

For $\forall \epsilon_{zt}^{g} \in \left[\epsilon_{kt}^{g}, \epsilon_{lt}^{g}\right]$, both $p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)$ and $g_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)$ decreasing on ϵ_{zt}^{g}. The substitution rate of the marginal utility is a constant, with

$$-\frac{\partial P^{t+1}(s_{t+1jz}^{p}, s_{t+1jz}^{g}, G_{jz}^{t+1})}{\partial G_{jz}^{t+1}} = \frac{c_{tjz}^{g}}{c_{tjz}^{p}} = \lambda.$$

Both households have lower utility values, as one gets a smaller income shock. Both $p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)$ and $g_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)$ are strictly decreasing on ϵ_{zt}^{g}.

Step 3 beyond the boundaries, partial risk sharing

For $\forall \epsilon_{zt}^{g} \in \left[\epsilon_{1t}^{g}, \epsilon_{kt}^{g}\right]$, $\frac{c_{tjz}^{g}}{c_{tjz}^{p}} < \lambda$.

If $\frac{c_{tjz}^{g}}{c_{tjz}^{p}} = \frac{\lambda + \varkappa_{jz}}{1+\omega_{jz}} \geq \lambda = \frac{c_{tjl}^{g}}{c_{tjl}^{p}}$, with grandparent's incentive binding, I get:

$$\begin{aligned} g_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right) &= v_{gt}^{aut}(s_t^g, \epsilon_{zt}^g) < v_{gt}^{aut}(s_t^g, \epsilon_{kt}^g) < g_t\left(s_t^g, s_t^p, \epsilon_{kt}^g, \epsilon_{jt}^p, G^t\right); \\ p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right) &> v_{pt}^{aut}(s_t^p, \epsilon_{jt}^p) = p_t\left(s_t^g, s_t^p, \epsilon_{kt}^g, \epsilon_{jt}^p, G^t\right). \end{aligned}$$

It means if set $\frac{c_{tjz}^{g}}{c_{tjz}^{p}} = \lambda$, decrease c_{tjz}^{g} and G_{jz}^{t+1} will increase $p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right)$ with $p_t\left(s_t^g, s_t^p, \epsilon_{zt}^g, \epsilon_{jt}^p, G^t\right) > v_{pt}^{aut}(s_t^p, \epsilon_{jt}^p)$. Contradiction to the conclusion of Lemma 2 that

$p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)$ is strict increasing on ϵ^g_{zt}, when $\frac{c^g_{tjz}}{c^p_{tjz}} = \lambda$. $\frac{c^g_{tjz}}{c^p_{tjz}} < \lambda$ means parent's incentive constraint binding. Use the same method, I can prove $\forall \epsilon^g_{zt} \in \left[\epsilon^g_{lt}, \epsilon^g_{Zt}\right]$, $\frac{c^g_{tjz}}{c^p_{tjz}} > \lambda$, with grandparent's incentive constraint binding. ■

<u>Proof of Proposition 6</u> **Proof.** In the full risk sharing region defined by proposition 5, there is $c^g_{tjn}/c^p_{tjn} = c^g_{tjm}/c^p_{tjm}$. I then want to prove when one IC is binding, if $\epsilon^g_{mt} < \epsilon^g_{nt}$, then $c^g_{tjn}/c^p_{tjn} > c^g_{tjm}/c^p_{tjm}$.

When ICp is binding

I want to prove, fixed ϵ^p_{jt}, if $\epsilon^g_{mt} < \epsilon^g_{nt}$, then $c^g_{tjn}/c^p_{tjn} > c^g_{tjm}/c^p_{tjm}$. Grandparent's incentive constraints are not binding in both cases, with $\varkappa_{jn} = \varkappa_{jm} = 0$. Parent's incentive constraints are binding, such that:

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) = p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt+1}(s^p_t, \epsilon^p_{jt})$$

By Lemma 2, I get:

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) > g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right).$$

Then I will prove that if $c^p_{tjn} > c^p_{tjm}$, then

$$U_{pt}(c^p_{tjn}, l^p_{tjn}, K^p_{tjn}, N^p_{tjn}) > U_{pt}(c^p_{tjm}, l^p_{tjm}, K^p_{tjm}, N^p_{tjm}).$$

Define $\kappa_t = \min\{w^g_t, w^p_t, p^o_t\}$, $\chi_t = \min\{w^p_t, p^o_t\}$. By the utility form, in the optimal contract, I get:

$$\frac{1}{c^i_t} = \frac{\alpha_t}{K^i_t \kappa_t} = \frac{\gamma_t}{\chi_t N^i_t}$$

If $c^p_{tjn} > c^p_{tjm}$, then

$$U_{pt}(c^p_{tjn}, l^p_{tjn}, K^p_{tjn}, N^p_{tjn}) > U_{pt}(c^p_{tjm}, l^p_{tjm}, K^p_{tjm}, N^p_{tjm}).$$

If $c^g_{tjn} > c^g_{tjm}$, then

$$U_{gt}(c^g_{tjn}, l^g_{tjn}, K^g_{tjn}, N^g_{tjn}) > U_{gt}(c^g_{tjm}, l^g_{tjm}, K^g_{tjm}, N^g_{tjm}).$$

In the interior region that IC is binding, fixed ϵ^g_z, as ϵ^p_j increase or decrease,

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt}(s^p_t, \epsilon^p_{jt}),$$

and

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{zt} + \epsilon, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt}(s^p_t, \epsilon^p_{jt} + \epsilon).$$

I get

$$\frac{\partial v^{aut}_{pt}(s^p_t, \epsilon^p_{jt})}{\partial \epsilon^p_{jt}} = \frac{\partial p_t\left(s^g_t, s^p_t, \epsilon^g_{zt}, \epsilon^p_{jt}, G^t\right)}{\partial \epsilon^p_{jt}},$$

then

$$\frac{1}{c^{paut}_{tjn}} = \frac{1}{c^{paut}_{tjm}} = \frac{1}{c^p_{tjm}} = \frac{1}{c^p_{tjmn}}.$$

I get

$$\begin{aligned} & \left(1 - \varrho_{t+8}\right) P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}) + \varrho_{t+8} V^{aut}_{pt+1}(s^p_{t+1jm} + s^g_{t+1jm}) \\ = \; & \left(1 - \varrho_{t+8}\right) P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) + \varrho_{t+8} V^{aut}_{pt+1}(s^p_{t+1jn} + s^g_{t+1jn}) \end{aligned} \tag{9}$$

and

$$U_{pt}(c^p_{tjn}, l^p_{tjn}, K^p_{tjn}, N^p_{tjn}) = U_{pt}(c^p_{tjm}, l^p_{tjm}, K^p_{tjm}, N^p_{tjm}). \tag{10}$$

If $s^p_{t+1jn} + s^g_{t+1jn} \leq s^p_{t+1jm} + s^g_{t+1jm}$, by Equation 9, I get:

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) \geq P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}),$$

and

$$G^{t+1}_{jn} \leq G^{t+1}_{jm}.$$

By $g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) > g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right)$, I get

$$U_{gt}(c^g_{tjn}, l^g_{tjn}, K^g_{tjn}, N^g_{tjn}) > U_{gt}(c^g_{tjm}, l^g_{tjm}, K^g_{tjm}, N^g_{tjm}),$$

and

$$c^g_{tjn} > c^g_{tjm}.$$

I get:

$$c^g_{tjn}/c^p_{tjn} > c^g_{tjm}/c^p_{tjm},$$

and

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})}{\partial G^{t+1}_{jn}} \leq \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial G^{t+1}_{jm}}.$$

Contradiction to Proposition 4.

$\forall \epsilon^g_{mt}, \epsilon^g_{nt} \in \left[\epsilon^g_{lt}, \epsilon^g_{Zt}\right]$, if $\epsilon^g_{mt} < \epsilon^g_{nt}$, then $s^p_{t+1jn} + s^g_{t+1jn} > s^p_{t+1jm} + s^g_{t+1jm}$. With $s^p_{t+1jn} + s^g_{t+1jn} > s^p_{t+1jm} + s^g_{t+1jm}$, assume $c^g_{tjn}/c^p_{tjn} \leq c^g_{tjm}/c^p_{tjm}$. By Equation 9, I get:

$$c^g_{tjn} \leq c^g_{tjm},$$

and

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) < P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm}).$$

If $s^p_{t+1jn}+s^g_{t+1jn}>s^p_{t+1jm}+s^g_{t+1jm}$, by Equation 9, I get $G^{t+1}_{jn}\geq G^{t+1}_{jm}$. It means:

$$\frac{\partial P^{t+1}(s^p_{t+1jn},s^g_{t+1jn},G^{t+1}_{jn})}{\partial G^{t+1}_{jn}}>\frac{\partial P^{t+1}(s^p_{t+1jm},s^g_{t+1jm},G^{t+1}_{jm})}{\partial G^{t+1}_{jm}}.$$

Contradictioj to Proposition 4. $\forall \epsilon^g_{mt},\epsilon^g_{nt}\in\left[\epsilon^g_{lt},\epsilon^g_{Zt}\right]$, if $\epsilon^g_{mt}<\epsilon^g_{nt}$, then $s^p_{t+1jn}+s^g_{t+1jn}>s^p_{t+1jm}+s^g_{t+1jm}$ and $c^g_{tjn}/c^p_{tjn}>c^g_{tjm}/c^p_{tjm}$. I get $c^g_{tjn}>c^g_{tjm}$.

When ICg is binding

I then want to prove that fixed ϵ^p_{jt}, $\forall\epsilon^g_{mt},\epsilon^g_{nt}\in\left[\epsilon^g_{1t},\epsilon^g_{kt}\right]$, if $\epsilon^g_{mt}<\epsilon^g_{nt}$, $c^g_{tjn}/c^p_{tjn}>c^g_{tjm}/c^p_{tjm}$. Grandparent's incentive constraints are not binding in both cases, with $\varkappa_{jn}=\varkappa_{jm}=0$. Parent's incentive constraints are binding:

$$g_t\left(s^g_t,s^p_t,\epsilon^g_{nt},\epsilon^p_{jt},G^t\right)=g_t\left(s^g_t,s^p_t,\epsilon^g_{mt},\epsilon^p_{jt},G^t\right)$$

and

$$p_t\left(s^g_t,s^p_t,\epsilon^g_{nt},\epsilon^p_{jt},G^t\right)>p_t\left(s^g_t,s^p_t,\epsilon^g_{mt},\epsilon^p_{jt},G^t\right)$$

Otherwise, all the allocation in state n is available at state m, and Pareto improvement exists for the contract at state n. Contradict to the contract is optimal. By Equation 9, I get:

$$\frac{1}{c^{gaut}_{tjn}}=\frac{1}{c^{gaut}_{tjm}}=\frac{1}{c^g_{tjm}}=\frac{1}{c^g_{tjmn}}.$$

If $g_t\left(s^g_t,s^p_t,\epsilon^g_{nt},\epsilon^p_{jt},G^t\right)=g_t\left(s^g_t,s^p_t,\epsilon^g_{mt},\epsilon^p_{jt},G^t\right)$, I get $G^{t+1}_{jn}=G^{t+1}_{jm}$. If $s^p_{t+1jn}+s^g_{t+1jn}\leq s^p_{t+1jm}+s^g_{t+1jm}$, by Equation 6, I get

$$P^{t+1}(s^p_{t+1jn},s^g_{t+1jn},G^{t+1}_{jn})\leq P^{t+1}(s^p_{t+1jm},s^g_{t+1jm},G^{t+1}_{jm}).$$

Define the total endowment by Equation 1, such that

$$\begin{aligned}
& E_{tjz} - s^p_{t+1jz} - s^g_{t+1jz} \\
= \quad & c^g_{tjz}\left(1+\eta_{t+6}+\alpha_{t+6}+\gamma_{t+6}\right) + c^p_{tjz}\left(1+\eta_t+\alpha_t+\gamma_t\right)
\end{aligned}$$

Because state n has more endowment, $E_{tjn} > E_{tjm}$. With $s^p_{t+1jn}+s^g_{t+1jn} \leq s^p_{t+1jm}+s^g_{t+1jm}$ and $c^g_{tjn} = c^g_{tjm}$, I get $c^p_{tjn} > c^p_{tjm}$.

$$\begin{aligned}
c^g_{tjn}/c^p_{tjn} \quad &> \quad c^g_{tjm}/c^p_{tjm} \\
\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})}{\partial G^{t+1}_{jn}} \quad &< \quad \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial G^{t+1}_{jm}}
\end{aligned}$$

Contradiction to Proposition 4. $\forall \epsilon^g_{mt}, \epsilon^g_{nt} \in \left[\epsilon^g_{1t}, \epsilon^g_{kt}\right]$, if $\epsilon^g_{mt} < \epsilon^g_{nt}$, then $s^p_{t+1jn}+s^g_{t+1jn} > s^p_{t+1jm}+s^g_{t+1jm}$.With $s^p_{t+1jn}+s^g_{t+1jn} > s^p_{t+1jm}+s^g_{t+1jm}$, by Equation 9,

$$P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn}) > P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})$$

I get

$$\frac{\partial P^{t+1}(s^p_{t+1jn}, s^g_{t+1jn}, G^{t+1}_{jn})}{\partial G^{t+1}_{jn}} > \frac{\partial P^{t+1}(s^p_{t+1jm}, s^g_{t+1jm}, G^{t+1}_{jm})}{\partial G^{t+1}_{jm}}.$$

By Equations 8, I get $c^p_{tjn} > c^p_{tjm}$.

$\forall \epsilon^g_{mt}, \epsilon^g_{nt} \in \left[\epsilon^g_{1t}, \epsilon^g_{kt}\right]$, if $\epsilon^g_{mt} < \epsilon^g_{nt}$, then

$$s^p_{t+1jn}+s^g_{t+1jn} > s^p_{t+1jm}+s^g_{t+1jm}$$

and

$$c^g_{tjn}/c^p_{tjn} > c^g_{tjm}/c^p_{tjm}$$

Using the same methods, I can prove fixed ϵ^g_{zt}, $\forall \epsilon^p_{mt} < \epsilon^p_{nt}$, $c^g_{tnz}/c^p_{tnz} \geq c^g_{tmz}/c^p_{tmz}$. ■

Corollary 2

Corollary 19 *Fixed ϵ_{jt}^{p}, if $\epsilon_{mt}^{g} > \epsilon_{nt}^{g}$, then $\omega_{jn} \leq \omega_{jm}$ and $\mu_{jn} \geq \mu_{jm}$. Fixed ϵ_{zt}^{g}, if $\epsilon_{mt}^{p} > \epsilon_{nt}^{p}$, then $\omega_{jn} \geq \omega_{jm}$ and $\mu_{jn} \leq \mu_{jm}$.*

Proof. By Proposition 6, with $\frac{c_{tjz}^{g}}{c_{tjz}^{p}} = \frac{\lambda+\varkappa_{jz}}{1+\omega_{jz}}$, $\frac{c_{t}^{g}}{c_{t}^{p}}$ is decreasing on ϵ_{j}^{p} and increasing on ϵ_{z}^{g}. By Proposition 2 at most one IC is binding. When $\varkappa_{jz} > 0, \omega_{jz} = 0$; when $\omega_{jz} > 0$, $\varkappa_{jz} = 0$.

$\epsilon_{mt}^{g} < \epsilon_{nt}^{g}, \frac{c_{tjm}^{g}}{c_{tjm}^{p}} = \frac{\lambda+\varkappa_{jm}}{1+\omega_{jm}} \leq \frac{\lambda+\varkappa_{jn}}{1+\omega_{jn}} = \frac{c_{tjn}^{g}}{c_{tjn}^{p}}$. We have either

$$\begin{aligned} \varkappa_{jn} &\geq \varkappa_{jm} > 0 \\ \omega_{jn} &= \omega_{jm} = 0 \end{aligned}$$

or

$$\begin{aligned} \varkappa_{jn} &= \varkappa_{jm} = 0 \\ \omega_{jm} &\geq \omega_{jn} > 0 \end{aligned}$$

or

$$\begin{aligned} \varkappa_{jn} &\geq \varkappa_{jm} = 0 \\ \omega_{jn} &\geq \omega_{jm} = 0 \end{aligned}$$

$\varkappa_{jz}$ is increasing on ϵ_{zt}^{g}; and ω_{jz} is decreasing on ϵ_{z}^{g}. Use the same method, I can prove $\varkappa_{jz}$ is decreasing on ϵ_{jt}^{p} and ω_{jz} is increasing on ϵ_{j}^{p}. ■

Proof of Proposition 7 **Proof.** Prove as grandparent gets bad income shock, both households will have lower utility values and consumption at time t.

Step 1: From the proof of Proposition 6, I get in the region with one's incentive constraint binding, big shock has no effect on the utility value of the household with incentive constraint binding and will increase the other household's utility level. Otherwise, exist a feasible Pareto improvement for the optimal contract. In the region with partial risk sharing, as one gets bad income shock, both households' utility value and consumption at time t will decrease.

Step 2: From the proof of Lemma 4, in the region that no incentive constraints binding, as one gets bad income shock two household's utility value will decrease. The lemma means in full-commitment, because the consumption ratio is fixed, as one gets bad income shock, two household's utility value and consumption at time t will decrease.

Step 3: Then, I need to prove that for two income shocks belonged to different regions, the relationship still holds. Fixed ϵ^p_{jt}, if $\epsilon^g_{nt} > \epsilon^g_{kt} > \epsilon^g_{mt}$,

$$g\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \geq g\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right),$$

$$p\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \geq p\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right),$$

and

$$c^p_{jn} \geq c^p_{jm}, c^g_{jn} \geq c^g_{jm}.$$

ϵ^g_{kt} is the marginal value defined by Proposition 5.

First, with $\epsilon^g_{kt} > \epsilon^g_{mt}$ and $\epsilon^g_{kt} \leq \epsilon^g_{nt}$, parent's incentive constraint is binding. With

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \geq p_t\left(s^g_t, s^p_t, \epsilon^g_{kt}, \epsilon^p_{jt}, G^t\right) = v^{aut}_{pt}(s^p_t, \epsilon^p_{jt})$$

and

$$c^p_{jn} = c^p_{jk}.$$

By the conclusion of step 1, $g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \geq g_t\left(s^g_t, s^p_t, \epsilon^g_{kt}, \epsilon^p_{jt}, G^t\right)$ and $c^g_{jn} > c^g_{jk}$. By the conclusion of Step 2, I get

$$g_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \geq g_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right),$$

$$p_t\left(s^g_t, s^p_t, \epsilon^g_{nt}, \epsilon^p_{jt}, G^t\right) \geq p_t\left(s^g_t, s^p_t, \epsilon^g_{mt}, \epsilon^p_{jt}, G^t\right),$$

and

$$c^p_{jn} \geq c^p_{jm}, c^g_{jn} \geq c^g_{jm}.$$

Use the same methods, I can prove as the parent gets a smaller income shock, both households will have lower utility values and consumption. ■

<u>Calculation of the Household Death Rate</u> I get average people death rate by age from government population statistics. In the data, female death rate is p^f_τ and male death rate is p^m_τ at age τ. In the model, each period has 4 years. The death rate household in period t is defined as that for the household with at least one household member alive in period t-1, the probability everyone die in period t. I use two steps to calculate the household death rate. In step one, I get the individual death rate in each period. The death rate, I get is

the probability that all household will die in the four years. In household i, wife's death rate is ϱ_{it}^{w} and husband's death rate is ϱ_{it}^{h}, which is defined as

$$\varrho_{it}^{h} = \sum_{\varkappa=4t}^{4t+3} \left[p_{\varkappa}^{m} \prod_{\varphi=1}^{\varkappa-1} \left(1 - p_{\varphi}^{m}\right) \right],$$

and

$$\varrho_{it}^{w} = \sum_{\varkappa=4t}^{4t+3} \left[p_{\varkappa}^{f} \prod_{\varphi=1}^{\varkappa-1} \left(1 - p_{\varphi}^{f}\right) \right].$$

In step two, I get the household death rate of each period. In period t, the probability both husband and wife are alive is

$$\Omega_{t}^{i,b} = \prod_{\tau=1}^{t} \left[\left(1 - \varrho_{i\tau}^{h}\right) \left(1 - \varrho_{i\tau}^{w}\right) \right];$$

the probability only wife alive is

$$\Omega_{t}^{i,w} = \sum_{\varkappa=1}^{t} \left[\varrho_{\varkappa}^{h} \prod_{\tau=1}^{t} \left(1 - \varrho_{i\tau}^{w}\right) \prod_{\varphi=1}^{\varkappa-1} \left(1 - \varrho_{\varphi}^{h}\right) \right];$$

the probability only husband alive is

$$\Omega_{t}^{i,h} = \sum_{\varkappa=1}^{t} \left[\varrho_{\varkappa}^{w} \prod_{\tau=1}^{t} \left(1 - \varrho_{i\tau}^{h}\right) \prod_{\varphi=1}^{\varkappa} \left(1 - \varrho_{\varphi}^{w}\right) \right].$$

So household i's death rate in period t+1 is given by:

$$\Psi_{t+1}^{i} = \frac{\Omega_{t}^{i,b} \varrho_{it+1}^{h} \varrho_{it+1}^{w} + \Omega_{t}^{i,w} \varrho_{it+1}^{w} + \Omega_{t}^{i,h} \varrho_{it+1}^{h}}{\Omega_{t}^{i,b} + \Omega_{t}^{i,w} + \Omega_{t}^{i,h}} \tag{11}$$

$\Omega_{t}^{i,b} + \Omega_{t}^{i,w} + \Omega_{t}^{i,h}$ is the probability that household i still has at least one people alive in period t. $\Omega_{t}^{i,b} \varrho_{it+1}^{h} \varrho_{it+1}^{w}$ is the probability that both household members die at period t+1 when both households members alive at period t. $\Omega_{t}^{i,w} \varrho_{it+1}^{w}$ is the probability that wife dies in t+1 when only wife alive at period t. $\Omega_t^{i,h}\varrho_{it+1}^h$ is the probability that husband dies at t+1 when only husband alive in period t. With death rate data from the National Population and Reproductive Health Science Data Center of China, I get the estimated household death rate by using the equations above. Figure 9 in the Appendix shows the estimated household death rate for each period. The individual death rate before age 40 is almost zero. I assume that the death rate in the age 21 is 1 and the parent's death rate before age 14 is 0.

<u>GLS Estimation on Preference Parameters</u> Given the linear relationship between the spending in each term, I use two steps generalized least squares regression (GLS) to get parameters. In the first step, I use ordinary least squares (OLS) regression to get the parameters and denote the vector of all parameter values as $\Lambda_t = (\alpha_t, \gamma_t, \eta)'$. I define the vector of dependent variables y_i, the vector of independent variables $x_i = \pi_{ii}^C \times I_3$, the vector of error term, $u_i = (u_{i1}, u_{i2}, u_{i3})'$, with $u_{i1} = \pi_{it}^K - \alpha_t\pi_{it}^C$, $u_{i2} = \pi_{it}^N - \gamma_t\pi_{it}^C$, and $u_{i3} = \pi_{it}^L - \eta\pi_{it}^C$. Correspondingly, moments for household i are thus given as

$$y_i = \Lambda_{it}x_i + u_i = (\Lambda_t + \xi_{it})\, x_i + u_i. \tag{12}$$

I use OLS regression to get $\tilde{\Lambda}_{it} = \left(x_i^T x_i\right)^{-1} x_i^T y_i$. With the distribution of $\tilde{\Lambda}_{it}$, I get the sample mean $\tilde{\Lambda}_t$and the sample variance $\tilde{\Sigma}_t$ of $\tilde{\Lambda}_{it}$.

In the second step, I use a generalized least squares estimation to get the best linear unbiased estimator of the random coefficient regression models. Using the variance estimated from the first step, I define the weight $\tilde{w}_i = \tilde{\Sigma}_t$. Using GLS regression, I get

$$\tilde{\Lambda}^g_{it} = \left(x^T_{ii} \tilde{w}^{-1}_i x_i \right)^{-1} x^T_i w^{-1}_i y_i. \tag{13}$$

I then get the sample mean $\tilde{\Lambda}^g_t$ and sample variance Σ^g_t of $\tilde{\Lambda}^g_{it}$, in which $\tilde{\Lambda}^g_t$ captures the average utility weight in each utility term and Σ^g_t captures the heterogeneous preference distribution.

<u>The GMM Estimation of Two-Types Models</u> The two types are specified as $\Lambda^1_{it} = \left\{\alpha^1_{it}, \eta^1_{it}, \gamma^1_{it}\right\}$ and $\Lambda^2_{it} = \left\{\alpha^2_{it}, \eta^2_{it}, \gamma^2_{it}\right\}$. I define $\bar{\alpha}^i_t = \frac{\alpha^i_t}{1+\eta^i+\alpha^i_t+\gamma^i_t}$, $\bar{\gamma}^i_t = \frac{\gamma^i_t}{1+\eta^i+\alpha^i_t+\gamma^i_t}$, and $\bar{\eta}^i = \frac{\eta^i}{1+\eta^i+\alpha^i_t+\gamma^i_t}$, with $i \in \{1, 2, g\}$. I draw N (N=10,000) pair of η^g_i, α^g_{it} and γ^g_{it} from the parameter distribution of the GLS distribution. I define the sample means: $\mu_{\alpha t} = \sum_{i=1}^{N} \frac{\bar{\alpha}^g_{it}}{N}$, $\mu_{\gamma t} = \sum_{i=1}^{N} \frac{\bar{\gamma}^g_{it}}{N}$, and $\mu_{\eta t} = \sum_{i=1}^{N} \frac{\bar{\eta}^g_{it}}{N}$. I define the sample variances: $\sigma^2_{\alpha t} = \sum_{i=1}^{N} \frac{\left(\bar{\alpha}^g_{it} - \mu_{\alpha t}\right)^2}{N-1}$, $\sigma^2_{\gamma t} = \sum_{i=1}^{N} \frac{\left(\bar{\gamma}^g_{it} - \mu_{\gamma t}\right)^2}{N-1}$, and $\sigma^2_{\eta t} = \sum_{i=1}^{N} \frac{\left(\bar{\eta}^g_{it} - \mu_{\eta t}\right)^2}{N-1}$. I define the sample covariance: $q_{\alpha\eta t} = \sum_{i=1}^{N} \frac{\left(\bar{\alpha}^g_{it} - \mu_{\alpha t}\right)\left(\bar{\eta}^g_{it} - \mu_{\eta t}\right)}{N-1}$. I draw N (N=10,000) pair of η^g_i, α^g_{it} and γ^g_{it} from the parameter distribution of the GLS distribution. I define the sample means $\mu^i_t = (\mu^i_{\alpha t},\ \mu^i_{\gamma t},\ \mu^i_{\eta t})$, the sample variances $\sigma^i_t = \left(\sigma^i_{\alpha t}, \sigma^i_{\gamma t}, \sigma^i_{\eta t}\right)$ and the sample covariance $q^i_t = \left(q^i_{\alpha\eta t}, q^i_{\alpha\gamma t}, q^i_{\alpha\gamma t}\right)$, for $\bar{\Lambda}^i_t$.The probability that i' is type 1 is ρ and that i is type 2 is $1-\rho$. The probability that i' is type 1 is ρ and that i is type 2 is $1-\rho$. Define $\theta = \left(\rho, \eta^1_i, \alpha^1_{it}, \gamma^1_{it}, \eta^2_i, \alpha^2_{it}, \gamma^2_{it}\right)$, and $Y_t = \left(\mu^g_{\alpha t}, \mu^g_{\gamma t}, \mu^g_{\eta t}, \sigma^g_{\alpha t}, \sigma^g_{\gamma t}, \sigma^g_{\eta t}, q^g_{\alpha\eta t}\right)$, I get the moments $m(\theta) = \frac{1}{T}\sum_{t=1}^{T} g(\theta, Y_t)$, such that:

$$\mu_{\alpha t} = \rho \bar{\alpha}^{1}_{it} + (1-\rho)\,\bar{\alpha}^{2}_{it} \tag{14}$$

$$\mu_{\gamma t} = \rho \bar{\eta}^{1}_{it} + (1-\rho)\,\bar{\eta}^{2}_{it} \tag{15}$$

$$\mu_{\eta t} = \rho \bar{\gamma}^{1}_{it} + (1-\rho)\,\bar{\gamma}^{2}_{it} \tag{16}$$

$$\sigma^2_{\alpha t} = \rho \left(\bar{\alpha}^{1}_{it} - \bar{\alpha}^{g}_{t}\right)^2 + (1-\rho)\left(\bar{\alpha}^{2}_{it} - \bar{\alpha}^{g}_{t}\right)^2 \tag{17}$$

$$\sigma^2_{\gamma t} = \rho \left(\bar{\eta}^{1}_{it} - \bar{\eta}^{g}_{t}\right)^2 + (1-\rho)\left(\bar{\eta}^{2}_{it} - \bar{\eta}^{g}_{t}\right)^2 \tag{18}$$

$$\sigma^2_{\eta t} = \rho \left(\bar{\gamma}^{1}_{it} - \bar{\gamma}^{g}_{t}\right)^2 + (1-\rho)\left(\bar{\gamma}^{2}_{it} - \bar{\gamma}^{g}_{t}\right)^2 \tag{19}$$

and

$$q^{i}_{\alpha\eta t} = \rho^2 \sigma^1_{\alpha t}\sigma^1_{\eta t} + (1-\rho)^2\,\sigma^2_{\alpha t}\sigma^2_{\eta t} + \rho\,(1-\rho)\,\sigma^1_{\alpha t}\sigma^2_{\eta t} + \sigma^2_{\alpha t}\sigma^1_{\eta t}] \tag{20}$$

Define the GMM estimator is:

$$\hat{\theta} = \arg\min_{\theta} \left(\frac{1}{T}\sum_{t=1}^{T} g(\theta, Y_t)\right)^T \hat{W} \left(\frac{1}{T}\sum_{t=1}^{T} g(\theta, Y_t)\right) \tag{21}$$

I use two-step GMM to solve the problems:

In the first step, take $\hat{W} = I$ (the identity matrix), I compute preliminary GMM estimate $\hat{\theta}_{(1)}$ by equation 21. This estimator is consistent for θ, although not efficient.

In the second step, I define

$$\hat{W}_T\left(\hat{\theta}_{(1)}\right) = \left(\frac{1}{T}\sum_{t=1}^{T} g(\hat{\theta}_{(1)}, Y_t)\right)^T \left(\frac{1}{T}\sum_{t=1}^{T} g(\hat{\theta}_{(1)}, Y_t)\right)^{-1} \tag{22}$$

as the weighting matrix, to estimate θ in equation 21. The estimator will be asymptotically efficient.

F test of Two-Type Models I calculate the average weight of spending on child care with continuous type $\bar{\alpha}_t^g$. I define the F statistic:

$$F_\alpha = \frac{\sum_{i=1}^{N} \left(\bar{\alpha}_{it}^{g} - \bar{\alpha}_{it}^{\rho} \right)^2}{\sigma_{\alpha t}^2 (N-2)} \tag{23}$$

Here, $\bar{\alpha}_{it}$ and $\bar{\eta}_i$ are the categories α_{it} and η_i fit in. For example, $\rho\%$ is the probability that $\alpha = \alpha_1$ and α_ρ is the ρ percentile. I define $\bar{\alpha}_{it}^{\rho} = \alpha_1$ if $\alpha_{it} > \alpha_\rho$; otherwise $\bar{\alpha}_{it}^{\rho} = \alpha_2$. N is the number of the observations.

Use the same methods, I define the F statistic of leisure, elder care weights at different age.

MLE Estimation of Substitution Rate Assume $\varepsilon_t^i = w_t^i - \iota p_t$, where ε_t^i satisfies normal distribution, with $\varepsilon_t^i \tilde{} N(0, \Omega)$. Let $\Phi\left(w_t^i - \iota p_t\right)$ denotes the cumulative distribution function. Then the log-likelihood function is:

$$\ln L(\iota) = \sum_{i=1}^{n} \left[o_t^i \ln \Phi\left(w_t^i - \iota p_t\right) + \left(1 - o_t^i\right) \ln\left(1 - \Phi\left(w_t^i - \iota p_t\right)\right) \right] \tag{24}$$

By using maximum likelihood estimation to estimate the probit model, I can get the substitution rate between the outside market care service and the household care service. Using the market service price and the substitution rate, I get the real market care price.

Proof of proposition 9 **Proof.** First, $s^*(\theta)$ must be continuos value.

By contradiction, assume $\exists\theta_1, \theta_2$ and there is a hole between $s^*(\theta_1)$ and $s^*(\theta_2)$ with $s^*(\theta_2) > s^*(\theta_1)$. It means $\forall\theta \in \left[\underline{\theta}, \bar{\theta}\right]$, $s^*(\theta) \notin (s^*(\theta_1), s^*(\theta_2))$

In this condition, $\exists s^{**}(\theta_2) = s^*(\theta_2) - \varepsilon$, ε is small enough, such that $\pi_j^i(s^{**}(\theta_2), \theta) > \pi_j^i(s^*(\theta_2), \theta)$

There must be $\frac{\partial\pi_j^i(s^*(\theta_2),\theta)}{\partial s^*(\theta_2)} \geq 0, otherwise$ if $\frac{\partial\pi_j^i(s^*(\theta_2),\theta)}{\partial s^*(\theta_2)} < 0$, I can get $\pi_j^i(s^{**}(\theta_2), \theta) > \pi_j^i(s^*(\theta_2), \theta)$,contradiction.

I get

$$\frac{\partial U\left[P_j^i(s^*(\theta_2)), s^*(\theta_2)\right]}{\partial s^*} + \frac{\partial U\left[P_j^i(s^*), s^*\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial s^*} - \theta_2 \geq 0$$

By the concavity of U, I get:

$$\begin{aligned}
& \left[U\left[P_j^i(s^{**}(\theta_2)), s^{**}(\theta_2)\right] - C(\theta_2, s^{**}(\theta_2))\right] - \left[U\left[P_j^i(s^*(\theta_2)), s^*(\theta_2)\right] - C(\theta_2, s^*(\theta_2))\right] \\
= & \left[\frac{\partial U\left[P_j^i(s(\theta_2)), s(\theta_2)\right]}{\partial \tilde{s}} + \frac{\partial U\left[P_j^i(s), s\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial \tilde{s}} - \theta_2\right]\varepsilon > 0,
\end{aligned}$$

$\exists\tilde{s} \in [s^{**}(\theta_2), s^*(\theta_2)] > 0$.In which,

$$\begin{aligned}
& \frac{\partial U\left[P_j^i(s(\theta_2)), s(\theta_2)\right]}{\partial \tilde{s}} + \frac{\partial U\left[P_j^i(s), s\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial \tilde{s}} \\
> & \frac{\partial U\left[P_j^i(s^*(\theta_2)), s^*(\theta_2)\right]}{\partial s^*} + \frac{\partial U\left[P_j^i(s^*), s^*\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial s^*} \geq \theta_2
\end{aligned}$$

, *by* the concavity of U.

Contradict to $s^*(\theta_1)$ and $s^*(\theta_2)$ are equilibrium value. The function is continuos.

Second, the , $s^*(\theta)$ must be strict decreasing on θ

By contradiction, assume $\exists\theta_1, \theta_2$, with $\theta_1 = \theta_2 - \varepsilon, \varepsilon$ is small enough. In the equilibrium, $s^*(\theta_2) > s^*(\theta_1)$.

By the concavity of U, we have

$$\left[UP_j^i(s^*(\theta_2)), s^*(\theta_2) - C(\theta_1, s^*(\theta_2))\right] - \left[U\left[P_j^i(s^*(\theta_1)), s^*(\theta_1)\right] - C(\theta_1, s^*(\theta_1))\right]$$
$$= \left[\frac{\partial U\left[P_j^i(\tilde{s}), \tilde{s}\right]}{\partial \tilde{s}} + \frac{\partial U\left[P_j^i(\tilde{s}), \tilde{s}\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial \tilde{s}} - \theta_2\right] > 0$$

Contradict to $s^*(\theta_1)$ and $s^*(\theta_2)$ are equilibrium value. The function is continuos. ■

<u>Proof of proposition 10</u> **Proof.** Given the utility function, in the final equilibrium, $s^*(\theta)$ is continuos and strict decreasing on θ.

Without quota, for individual with talent θ, I have the number of people with type lower than his type in the equilibrium is:$1 - F_{all}(\theta)$. In the equilibrium, $P_i^w(s^*(\theta)) = (1-\mu)e^{-\lambda_1\theta} + \mu e^{-\lambda_0\theta}$. and $P_i^w(s^*(\underline{\theta})) = 0, P_i^w(s^*(\overline{\theta})) = 1$. In the final equilibrium $s^*(\theta)$ is:

$$s^*(\theta) = \left[(1-\mu)e^{-\lambda_1\theta} + \mu e^{-\lambda_0\theta}\right]\theta^{\frac{-1}{\alpha}}$$

In the all pay auction, all individual's payoff is 0 in the equilibrium.

Assume in the equilibrium, $\exists\theta', st\ \ \pi_j^i(s^*(\theta'), \theta') > 0$.

$\exists\ \theta'', \theta'' = \theta' + \varepsilon, , \varepsilon$ is small enough and $s^*(\theta'') = s^*(\theta') - \kappa, \epsilon$ is small enough.

And $s^{**}(\theta'') = s^*(\theta') + \epsilon, \epsilon$ is small enough. Such that

$$\left[U\left[P_j^i(s^{**}(\theta'')), s^{**}(\theta'')\right] - C(\theta'', s^{**}(\theta''))\right] - \left[U\left[P_j^i(s^*(\theta_1)), s^*(\theta_1)\right] - C(\theta_1, s^*(\theta_1))\right]$$
$$= \left[\frac{\partial U\left[P_j^i(s(\theta_2)), s(\theta_2)\right]}{\partial \tilde{s}} + \frac{\partial U\left[P_j^i(s), s\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial \tilde{s}} - \theta_2\right](\varepsilon + \kappa)$$

and $\exists \tilde{s} \in [s^{**}(\theta_2), s^{*}(\theta_2)] > 0$.In which,

$$
\begin{aligned}
& \frac{\partial U\left[P_j^i(s(\theta_2)), s(\theta_2)\right]}{\partial \tilde{s}} + \frac{\partial U\left[P_j^i(s), s\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial \tilde{s}} \\
> \quad & \frac{\partial U\left[P_j^i(s^*(\theta_2)), s^*(\theta_2)\right]}{\partial s^*} + \frac{\partial U\left[P_j^i(s^*), s^*\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial s^*} \geq \theta_2
\end{aligned}
$$

Contradict to $s^*(\theta^{''})$and $s^*(\theta^{'})$are equilibrium value. $\pi_j^i(s^*(\theta'), \theta') = 0$. Given the strict decreasing property, the optimal s is given by

$$
\frac{\partial U\left[P_j^i(s^*(\theta)), s^*(\theta)\right]}{\partial s^*} + \frac{\partial U\left[P_j^i(s^*), s^*\right]}{\partial P_j^i}\frac{\partial P_j^i}{\partial s^*} = \theta
$$

The lower θ is, the higher s and p people get. The college seat allocation function is:

$$
P_i^w(s) = F_P^{-1}(G_{all}^w(s)), i \in \{0, 1\}
$$

and

$$
P_i^q(s) = F_P^{-1}(G_i^q(s)), i \in \{0, 1\}
$$

which means in each tournament, the reverse rank of talent are exactly same to the rank of college seats. The lower θ is, the higher s and p people get.

Without quota, for individual with talent θ, I get the number of people with type lower than his type in the equilibrium is:$1 - F_{all}(\theta)$. In the equilibrium,

$$
P_i^w(s^*(\theta)) = (1-\mu)\, e^{-\lambda_1 \theta} + \mu e^{-\lambda_0 \theta}
$$

I get $P_i^w(s^*(\underline{\theta})) = 0$, and $P_i^w(s^*(\bar{\theta})) = 1$. In the final equilibrium $s^*(\theta)$ is:

$$
s^*(\theta) = \left[(1-\mu)\, e^{-\lambda_1 \theta} + \mu e^{-\lambda_0 \theta}\right] \theta^{\frac{-1}{\alpha}}
$$

With quota, assume the seat is given to the group 0, with distribution

$$F_0(p) = \mu p^{\tau};$$

and

$$F_1(p) = p - \mu p^{\tau}.$$

$\tau > 1$ means 0 is the disadvantage province with higher portion of students ending up with lower p than 1. Otherwise, 0 is the advantage province with higher portion of students ending up with higher p than 1. For an individual with talent θ, the number of people with type lower than his type in the equilibrium is: $1 - F_i(\theta)$. I get:

$$P_0^q(s^*(\theta)) = e^{-\lambda_0 \theta} = p^{\tau}$$

and

$$P_1^q(s^*(\theta)) = e^{-\lambda_1 \theta} = \frac{p - \mu p^{\tau}}{1 - \mu}$$

In the equilibrium, I get:

$$s_0^*(\theta) = e^{-\frac{\lambda_0 \theta}{\tau}} \theta^{\frac{-1}{\alpha}}$$

In the all pay auction game, all individual's payoff is 0 in the equilibrium. ■

Proof of proposition 11 **Proof.** Define $\Delta s(\theta) = s_0^*(\theta) - s^*(\theta) = \left(e^{-\frac{\lambda_0 \theta}{\tau}} - \left[(1-\mu) e^{-\lambda_1 \theta} + \mu e^{-\lambda_0 \theta}\right]\right) \theta^{\frac{-1}{\alpha}}$.

If $\tau > 1$, I get $\Delta s(\theta) > 0$,every one in group 0 will take more s than the case without quota case.

If $\tau < \frac{\lambda_0}{\lambda_1}$,I get $\Delta s\left(\theta\right) < 0$, every one in group 0 will take less s than the case without quota case.

If $\tau \in \left[\frac{\lambda_0}{\lambda_1}, 1\right]$, I get $\lim_{\theta\to 0} \Delta s\left(\theta\right) = +\infty$ and $\lim_{\theta\to\infty} \Delta s\left(\theta\right) = -\infty$. In addition, Δs is a continuos function with $\frac{\partial \Delta s(\theta)}{\partial \partial} < 0$ for $\forall \theta \in [0, \infty]$. $\exists \theta^* \in [0, \infty]$, such that $\Delta s\left(\theta^*\right) = 0$. I get $\Delta s\left(\theta\right) > 0$, if $\theta < \theta^*$ and $\Delta s\left(\theta\right) < 0$, if $\theta > \theta^*$. Anyone one in group 0 with θ larger than θ^*incentive will take more s than the case without quota case. anyone one in group 0 with θ smaller than θ^*incentive will take less s than the case without quota case. ■