Internal migration and extended families in China

Ling Zhong*

Cheung Kong Graduate School of Business and IZA

May 30, 2024

Abstract

Internal migration from rural to urban areas can have large welfare effects on the migrants and their extended families. In China, temporary migrants leave family members behind and support them with remittances from the urban area, while the grandparents look after children in the rural area. I study how internal migration opportunities for rural workers affect their extended family's behavior and welfare. Using micro-level data from China, I develop and estimate a model featuring three generations and migration choices. The model features an informal limited-commitment contract over child care, financial transfers, and commitments to elder care. I find that the informal contract generates welfare gains. I then use the model to evaluate hypothetical policies targeted at specific generations. Counterfactual outcomes imply that policies affect all household members via the informal contract. Extended models incorporating altruism underperform the baseline model in matching the moments, indicating that exchange motives predominantly drive intrahousehold behaviors in rural Chinese households.

Keywords: Internal migration, China, intrahousehold, three generations, left-behind children

JEL Codes: D1, J1, R2, O1

---

*Cheung Kong Graduate School of Business, Oriental Plaza Tower E3, 3F, One East Chang An Avenue, Beijing, 100738, China. E-mail: lingzhong@ckgsb.edu.cn. I thank Joseph Altonji, Emily Oster and Cormac O'Dea for their invaluable support and guidance. Thanks also to Elyse Adamic, Isaiah Andrews, Joy Chen, Christian Dustmann, Tianyu Fan, Ran Gu, Naijia Guo, John Eric Humphries, Louise Laage, Costas Meghir, Yusuke Narita, Eddie Ning, Jesse Shapiro, Joseph Shapiro, James Thomas, Jaya Wen, Chenggang Xu, Leslie Young, Xiaoyi Zhang, Xiaodong Zhu and seminar participants at AASLE, Brown University, Cheung Kong Graduate School of Business, Chinese University of Hong Kong, McMaster University, Monash Business School, New York University, Peking University, Renmin University, Tsinghua University, Yale University for their helpful comments. I gratefully acknowledge financial support from the Cowles Foundation Structural Micro Program. All remaining errors are my own.

1 Introduction

The rural-to-urban migration of workers in China, involving more than 250 million individuals, is the largest migration in human history. This massive migration movement has significantly contributed to China's economic growth over the past two decades (see, e.g., Tombe and Zhu (2019)). However, the impact of migration extends beyond the economic gains for the migrants themselves to affect their families, often left behind in rural areas. Rural-urban migration increases the income levels of migrant workers, but imposes considerable costs on their families. It has separated over 60 million children from their parents, leading to poor educational outcomes and potential long-term human capital inequalities in China's labor force (Rozelle, 1994). Furthermore, approximately 40 million elderly are at risk of lacking private care typically provided by adult children.

To fully assess the welfare implications of rural-urban migration in China, this paper explores several critical questions: how do multigenerational rural families make decisions about migration, financial transfers, children's education, and elder care? How do financial and institutional constraints affect these decisions? What policies can improve the welfare of migrants and their extended families?

In this paper, I develop and estimate a three-generational household model with migration choices to address these questions. The household includes a non-migrating grandparent, a parent with the option to migrate, and a child who may either migrate with the parent or left behind with the grandparent. Figure 1 illustrates the intricate interactions among the three generations. In the model, the rural parent's choices include not migrating, migrating alone, or migrating with the child. The parent also decides on investments in the child's education, financial support, and elder care for the grandparent, with the latter prohibiting migration. The grandparent, facing illness and mortality risks, decides on childcare and healthcare consumption. The child, while passive in decision-making, is modeled as an educational production function with school dropout risks.

The model highlights both the positive and negative effects of migration on each gen-

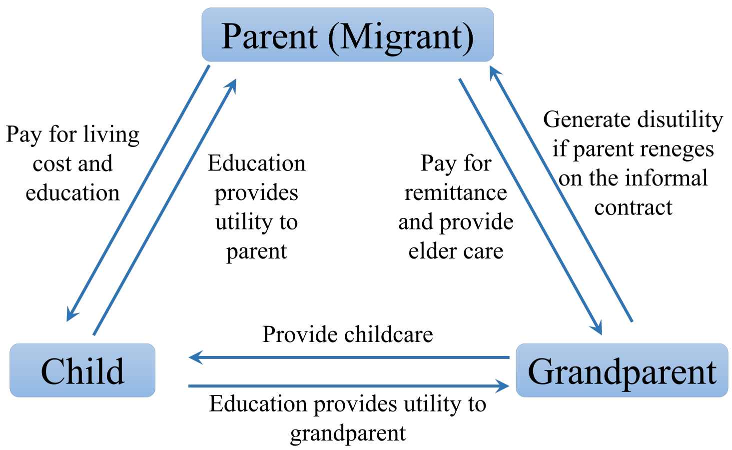


Figure 1: Interactions between the migrant

eration within a rural family: My reduced-form analysis suggests that children left behind have poorer school performance and that ill grandparents suffer higher mortality rates due to inadequate care when parents migrate. Compared with parents in households with healthy grandparents, parents are less likely to migrate to care for ill grandparents, especially if these grandparents had previously provided childcare, suggesting the influence of grandparents' past behaviors on parents' decisions. To capture the intertemporal dynamics between migrating parents and left-behind grandparents, I define an informal limited-commitment contract that includes childcare, transfers, and elder care commitments. To accurately formalize decision-making, the model incorporates realistic features of a typical Chinese rural family and employs a novel computational strategy to expand estimation capabilities.

To facilitate quantitative analyses on rural family members' preferences and the potential effects of policy interventions, I estimate the model parameters that reflect parent's preferences for higher income from migration, the child's educational outcomes, and their obligations to grandparents. I find that both parents and grandparents value the child's education highly. In most households, the guilt associated with reneging on informal contracts sufficiently motivates parents to provide financial support and care for ill grandparents. Additionally, I find that poorer households are more likely to adopt the informal contract, while wealthier rural families show little incentive to migrate.

I perform counterfactual experiments on the model to evaluate the impacts of various policies on migrants and their extended families. The first policy examined is an expansion of health insurance, which increases the government's contribution to healthcare co-payments. This policy generates a welfare trade-off: while some grandparents gain better access to healthcare and withdraw from informal contracts, parents face increased costs from raising children in urban areas due to the loss of childcare options in their rural homes. As a result, this policy reduces both the number of migrant workers and the fraction of left-behind children. I also evaluate policies aimed at decreasing educational costs and childcare time. The policy that reduces the time associated with childcare emerges as more effective in improving educational outcomes for children. The experiments show that policies targeting children alone do not significantly resolve the issue of left-behind children due to intrahousehold behaviors.

To show that the model accurately captures the key dynamics of intrahousehold behaviors in rural Chinese families with migration opportunities, I extend the model to include altruism as a partial or full motive for intergenerational behaviors such as financial transfers and private care. Re-estimating the extended model and comparing the moments prediction reveals that the baseline model, which emphasizes the exchange motive, produces the most accurate predictions. The analysis further indicates that the exchange motive is the primary driver of financial and service-related behaviors in rural Chinese households with migrants in China, in line with studies on motives for intrahousehold dynamics in other developing countries (Cox et al., 1998; Kazianga, 2006).

This paper makes an important contribution as the first to develop a three-generational household model with migration choices. First, working with a three-generational model marks a substantial step forward in analyzing internal migration in developing countries like China. This model formalizes the economic and behavioral trade-offs typical rural families encounter and accommodates complex intergenerational dynamics. For example, it captures the simultaneous impact of a parent's migration decision on both the child's education and the grandparent's healthcare. Second, this paper facilitates counterfactual simulations to make policy predictions on the flow of migration and the welfare of each generation in the migrants' extended families. For effective policy analysis concerning internal migration in China, having a model that comprehensively addresses first-order microeconomic factors —such as migration rates, consumption levels, children's education, and elderly healthcare —is essential.

This paper fits in three literatures. It contributes primarily to the literature on structural models on migration. Existing models in this area typically focus on households with one or two generations (Barham and Boucher, 1998; Thom, 2010; Gemici, 2011; Görlach, 2016; Morten, 2016) or on the labor market outcomes of migrants (Bayer and Juessen, 2012; Yoon, 2017; Gai et al., 2021). This study advances the literature by constructing a three-generational household model, which unveils many intrahousehold effects previously unobservable in less complex frameworks.

This paper also contributes to the literature on household models with intergenerational behaviors. Prior works have focused on financial transfers and elder care between the elderly and their adult children (Becker, 1974; Altonji et al., 1997; Costa, 1999; Pezzin et al., 2007; Wiemers et al., 2017; Barczyk and Kredler, 2018; Mommaerts, 2018), as well as on consumption smoothing for young children supported by their parents (McElroy, 1985; Buck and Scott, 1993; Rosenzweig and Wolpin, 1993; Ermisch and Di Salvo, 1997; Ermisch, 1999; Goldscheider and Goldscheider, 1999; Manacorda and Moretti, 2006; Kaplan, 2012; Sieg et al., 2023). My research expands this scope by incorporating the interplay among all three generations, with a particular emphasis on migration decisions. Within the field of household economics, a particular strand of theoretical and empirical research investigates the motives behind intergenerational behaviors, building on the foundational work of Becker (1974) and Cox (1987). Numerous studies emphasize the altruism motive (Altonji et al., 1997; Foster and Rosenzweig, 2001; Barczyk and Kredler, 2018), while others suggest these behaviors are driven by exchange motives (Cox et al., 1998; Kazianga, 2006) or a combination of both (Andreoni, 1989). However, most empirical research, including studies like Wilhelm (1996) and Kazianga (2006), have found little evidence to support altruism motives. In this paper, the baseline model suggests the intergenerational behaviors are driven purely by the exchange motive. I assess alternative mechanisms by comparing models based purely on exchange, purely on altruism, and on combined motives. The findings indicate that the pure exchange model most effectively captures the intergenerational behaviors in Chinese households with migrants.

Finally, this paper relates to the literature on internal migration in China. This field has extensively explored migrants' labor market integration and urban living conditions (Zhao, 2003; Song et al., 2008; Song and Zenou, 2012; Lagakos et al., 2020; An et al., 2024), and a broad array of migration behaviors (Massey, 1990; Li and Zahniser, 2002; Zhao, 2002; Zhu, 2002; Taylor et al., 2003; Du et al., 2005; Huang and Zhan, 2005; De Brauw and Rozelle, 2008; Démurger and Xu, 2011; Mullan et al., 2011) and the dynamics within migrants' households (Brauw et al., 1999; Zhao, 1999; Zhu, 2002; Biao, 2007; Mu and Van de Walle, 2009; Qin and Albin, 2010; Chang et al., 2011). To the best of my knowledge, this is the first paper to model the interaction between migration behavior and extended family characteristics of migrants, offering a novel framework for policy analysis through an intrahousehold game equilibrium approach.

The paper proceeds as follows. Section 2 introduces the datasets I use. Section 3 describes the economic and behavioral characteristics of rural households with migration opportunities. Section 4 presents a household model with a migration option and an informal contract option. Section 5 discusses the estimation procedure and the model estimates. Section 6 uses the model estimates to evaluate the role played by the informal contract. Section 7 simulates various policies targeting on the grandparents or the children, and discuss the policy effects on each generation and the migration decisions. Section 8 discusses alternative explanations for intergenerational behaviors by comparing the baseline model with exchange motives against models with altruism. Section 9 concludes the paper.

2 Data sources

I use data from five micro-level surveys conducted in China for two purposes. First, I describe the economic environment of internal migration and the migrants' households to obtain estimates for parameters I can estimate directly. Second, I construct the data moments that will be used to identify the rest of the model's parameters.

I briefly discuss the relevant features of each data set, and the role that they play in my analysis.[1] I also provide additional details on the data processing in Appendix B.

China Family Panel Studies (CFPS): As a rich panel data set collected by the Institute of Social Science Survey (ISSS) of Peking University in China, the CFPS data has community-, household- and individual-level information on demographics, economic activities, education, and employment, as well as health and nutrition information. The study was launched in 2010 and follow-up data were collected in 2012. It contains 56,121 rural individuals from 13,355 households with a good follow-up rate of 90%. The CFPS fits my needs in the following aspects: first, the panel data structure follows everyone in the household regardless of their residential locations, enabling identification of rural households with migrants. Second, the CFPS captures various household arrangements including childcare and remittance. Third, it documents family sizes and agricultural income, allowing me to recover the income distribution in rural China.

Rural-Urban Migration in China (RUMIC): The RUMIC 2008-2009 data targets internal migrants with a survey specifically designed for rural-to-urban migrants. I treat it as cross-sectional data. The 2009 rural migrant survey provides the most relevant information for this paper. It contains 5,426 individuals over 15 years old. The data provides information on migration movements, migrants' labor market conditions, remittance, and left-behind children.

China Health and Nutrition Survey (CHNS): One of the longest-running panel

[1] Although some of the data sets include follow-up surveys from 2018 to 2024, I have confined the time frame for the reduced form analysis and the data moments used in the structural estimation to the period 2010-2016. This restriction ensures consistency in the policy environment across all data sources and years.

data sets in China (1989 to 2011), CHNS includes around 19,000 individuals from 4,400 households. The survey asks for detailed health and nutrition information, in addition to collecting precise health outcomes by offering a physical examination to all survey participants. I use variables related to time allocation to capture the variation in time consumption associated with work, agricultural production, and daily care for family members. I also take advantage of its emphasis on health status and healthcare to estimate parameters related to illness and mortality risks and private consumption of medical services.

China Health and Retirement Longitudinal Study (CHARLS): The CHARLS data, as a sister study of the American Health and Retirement Study (HRS), surveys people over 45 years old to support research on the aging population in China. It is a biannual panel data launched in 2011 with about 17,500 interviewees from 10,000 households and biennial follow-ups. The data contains variables on family structure and transfers, health status and expenses, labor market participation and income, and consumption and savings. The unique information provided by the CHARLS data is its detailed records about the financial transfers and daily care from the elderly to their grandchildren and from the adult children to the elderly. It provides useful statistics on correlations among various intergenerational behaviors and remittances.

China Household Finance Survey (CHFS): The CHFS data is a cross-sectional survey with two waves focusing on household financial and physical assets, income, expenditure, and intergenerational transfers. The published waves include a total of 3,002 rural households and 5,434 urban households. The expenditure variables allow me to estimate the difference in the price levels in rural and urban areas. The consumption records by category also enable estimation of subsistence consumption level net of spending on education and healthcare for each location.

Summary: To show how the five datasets together contribute to my paper, I summarize the usage of each data source by components of my research topic in Table A2. Although a single comprehensive dataset that would allow for a multivariate regression with adequate controls does not exist, the collection of datasets used in this paper is sufficient to provide information on all important aspects of a three-generational Chinese household with migrants.

3 Facts about migration and extended families

In this section, I document empirical facts that aim to align the key features of the household structure of the migrants and the social and economic environments they face. Specifically, I present facts about rural-to-urban migration in China in Section 3.1, Chinese rural households in Section 3.2, and how these rural households behave under the massive migration movement in Section 3.3. They are based on reduced-form analyses of my own using various datasets summarized in Section 2 and the results of other papers on these topics. I then introduce the concept of the informal contract as a tool to connect intergenerational behaviors in rural Chinese households with migrants.

I summarize the most relevant facts that will be addressed in my model: (1) Migration decision is a decision between two distinct economic environments. (2) Parents' migration decisions depend on the education of the child and the health of the grandparent. (3) Left-behind children have worse education outcomes. (4) Grandparents rely on parents for daily care and funding for medical treatment. (5) The informal contract between the migrating parent and the left-behind grandparent has two main components: the first is pecuniary, in which the parent with migration experience provides financial support to the left-behind extended family; and the second is non-pecuniary, in which the grandparent provides childcare in exchange for the parent's provision of daily care when she is ill.

3.1 Internal migration in China

Migration is mostly temporary for rural residents without a college degree. Rural-to-urban migration within China was made possible as the government gradually relaxed the

Hukou (household registration) system in the mid-1990s (Zhao, 2005). Currently, this policy change is used as a means to balance the excess rural labor supply with the excess urban labor demand. However, the migration had to be temporary.[2] Permanent settlement in the urban area, or joining the urban Hukou, remains extremely difficult for rural migrants, regardless of how long they live in the urban area (Wang, 2004). The temporary nature of the migration movement is further reinforced by the migrants' highly limited access to the urban public service system, including public schools and healthcare (Müller, 2016; Zhou and Cheung, 2017).

Most rural migrants are employed workers in the urban area. My model considers high-income job opportunities as the sole incentive to migrate to the cities. According to RUMIC 2009 data, 85% of rural parents stay in urban areas for this reason.

Rural and urban areas have very different labor markets and living costs. In choosing between rural and urban settings, a potential migrant is implicitly deciding between two very different sets of jobs and expenses.[3] While the urban labor market offers a much higher income, it also requires longer hours. Urban life is also more expensive in terms of price level, as well as offering limited access to public health insurance coverage. Migrants have the choice to bring their children to the urban area. But migrants' children go to under-regulated private schools with higher tuition than rural schools and lower teacher quality than urban public schools (Li et al., 2010).

3.2 Chinese rural households

Many people live in three-generational households. As discussed in Hu and Peng (2015), 57% of elderly Chinese (aged 65 and above) live with their children and grandchildren.

[2] Changing Hukou from "rural" to "urban" is extremely hard and rare for rural people. I abstract from it in my model and assume that nobody changes their Hukou. Therefore, in my model, all migrants must return to the rural area in the end. Song (2014) provides a comprehensive summary of the Hukou system.

[3] Table A3 presents a list of location-specific economic factors that are key to the migrants. On average, the nominal annual income are 23,428 RMB in urban and 2,702 RMB in rural, the weekly hours at work is 66 in urban and 49 in rural, and the annual living expenses are 16,464 RMB in urban and 2,389 RMB in rural.

This co-residence pattern is more common in rural areas than in cities.

Expenditure concentrates on food, healthcare and education. Table B2 in Appendix C disaggregates the annual expenditure of rural households by spending category. It shows that 68% of household-level spending is on food and housing, healthcare, and education. I model household consumption into a continuous choice on daily consumption and two discrete choices on grandparents' healthcare and children's education.

Most grandparents rely on their adult children when they are ill. The 2005 Chinese Longitudinal Healthy Longevity Survey reports that 63% of ill grandparents are taken care of by their adult children.[4] The 2015 National 1% Sampling Survey reports that 46% of ill grandparents financially depend on their children.

Children who are taken care of by grandparents have poorer educational outcomes. Table A4 shows that children's school enrollment status and course performances significantly depend on the years of schooling completed by the child's primary caretaker, while the relationship between the children and their primary caretaker does not matter. Children who are primarily cared for by grandparents have a worse educational outcome because rural grandparents are less educated than the children's parents.[5]

3.3 Rural households with migrants

Migrating workers earn more in the urban area, and their rural household structure extends the effect of the parent's migration decision onto the welfare of the grandparents and children. These chain effects, in turn, become important factors in the parent's migration decision. In this section, I present descriptive statistics about rural households with migrating parents.

Migration decisions depend on migrants' wealth and demographics, children's

[4]Note that in all the datasets I use, health status is divided into five categories, from "very good" to "very ill". I combine the "ill" and "very ill" to define "illness" in my empirical analyses and calculation of the moments that I match my model to in the data. In my model as well as my empirical analyses, "illness" means not being able to provide childcare, and requiring private care in the grandparent's daily life.

[5]On average, grandparents' years of schooling are 2.58 years shorter than parents' (my analysis, CFPS 2010-2014). This gap lowers the odds ratio of the child's enrollment by 29%, the math score by 0.08 points, and the Chinese score by 0.05 points. See Appendix C.6 for details.

education and grandparents' health. Three key determinants relate to my paper. First, rural adults are more likely to migrate out if they are young, less educated, or come from poor households (Zhao, 1999; Yang, 2000; Li and Zahniser, 2002; Du et al., 2005). Symmetrically, they are more likely to remain in rural areas when they are old (Zhao, 2002; my analysis, RUMIC 2008 and 2009). Second, rural parents migrate more if their children are enrolled in school after controlling for the children's age (my analysis, CFPS 2010-2014).[6] Third, when grandparents are ill, rural parents are more likely to stay in the rural area only if the grandparents have provided childcare in the past (my analysis, CHARLS 2013 and 2015). In Appendix C, I provide the results of my regression analyses and a table summarizing the data sources and key findings of other papers on this topic.

Therefore, I differentiate rural families by the per capita agricultural income of the household; use the children's age (rather than the parent's age) as the time index in the model; and model the heterogeneity in children's education enrollment and grandparent's health, in order to account for the key factors in the parent's migration behaviors.

Transfers from the parents to the grandparents depend on migration, childcare, and health. Remittance from migrants to their left-behind families is a consumption-smoothing and welfare-improving device (Katz and Stark, 1986; Massey, 1990; Zhao, 2002; Taylor et al., 2003; Li, 2006; Biao, 2007; Zhu and Luo, 2010; Rong et al., 2012). I construct a panel data set from CHARLS 2008-2015 to focus on parents' migration, grandparents' childcare, and health status. First, the financial transfer between rural parents and grandparents is mostly one-directional. Thirty-nine percent of grandparents receive money from their children, while only 2% give money to their children. Second, I run a Logit regression for whether parents send money to grandparents and an OLS regression for how much money they send, with controls on gender and age of the parent and grandparent (Table B6 in Appendix C). Parents with migration experience are more generous, even after they

[6] One interpretation is that forward-looking parents are more incentivized to migrate for higher wage income to fund future education costs for the children, while parents of children who have dropped out of school do not have such future expenditure.

return to the rural area. Parents' financial transfer behavior is also positively associated with the grandparents' childcare experience and illness.

Left-behind children suffer from being apart from their parents. Left-behind children come from 74% of households with migrants (CFPS 2010-2014) and account for 28% of all rural children (Jia and Tian, 2010). The RUMIC survey asks the migrants for "the main reason why your children do not live with you". The top three reasons are high cost of living, lack of childcare, and high tuition (Table B7 in Appendix C). When children live with their parents, regardless of where they live, less than 20% are cared for by grandparents. If the children are left behind, 71% of them are cared for by grandparents (CFPS 2010-2014).

Left-behind children in China have been studied by psychologists, education researchers, economists, and nutrition scientists. While research has shown that left-behind children are disadvantaged in psychological and nutritious conditions (De Brauw and Mu, 2011; Ye and Lu, 2011; Wen and Lin, 2012; Su et al., 2013), the educational consequences of leaving the children behind remain unclear. My paper models two main effects of migration on left-behind children: (1) The positive effect of migration on financial resources for education as suggested by Meyerhoefer and Chen (2011). (2) The negative effect of grandparents' care for children on their educational outcome as suggested by Wen and Lin (2012).

3.4 Informal contract

The behavioral correlation between the parent's migration and grandparents' childcare and health status can be considered as an implicit contract between the two agents. The contract obliges the healthy grandparent to take care of the child if the parent chooses to migrate to the urban area while leaving the child behind. It also obliges the migrating parent to send remittance to the grandparent and come back to the rural area when the grandparent is ill, to provide daily care and assist with the grandparent's healthcare expenses.

The agents, especially the parent, are incentivized to commit to the contract to maintain their reputation in the local society in China. A rural individual's reputation in a social network influences trust, reciprocity, and mutual help from others. A good reputation promotes the person's long-term economic status, health, and well-being (Chen and Silverstein, 2000; Ma, 2002; Yip et al., 2007). Most agents have to live in the rural area after temporary migration to the cities.[7] Thus, migrants have no escape from the consequences of their reputation in their home villages.

4 Model

In this section, I propose a model featuring a sequential game between a parent and a grandparent in a typical three-generational rural Chinese household. It highlights their interactions through monetary transfers and private care exchanges, and organizes these behaviors into an informal contract of limited commitment.

Model overview: Each household consists of an active grandparent, an active parent, and a passive child. The active agents make sequential choices. Initially, the grandparent decides whether to offer the informal contract. Subsequently, after the child's birth, the parent first determines migration and informal contract status, transfers, consumption, and the child's education. The grandparent then decides on consumption and healthcare spending.

The parent's utility depends on (a) his personal consumption and leisure, (b) child's educational outcomes, and (c) commitment to the informal contract. The grandparent's utility depends on (a) her personal consumption and leisure, (b) grandchild's education, (c) her health status and healthcare expenditures. The household faces risks related to the grandparent's health, child's school dropout potential. Households vary by rural income per capita determined by family endowment. Government can affect these dynamics through subsidizing the costs on healthcare and school tuition, and providing social security for those below the poverty line.

[7] Over 70% of migrants stay for less than 10 years in an urban area (my analysis, RUMIC 2009). I provide empirical evidence on the duration of migration in Appendix C.4.

4.1 Household structure, utility and budget constraints

Demographics: The model starts at the birth of the child. Time in the model is discrete, denoted by the child's age, and divided into five periods by the child's potential education level with varying length. Periods 1 and 2, each spanning 6 years, correspond to pre-school and primary school ages, respectively. Periods 3 and 4, each lasting 3 years, cover the middle and high school ages. The final period, Period 5, lasts 4 years, covers the college age when the child may still financially depend on her rural family.

The model's basic unit is a three-person household: a grandparent, a parent, and a child. Households are categorized into J types, each indexed by j and characterized by a uniform annual agricultural income per capita, A_j.[8] These types, determined before the start of the model, imply different levels of the flow income of non-migrants.

The three members of the household are denoted by C for child, P for parent, and G for grandparent. The state of a household is summarized by the state vector X_t:

$$X_t = \left\{contract^G, contract_t^P, urban_t^P, urban_t^C, s_t^G, s_t^P, h_t, g_t, enroll_t, edu_t\right\} \tag{1}$$

in which $\left(contract^G, contract_t^P\right) \in \{(0,0), (1,0), (1,1)\}$ are the contract status of the grandparent and the parent with $accept = 1$ and $reject = 0$. The grandparent makes a one-time decision when the model starts, while the parent may opt into the contract at any feasible period. $\left(urban_t^P, urban_t^C\right) \in \{(0,0), (1,0), (1,1)\}$ indicate the locations of the parent and the child with $urban = 1$ and $rural = 0$. The child may migrate only if the parent migrates. s_t^G and s_t^P are wealth of the grandparent and the parent. $h_t \in \{healthy = 0, ill = 1, death = 2\}$ is the health status of the grandparent. g_t is the level of guilt of the parent if he reneges on the informal contract. $enroll_t$ is an indicator for whether the child attends school. edu_t is

[8] I set the rural labor income as a constant, which does not depend on the number of adults in the rural household. This assumption is supported by findings from the CFPS 2010-2014 data, which I provide in Appendix C. It also aligns with Rong et al. (2012), who find that the rural household's income is a roughly constant amount per capita. Thus, agricultural income per capita for rural residents is not affected by whether the household has migrants. The intuitive interpretation of this fact is that the current binding constraint in agricultural production is human capital instead of land or production equipment, as the population density in rural China has decreased in the last several decades due to the One Child Policy and internal migration.

the education attained by the child, represented by years of schooling completed by the end of the current period.

Child: The child in this model is passive. The child consumes on two goods: education, with tuition at price $tuition_t$ paid by the parent; and other goods with a total consumption of c_t^C paid for by the co-residing adult —either the parent or the grandparent. The child's role is simplified into an education production function, and the educational outcome affects the utilities of the adult agents in the household.

Figure 2 illustrates the evolvement of the child's education, represented by enrollment status $enroll_t$ and attainment edu_t. Exiting education is irreversible. If the child was not enrolled in the previous period, she cannot be enrolled in current period, leaving her attainment unchanged. If the child was previously enrolled her parent pays for the tuition, her enrollment continues. Enrolled child faces a positive dropout probability $p_t^{dropout}$. If the child drops out, attainment does not update; otherwise, her educational outcome progresses.

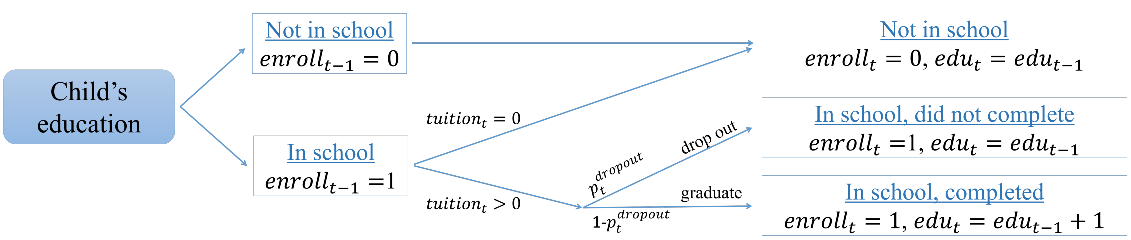


Figure 2: Child's education

The child's education production function captures several effects of parental decisions on education outcomes evident from Section 3: the continuous enrollment requirement makes education investment an intertemporal decision. Categorizing tuition payment as a separate consumption choice highlights its importance in rural household budgets. Dropout risk increases over time but is mitigated by the caretaker's educational level.

The child's consumption of other goods is a fixed portion of the caretaker's consumption, determined by the OECD equivalence scale for China (Liu and Li, 2011). Therefore, the caretaker allocates

$$c_t^{\text{child}} = \pi \times c_t^{\text{caretaker}} \tag{2}$$

to the child's daily consumption, where π is the equivalence scale parameter.

Parent's preferences and budget constraint: The parent has time-separable utility preferences. His flow utility U_t^P depends on his private utility and two factors representing his altruism towards the child and the grandparent:[9]

$$\begin{aligned} U_t^P &= u^P\left(\ \tilde{c}_t^P, l_t^P, enroll_t, g_t\ \right) \\ &= \underbrace{u^{private}\left(\tilde{c^P}, l^P\right)}_{\text{(private utility)}} \times \underbrace{\left(1 + u^{edu}\left(enroll_t, t\right)\right)}_{\text{(child's education)}} \times \underbrace{\left(1 - g_t\right)}_{\text{(contract with grandparent)}} \\ &\quad + \delta \times \left(1 + u^{edu}\left(enroll_t, t\right)\right) \times \left(1 - g_t\right) \end{aligned} \tag{3}$$

in which c_t^P is the parent's consumption on private goods[10] and l_t^P is leisure; $u^{private}(c,l) = \frac{\left(\left(c^{\theta}l^{1-\theta}\right)^{1-\gamma}\right)-1}{1-\gamma}$ is the private utility function; $u^{edu}(enroll_t, t) = \varphi \times \mathbb{1}_{enroll_t=1} \times \mathbb{1}_{t=4}$ is the utility gain from the child's high school enrollment;[11] g_t is the utility loss if reneging on the informal contract, so $g_t = 0$ when the contract is fulfilled.

Fulfilling the informal contract has two effects on the parent's utility. First, they directly cause utility gain/loss, they are added to the parent's private utility by a factor δ. Second, they change the private utility levels, so they are multiplied with the private utility term. The addition component represents the non-substitutability of components other than daily consumption and leisure. When $\delta = 0$, parents' utility gain/loss from the altruistic terms are proportional to their private utility.

The parent's intertemporal budget constraint is

$$\begin{aligned} s_{t+1}^P = (1+r)\left(s_t^P + income\left(urban_t^P\right) + Bequest \times \mathbb{1}_{g_t=0} \right. \\ \left. - c^{hh}\left(c_t^P, \mathbb{1}_{urban_t^P = urban_t^C}\right) - tuition_t - Tr_t\right). \end{aligned} \tag{4}$$

[9]His terminal utility function is defined later in equation (15).

[10]To account for the much higher price level in the urban area, consumption is normalized by the price ratio between the urban and rural areas before it enters the parent's utility function. i.e. $\tilde{c}_t^P = c_t^P \times \frac{p^{rural}}{p^{urban}}$ if $urban_t^P = 1$, in which c^P is the individual level consumption of the parent.

[11]Compulsory education in China is nine years. Rural children who quit education after middle school can help with agricultural production or work in the urban area. Therefore, investing in a time- and financially-costly and non-compulsory education period should yield a welfare gain.

in which s_t^P is the wealth of the parent; $income\left(urban_t\right)$ is the labor income that equals to A_j, the constant flow agricultural income for non-migrants if $urban_t = 0$, or w_t, the wage rate in the urban labor market if $urban_t = 1$; $Bequest$ is the bequest from the grandparent;[12] $c^{hh}\left(c_t, coreside_t\right) = c_t \times \left(1 + \pi \times coreside_t\right)$ is the household consumption on commodities, accounting for the total of the adult and potentially co-residing child ($coreside_t$ is a dummy indicating that the child lives with the parent), following equation (2); $tuition_t$ is the child's education cost; and Tr_t is the financial transfer to the grandparent.

Grandparent's preferences and budget constraint: The grandparent's flow utility U_t^G values private utility, the child's education, and her health status and consumption on healthcare:

$$\begin{aligned} U_t^G &= u^G\left(\ c_t^G, l_t^G, enroll_t, c_t^h\ \right) \\ &= u^{private}\left(c^G, l^G\right) \times \left(1 + u^{edu}\left(enroll_t, t\right)\right) \times HC_t \\ &\quad + \delta \times \left(1 + u^{edu}\left(enroll_t, t\right)\right) \times HC_t \end{aligned} \tag{5}$$

in which c_t^G is consumption on private goods and l_t^G is leisure; HC_t is the factor on health status and healthcare as specified below:

$$HC_t = \mathbb{1}_{h_t=0} + \mathbb{1}_{h_t=1} \times \left(1 - \eta_1 + \eta_2 \times \mathbb{1}_{c_t^h>0}\ \right) \tag{6}$$

in which η_1 is the utility loss of being ill and η_2 is the utility gain from healthcare.[13] So HC_t equals 1 when the grandparent is healthy and falls below 1 when she is ill, with specific values depending on her healthcare expenditures; δ is the same utility gain from the child's education as in parental utility.[14]

The grandparent's intertemporal budget constraint is:

$$s_{t+1}^G = (1+r)\left(s_t^G + income(0) \times \mathbb{1}_{h_t=0} + Tr_t - c^{hh}\left(c_t^G, \mathbb{1}_{urban_t^P \neq urban_t^C}\right) - c_t^h\right) \tag{7}$$

[12]The role of g_t in the budget constraint, associating with the bequest, is to incentivize contract commitment, such that the parent does not receive any bequest if he reneges on the contract, or if he does not enter the contract.

[13]The consumption on healthcare is discretized into a homogeneous annual cost paid by the grandparent, i.e. $c_t^h \in \left\{0, \bar{c}^h\right\}$ and another lump sum cost right after the grandparent's death, paid for primarily out of the grandparent's bequest and then out of the parent's savings. I describe this specification in detail in Section 4.2.

[14]I also estimated an alternative model with two different additive factors for the parent and grandparent. The model estimates suggest that the values of the two additive factors do not differ significantly.

which is different from the parent's budget constraint in three aspects: (1) the grandparent is the receiver of the transfers, (2) she never migrates and can only earn from agricultural production when she is healthy, and (3) she may spend c_t^h on healthcare when ill.

4.2 Economic and social environment

Labor market and earnings: The parent's migration decision determines access to two distinct labor markets. Non-migrants work on household land with heterogeneous income indexed by type j but fixed hours. Migrants access the urban labor market, earning a uniform wage $\bar{w}$ when employed. They are assumed employed full-time from periods 1 to 4, with potential unemployment in period 5 at a probability of pr^{ump}, leading to an expected wage of $w_t = \bar{w} \times (1 - pr^{ump} \times \mathbb{1}_{t=5})$.[15] In the model, the migration decision, $urban_t^P \in \{0, 1\}$, chooses between two wage-hour pairs: (A_j, T_{rural}) and (w_t, T_{urban}). The grandparent, who does not migrate, earns A_j when healthy and nothing when ill.

Daily consumption: Rural and urban areas have segregated goods markets, so I use a price ratio $\frac{p^{urban}}{p^{rural}}$ to convert nominal consumption into actual living standards.[16]

Tuition: As discussed in Section 3.2, education costs, including tuition and fees, differ by location and level of schooling. Accordingly, the model sets education investment as a binary choice between a standard tuition rate, $\overline{tuition}_t(urban_t^C)$, and zero.

Healthcare: The model assumes that only the ill grandparent consumes healthcare. When she becomes ill, she decides whether to spend on a homogeneous annual cost $c_t^h \in \left\{\bar{c}_t{}^h, 0\right\}$. To account for the increased medical expenses in the grandparent's final year, households face a lump-sum cost immediately after the grandparent's death.[17]

Leisure: The leisure time of the parent l_t^P and of the grandparent l_t^G are defined as the

[15] Consequently, the heterogeneity in the migrants' earnings in my model is consistent with the distribution of monthly income of the migrants I summarized from the RUMIC 2009 data (Appendix C.1).

[16] For further details on the ratio's definition and estimation, see Section 5 and Appendix C.

[17] The lump sum cost is deducted from the grandparents' savings when she dies. If the cost exceeds her assets, the remaining cost is transferred to the parents as a debt. This extra healthcare cost is implemented in the model but left out of the budget constraint equations of the parents and grandparents for clarity. I provide supporting evidence of the specifications on the homogeneous annual cost and the lump sum cost in Appendix D.4.

number of hours net of labor supply, child care, and private care for the ill grandparent:

$$\begin{cases} l_t^P & = T_{total} - \underbrace{T_{rural}\mathbb{1}_{urban_t^P=0} - T_{urban}\mathbb{1}_{urban_t^P=1}}_{\text{(labor supply)}} - \underbrace{T_t^C\mathbb{1}_{urban_t^P=urban_t^C}}_{\text{(childcare)}} - \underbrace{T^G\mathbb{1}_{urban_t^P=0}\mathbb{1}_{h_t=1}}_{\text{(elder care)}} \\ l_t^G & = T_{total} - \underbrace{T_{rural}}_{\text{(labor supply)}} - \underbrace{T_t^C\mathbb{1}_{urban_t^P\neq urban_t^C}}_{\text{(childcare)}} \end{cases} \tag{8}$$

in which T_{total} is the weekly endowment of time, T_{rural} and T_{urban} are location-specific hours spend on working, T_t^C is the period-specific hours spent on childcare, and T^G is the parent's hours spent on private care for ill grandparent.[18]

Uncertainty: The model incorporates two main uncertainties affected by the decisions of parents and grandparents. It assumes grandparents' health changes over time, with transition probabilities depending on their age and whether the parent fulfills the contract by providing daily care. There is no recovery or sudden death, except in the final period.[19] The transition probabilities between health statuses are defined as follows:

$$\begin{aligned} pr(h_{t+1}=1) &= \mathbb{1}_{h_t=0} \times pr_t^{ill} && \text{(illness risk)} \\ pr(h_{t+1}=2) &= \underbrace{\mathbb{1}_{h_t=1}}_{\text{(no sudden death)}} \times pr_t^{death} \times (1 + \underbrace{\rho^{death} \times \mathbb{1}_{urban_t^P=1} \times \mathbb{1}_{contract_t^P=1}}_{\text{(higher mortality risk without elder care)}}). && \text{(mortality risk)} \end{aligned} \tag{9}$$

in which ρ^{death}, the Probit model coefficient, measures how the absence of the parent's care under an informal contract impacts the grandparent's mortality risk.

Second, the education of enrolled children may terminate due to poor performance, proxied by a dropout probability. As discussed in Section 3.3, this probability, determined by the child's education level and whether she is left behind, is defined as:

$$\begin{aligned} pr_t^{dropout} &= pr(\ edu_{t+1} = edu_t \ \Big|\ enroll_t = 1, \mathbb{1}_{urban_t^P \neq urban_t^C}\ ) \\ &= f^{edu}(t, \mathbb{1}_{urban_t^P \neq urban_t^C}). \end{aligned} \tag{10}$$

[18]I set $T_{rural} > T_{urban}$ as noted in Section 3.1. I set $T_1^C > T_2^C = T_3^C = T_4^C > T_5^C = 0$ because 0–5 years-olds demand more time, and adult children demand no time as they leave home for college or work

[19]Supporting evidence of this assumption is provided in Appendix D.3.

4.3 Informal contract and the sequential game

The informal contract involves the grandparent providing childcare in exchange for financial transfers and elder care from the parent. Only healthy grandparents are capable of looking after children and thus can propose the informal contract. Under it, the grandparent commits to caring for the child if the parent migrates, and continues as long as she remains healthy.[20] If the parent accepts the contract, he is obliged to provide remittances and support the grandparent financially and through elder care when she is ill. Reneging on this agreement results in guilt for the parent.

Table 1: Timeline of the model

Stage of the household	Child's Birth	3-generational household
t (age of child) =	0	1 ⋯ 22 T
Choice Variables		
Grandparent	$contract^G$	$c_t^G,\ c_t^C,\ c_t^h$
Parent	-	$urban_t^P, urban_t^C,\ c_t^P,\ c_t^C,\ tuition_t,\ Tr_t$
Resources		
Grandparent	s_0^G	$s_t^G + A_j \mathbb{1}_{h_t=0} + Tr_t$
Parent	0	$s_t^P + A_j(1 - urban_t^P) + w_t urban_t^P +$ Bequest

Sequential game played between the parent and grandparent: The behaviors of the two agents around the informal contract forms a sequential game. Table 1 summarizes the timeline and presents the choice and state variables in each period of the model. I describe the sequence of the actions in detail:

1. When the child is born, the grandparent first decides whether to propose the contract.
2. If the grandparent proposes the contract, the parent may accept the contract at any time, signaled by his decision to migrate without the child.
3. In every subsequent period, the parent moves first by deciding on migration, consumption, and transfer. The grandparent receives transfers, and then moves next to make

[20]Only healthy adults can look after the child. Thus, when the grandparent is ill, the parent has to be the child's co-residing caretaker.

consumption decisions.

4. When the parent migrates alone, the grandparent looks after of the child. The parent sends remittance to the grandparent to cover the child's daily consumption.

5. When the grandparent is ill, the parent must stay in or return to the rural area to take care of the grandparent and pay for the grandparent's total expenditure.

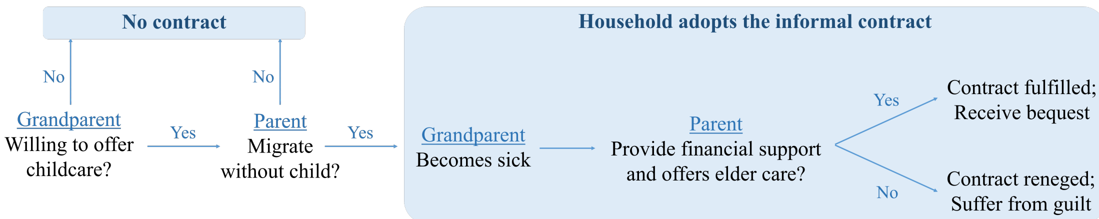


Figure 3: Sequential game and the informal contract

Contract: The contract is binding and irreversible once accepted by both agents, enduring until the end of the model. Figure 3 illustrates how the agents' choices determine the informal contract status. The grandparent fully commits to the contract, denoted by $contract^G \in \{0, 1\}$ where 1 indicates a proposal. In contrast, the parent has limited commitment. The status of the parent's contract, $contract^P_t$, evolves as follows:

$$contract^P_t = \max \{ \underbrace{contract^P_{t-1}}_{\text{(contract held in the previous period)}} , \underbrace{contract^G}_{\text{(grandparent proposes)}} \times \big(\underbrace{urban^P_t - urban^C_t}_{\text{(parent leaves child behind)}}\big) \} \tag{11}$$

Transfer: A parent under the informal contract must pay remittances in two specific scenarios: First, when migrating alone and leaving the child with the grandparent, the remittance should cover the child's daily consumption, calculated as $Tr_t = \pi \times c^G_t$ if $urban^P_t - urban^C_t = 1$.[21] Second, when the grandparent is ill, the remittance covers both daily and medical expenses, expressed as $Tr_t = (c^G_t + \bar{c}^h_t)$ if $h_t = 1$. Additionally, a migrating parent who takes the child, but leaves an ill grandparent behind, must still remit funds covering the grandparent's daily expenses, denoted by $Tr_t \geq c^G_t$ if $h_t = 1$ and $urban^P_t = 1$.

[21] Specifically, the grandparent takes the amount of transfer as given when she makes the consumption choice for herself and, implicitly, for the child. If her consumption choice results in the consumption on the child exceeding the amount of transfer, then I do not allow the parent to choose that transfer amount.

Guilt: Parents who renege on any part of the contract suffers from guilt, which reflects failures to meet two separate obligations: κ_1 represents guilt from failing the financial obligations, while κ_2 covers the elder care aspect. This guilt is cumulative; each new breach adds to existing guilt, and the effects of past breaches persist over time.

$$g_t = \begin{cases} max\left\{\underbrace{g_{t-1}}_{\text{(last period's guilt)}}, \kappa_1 \times \underbrace{\mathbb{1}_{Tr_t < c_t^G + \bar{c}_t^h}}_{\text{(low transfer)}} + \kappa_2 \times \underbrace{\mathbb{1}_{urban_t^P = 1}}_{\text{(no elder care)}}\right\} & \text{if } \underbrace{\text{contract}_t^P = 1 \text{ and } h_t = 1}_{\text{(contract holds, grandparent is ill)}} \\ g_{t-1} \quad \text{(guilt persists)} & \text{if contract}_t^P = 1 \text{ and } h_t \neq 1 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

4.4 Government

This model captures three key policy channels impacting rural Chinese households, reflecting the policy environment from 2008 to 2014.

Health insurance: Government pays a fixed fraction ρ^h of the total medical cost $\tilde{c}^h$ when the grandparent is ill,[22] so the homogeneous private cost of healthcare $\bar{c}_t^h$ is:

$$\bar{c}_t^h = \left(1 - \rho^h\right) \times \tilde{c}^h \tag{13}$$

Education subsidy: Compulsory education (up to the 9th grade) is less expensive at a child's Hukou location. Migrants face higher tuition fees when they bring their children to urban areas and choose to continue their education.[23] Thus, the child's education cost is

$$\overline{tuition}_t = tuition_t^R \times \mathbb{1}_{urban_t^C = 0} + tuition_t^U \times \mathbb{1}_{urban_t^C = 1} \tag{14}$$

in which $tuition_t^R$ and $tuition_t^U$ are rural and urban tuition by education level indexed by t.

Rural social security system: Let c_{min} denote the minimum consumption level speci-

[22]Most Chinese citizens with rural Hukou were covered by the New Rural Cooperative Medical Care System (NRCMCS) by 2010 ("Enrollment Rate for Three Basic Medical Insurance was more than 95% in 2014." CNR.cn). The government gradually adjusted the coverage to increase its copay fraction.

[23]In addition, urban schools admitting rural children are under-regulated (Hernández, Javier C. and Iris Zhao. "One Target in Beijing's Migrant Crackdown: Schoolchildren". New York Times (2017)) and more expensive than rural public schools (Average out-of-pocket tuition fees for rural children in urban schools are listed in Table A3.)

fied in the security system.[24] Should a rural resident's total resources fall below this threshold ($s_t + A_j < c_{min}$), it is assumed they exhaust all savings ($s_{t+1} = 0$) on essential needs like housing and food, but not on education or healthcare. The government then supplements their consumption to meet the poverty line ($c_t = c_{min}$).

4.5 State and choice variables and the maximization problem

Initial condition and terminal condition: The model begins with the child's birth, parents starting with no savings ($s_0^P = 0$), and grandparents having wealth proportional to their agricultural income ($s_0^G = \omega A_j$). Grandparent's initial health status is determined by probabilities $pr(h_0 = 0)$ for healthy, $pr(h_0 = 1)$ for ill, and $pr(h_0 = 2)$ for deceased. It ends when the child turns 22, achieving financial independence and completing their education (edu_T). Parents remain in rural areas thereafter, with terminal utility calculated as the present discounted value of their flow utility for the next 20 years:

$$U_{j,T}^P = \sum_{i=0}^{19} \beta^i \Big(u^{private}({c_j}_T^P, l_T^P) \times \frac{edu_T^{1-\lambda}}{1-\lambda} \times \ (1 - g_T) \Big) \tag{15}$$

in which ${c_j}_T^P$ satisifies $\sum_{i=0}^{19} \frac{{c_j}_T^P}{(1+r)^i} = s_T^P + \sum_{i=0}^{19} \frac{A_j}{(1+r)^i}$, meaning that he completely smoothes consumption throughout those years; $l_T^P = T_{total} - T_{rural}$ as he only works in the rural area.

Grandparents are assumed deceased in the terminal condition,[25] with their terminal utility as the isoelastic utility gain from the child's educational outcome upon their death:[26]

$$U_{t_{death}}^G = \delta \frac{edu_{t_{death}}^{1-\lambda}}{1-\lambda} \tag{16}$$

[24] The Chinese government has a Rural Minimum Living Security (*Dibao*) system ("China's new approach to beating poverty." The Economist (2017)). The system subsidizes rural people below the poverty line to ensure their annual income reaches 1,210 RMB ($174.5) ("Annual report of statistics of social services", Ministry of Civil Affairs of the People's Republic of China (2009)). In the estimation, I set $c_{min} = 1,210$RMB. Note that the social security eligibility is tied to the *Hukou* system. As a result, migrants are ineligible for the social security system in either rural or urban areas.

[25] I provide the distribution of the health status of the grandparents by children's age in Appendix E.1.5 (my analysis, CFPS 2010-2014). In addition, the life expectancy is 58.99 for rural men and 72.46 for rural women (Shen, 1993).

[26] In the terminal condition, the flow utility of the grandparent after his death is set to zero. The factor δ is difficult to identify in my model given the data available. So I set $\delta = 1$ in my estimation.

Parent's optimization problem: The parent's choice variables can be summarized as $choice_t^P = \left\{urban_t^P, urban_t^C, c_t^P, tuition_t, Tr_t\right\}$, including his migration decision for himself and his child, and the daily consumption for himself,[27] the child's tuition, and the remittance to the grandparent. The parent's maximization problem is

$$V_t^P\left(X_t\,|j\right) = \max_{choice_t^P}\left(u^P\left(c_t^P, l_t^P, enroll_t, g_t\right) + \beta E_t\left[V_{t+1}^P\left|X_t, j, choice_t^P\right.\right]\right) \tag{17}$$

Grandparent's optimization problem: The grandparent's choice variables are the one-time contract decision, private consumption for herself and consumption on healthcare, summarized as $choice_0^G = \left\{contract^G\right\}$ and $choice_t^G = \left\{c_t^G, c_t^h\right\}$ for $t > 0$. Equation (18) below states that upon the child's birth, the grandparent decides whether to propose the contract by comparing the expected value functions of two potential futures, conditioning on her household type j. Equation (19) states how, during the child's upbringing, the grandparent chooses c_t^G and c_t^h in each period to maximize her value function, conditional on j, $state_t$ including her previous contract choice, and $choice_t^P$.

$$t = 0:\ V_0^G\left(j\right) = \max_{contract^G}\left\{E_0\left[V_1^G\left|contract^G = 0, j\right.\right], E_0\left[V_1^G\left|contract^G = 1, j\right.\right]\right\} \tag{18}$$

$$\begin{aligned} t > 0:\ V_t^G\left(X_t\,|\,j, choice_t^P\right) = \max_{choice_t^G}\Big(& u^G\left(c_t^G, l_t^G, enroll_t, HC_t\left|choice_t^P\right.\right) \\ &+ \beta E_t\left[V_{t+1}^G\left|X_t, j, choice_t^P, choice_t^G\right.\right]\Big) \end{aligned} \tag{19}$$

This model framework, including many heterogeneity and uncertainties, migration choices, three-generational family structure, and various sequential game dynamics, is computationally demanding. It excludes some rural family dynamics, such as uncertain urban labor markets and family structures with multiple migrants, to focus on the core interactions between migrating rural parents and their dependents needing education and healthcare.

[27] For both the parent and the grandparent below, once c_t^P or c_t^G are chosen, the co-residing child's consumption is determined by the equivalence scale. Therefore, c_t^C is not a separate choice variable for the guardian's consumption choice.

5 Estimation and results

5.1 Structural Estimation

In this section, I describe the two-step structural estimation procedure in detail. In the first step, I estimate the parameters that can be cleanly identified outside the model, or set them using values from the literature. In the second step, I estimate 9 preference parameters and one heterogeneity parameter using the Generalized Method of Moments.

5.1.1 Externally estimated parameters

In the structural estimation, I focus on the preference parameters and the marginal effect of intergenerational behaviors on the welfare of all three generations. Many other parameters in my model can be cleanly identified using various data sources introduced in Section 2. This subsection describes these parameters and how they are identified from the data. In Appendix E, I provide additional information on the estimation of the parameters.

Distribution of agricultural income: The distribution of per capita agricultural income skews to the left with 90% of the households' income level falling between 500 and 10,000 RMB/year (roughly \$73 and \$1,459). I discretize the distribution into 10 levels with comparable population shares as listed in Table B11

Price ratio and subsistence consumption level: A price ratio between the segregated rural and urban goods markets is needed to transform nominal daily consumption levels into purchasing powers or standards of living, but an estimate of the price ratio for consumption net of education and healthcare is not available. Therefore, I use CHFS data to estimate the price ratio from the Engel curves for food expenditure among the total daily consumption (Hamilton, 2001; Almås et al., 2018)[28] and obtain a ratio of 6.48. Another way to do this: use expenditure variation and relative food price between urban and rural to estimate relative non-food price.

[28] Appendix E details the estimation process.

Time allocation: The average hours worked per week is 66 for migrants and 50 for non-migrants, estimated from the rural-to-urban migrant and rural resident subsamples of the RUMIC 2009 data. The average number of hours needed to care for a child under 15 years old and an ill grandparent are both 12 hours per week, respectively, estimated from the CHNS data. I compute the number of hours of leisure time by assuming an endowment of 12 hours per day and subtracting the hours taken by each individual labor and home production activity from the endowment.

Cost of healthcare: Patients pay the total cost of healthcare initially, then the government's co-pay system can reimburse 34% of the expenses (Deng et al., 2017). The total costs of healthcare is highly bimodal at zero and 3,305 RMB ($483) per year. Rural households spend an additional 8,209 RMB ($1,198) on healthcare in the year in which an elder family member dies.

Baseline probabilities on health status: Amid the empirical difficulties,[29] I estimate the grandparent's probability distribution over health status by children's age range in two steps using the CFPS data. First, I take the subsample of grandparents whose adult children never migrated, and restrict the sample to grandparents who spent on healthcare when they are ill to estimate the distribution of the grandparents' health status by the children's age group.[30] Second, I recover the health transition probabilities from the static distribution.

In addition, I estimate the effect of the parent's migration on the grandparent's mortality by running a Logit regression of the grandparent's survival based on whether the parents took care of them in the past using the CFPS data. The analysis suggests that lack of adult children's private care increases the grandparent's mortality rate by 67%.

Dropout probability by education level and caretaker: Let $pr^{dropout}_{t_c,i}$ denote the

[29]The number of observations with a transition in health status is very limited in the five datasets I use, which are either cross-sectional or short panel datasets with 2 or 3 interviews. Moreover, the intertemporal probabilities of changes in health status of the elderly need to be organized by children's age groups, in order to be compatible with my model. The CFPS data is the only dataset that has informative variables on all three generations.

[30]In order to focus on the transition of health, I eliminate the effects of poor financial condition or lack of private care on mortality.

dropout probability for a specific education level t_c and caretaker $i \in \{\text{parent}, \text{grandparent}\}$, and let edu_g denote the education level of the caretaker.[31] The dropout probability based on the child's education level and the caretaker's can be computed as follows:[32]

$$pr^{dropout}_{t_c,i} := \sum_{t=0}^{16} pr(dropout\,|t_c, edu_g = t\,) \times pr(edu_g = t\,|i\,) \tag{20}$$

Risk in college admission: In period 5, if the parent decide to invest in the child's college education, the child could participate in the National College Entrance Exam. The estimated probability of being admitted by any college conditioning on participation is 27%.[33]

5.1.2 General Method of Moments (GMM) estimation

I use the generalized method of moments to estimate the model (Hansen, 1982). The method searches for the set of parameters that best match the theoretical moments on people's behavior predicted by the model with the empirical patterns we observed in the data. The set of parameters is denoted by a vector $\vec{\theta}$. I denote the set of moments used to describe people's decisions as Q_0 for the data moments and $Q(\vec{\theta})$ for the model moments. Therefore, the set of parameter $\hat{v}_{GMM}$ is defined in equation (21):

$$\hat{\theta}_{GMM} = \arg\min_{\vec{\theta}}\ (Q_0 - Q(\vec{\theta}))' W^{-1} (Q_0 - Q(\vec{\theta})) \tag{21}$$

[31] The sample size is too small to estimate the dropout probability for each combination of the child's level of education, the grandparent's role as caretaker, and the caretaker's education. My findings in Section 3.2 show that it is sufficient to characterize the dropout probability using the effect of the education level of the caretaker on the child's dropout probability, and the caretaker-specific education attainment distributions. They can be measured from the CFPS data with convincing sample sizes.

[32] Sociology research argues that migrants' children are more eager to quit school earlier to become migrant workers as they see their parents as examples (China Women's Federation Children's Work Department, 2011). To separate involuntary dropouts caused by poor economic conditions from dropouts caused by poor performance or lack of incentive to continue education, when estimating the right-hand side probabilities in equation (20), I exclude children who claimed that they dropped out because of economic difficulties.

[33] I compute the exam passing rate using national level data on the number of rural children who participated in the exam and the number of rural children who were newly enrolled in college in 2004. In the model, if the child passes the exam, then the parents pay the tuition, and the child receives a college degree by the end of period 5. If she does not pass, the tuition money that her parent prepared for her college education remains in his stock of savings.

5.1.3 Parameters, moments and identification

Most parameters that I structurally estimate using the GMM method have straightforward sources of identification. For example, the discount factor, the CRRA coefficient, and the consumption-leisure trade-off parameter in the utility function affects consumption behavior, and motivates migration and participation in the informal contract. Ill grandparent's utility gain from healthcare directly influences the rural elderly's healthcare consumption. I list the 11 structurally estimated parameters and their key sources of identification in Table B17 and provide a complete list of 29 moments in Table 3. Because the identification of parameters related to the informal contract could be less intuitive, I elaborate on them in detail.

Children's education parameters (φ and λ in equations (3), (15), and (16)) directly influence parental decisions on tuition and migration, with those valuing education likely to migrate for better incomes to finance schooling. Children's better school performance when living with parents motivates decisions against leaving the children behind and affects informal contract commitments. Moreover, grandparents gain higher utility from extended life during their grandchildren's schooling, with φ boosting utility from high school attendance and λ affecting all educational stages. Notably, φ specifically enhances utility from high school enrollment, while λ impacts returns at all educational levels. These dynamics are quantified using data on education enrollment rates and educational achievements.

Guilt associated with reneging on financial and care obligations in the informal contract (κ_1 and κ_2 in equation (12)) influences parental behavior towards ill grandparents. High κ_1 values lead parents to increase financial transfers during grandparent's illness. High κ_2 values encourage parents to remain in rural areas to provide elder care. So the proportions of parents who fulfill each contractual component identify the two guild parameters.

Weighting matrix in GMM estimation: For the GMM estimation, I use an identity weighting matrix for the 26 moments on various fractions. I assign smaller weights to three specific moments: the average rural consumption, the average financial transfers to grandparents, and the average educational attainment of rural children. The small weights for the three quantities are determined from the standard errors from data estimates.[34]

5.1.4 Computational strategy

Conceptually, the analytical solution to the model exists for any given set of parameter values, because it does not involve any continuously distributed idiosyncratic shocks. Finding the theoretical moments predicted by the model means computing the probability distribution over all possible profiles of choice and state variables, i.e. over all distinct values for the vector $\{j, contract^G, \{X_t, choice_t^P, choice_t^G\}_{t=1}^5\}$. Calculating the probability distribution and searching for the optimal set of parameters using the GMM technique demands extremely complex computations. I explain in this subsection the major challenges I faced in the estimation process and introduce my solutions to them.

The challenges are two-fold. First, the three-generational household structure with two decision-makers and a sequential game structure embeds a large set of choices. Specifically, within each period, there are up to 3,240 distinct pairs of choices for the two agents.[35] Furthermore, the pair of per-period value functions for the two agents are evaluated 2.2×10^{15} times to obtain the set of moments for a given set of parameters.[36] I use various techniques to improve efficiency. The resulting algorithm takes 17.164 hours for a one-core computer to complete one iteration.[37] Second, the objective function specified in equation (21) is not

[34] I do not use the optimal weighting matrix. Most moments I match in the data do not have substantial sampling errors, considering they are nearly perfectly estimated using nationally representative survey data. A diagonal weighting matrix based on the standard errors from data estimates of the moments would result in imprecisely estimated moments, having essentially no effects on identification in the structural estimation procedure. I want the model to match all the moments with equal weights, so I assign higher weights to moments estimated on smaller samples than their weights in the optimal weighting matrix. I report the specific values of the weights in Appendix E.3.

[35] The number of evaluations is for the baseline setting of the model. In this setting, the daily consumption of the parent and the grandparent (and thus the savings for the two agents) are the only continuous choice variables in the model, with 9 and 15 points on the grid, respectively.

[36] Note that, because no two periods are exactly the same, the computation for flow utilities in a period cannot be used in another period to save computation time. For this reason, the per-period value functions are calculated as many times as the total number of possible profiles of choice and state variables.

[37] (a) Techniques I applied include the Branch optimization, Common sub-expression elimination, Constant propagation, Code Inlining, Instruction scheduling, Inter-procedural analysis, and Heap Memory Management. Without these techniques, the time cost would have been 550.27 hours.

(b) The one-core computer is one out of 24 cores of a 3.00 GHz Intel Xeon Platinum 8158 Processor.

globally concave. It disqualifies gradient estimation methods, as the derivative-based search would likely fall into local optima that may not be the global optimum. A grid search would be the most desirable approach but is infeasible considering the computation costs.

To overcome the challenges caused by the complicated model structure and the non-monotonic objective function, my computation strategy involves two key techniques: My first technique is the Genetic Algorithm (Mitchell, 1998) followed by the Nelder-Mead (Nelder and Mead, 1965) method.[38] My second technique is, within each group of estimation in the Genetic Algorithm, to divide each iteration into 40 non-overlapping optimal sub-tasks,[39] and apply the concept of a distributed system design. The system uses a one-core machine to control as many mutually independent high-performance-computers (HPCs), each equipped with multiple CPUs, as desired to concurrently estimate a generation of 50 to 70 iterations (2,000 $\sim$ 3,000 sub-tasks in total) when running the Genetic Algorithm. The hierarchical system makes nearly full use of all CPUs on all HPCs for most of the time by centralized resource allocation and the efficient assignment of sub-tasks to computers. Figure B6 in Appendix F provides an example of the system's performance, illustrated by the CPU usage of ten 72-core HPCs controlled by a one-core machine on Amazon Web Services.

The model simulation produces a probability distribution of a set of representative agents over all possible multi-period game outcomes. The model moments are probabilities and weighted averages of game outcomes that can uniquely identify the choice of parameter values. Their corresponding data moments are macro-level fractions and average values of economic terms. These statistics can be estimated with more confidence, despite the

[38]The genetic algorithm is used in only a few economic studies (Holland and Miller, 1991), and is used much more extensively in natural science and engineering literature. My implementation of the Genetic Algorithm applies a combination of the Roulette Wheel Selection (Goldberg and Deb, 1991) and the Elitism Selection rules (Baker, 1985). It also applies the optimal mutation and crossover probabilities as well as the optimal population size from the literature on the Mathematical Theory of Computation (Alander, 1992; Stanhope and Daida, 1998).

[39]An iteration can be divided into sub-tasks by household type, the grandparent's initial health status, and her choice on the informal contract. There are 10 household types, 3 health statuses, and 2 contract choices. If the grandparent is healthy when the model starts, she has two options on the informal contract: accept or reject. An ill or dead grandparent, meanwhile, cannot enter the contract. Therefore, each iteration has 40 non-overlapping sub-tasks.

challenge of identifying and tracking migrants and their families in survey data.

To demonstrate equivalence between the dynamic programming approach and the computation approach I developed, I numerically solve a lifecycle model with stochastic income and a borrowing constraint using both approaches, as detailed in Appendix F.2. Additionally, I report the time costs of the two approaches as the model complexity increases, illustrating the higher efficiency of my approach and thereby its advantage in solving complex models.

5.2 Estimation Results

5.2.1 Parameter estimates

Table 2 gives the estimates of the parameters. My estimates for the standard errors follow Lockwood (2018)'s approach. Prior works on the elasticity of intertemporal substitution of rural people in China found CRRA coefficients between 0.5 and 3.49 (Zhang, 2011; Liao, 2013; Zhu et al., 2014; Yang and Qiu, 2016; Wang et al., 2017). My estimate of $\gamma = 1.235$ is in line with existing literature. The discount factor, $\beta = 0.866$, is also within the wide range of existing estimates on rural Chinese people's discount factor (Liao, 2013; Zhu et al., 2014; Yang and Qiu, 2016), with values between 0.395 and 0.955. The value of $\eta_2 = 0.349$ indicates that grandparent gains from healthcare, which provides an incentive for the informal contract. The estimates of parameters governing the parent's and grandparent's altruism toward the child's education ($\lambda = 2.649$) reveal a high marginal return to additional years of schooling. High school enrollment ($\varphi = 0.025$) yields a utility gain that is equivalent to a 29% increase in consumption. The guilt parameters governing the parent's attitude towards the informal contract (κ_1 and κ_2) are small, but are sufficient for many parents to fulfill the contract. The guilt from reneging on the financial component of the informal contract ($\kappa_1 = 0.006$) is equivalent to a 6% decrease in consumption. The guilt from reneging on the private care component of the contract ($\kappa_2 = 0.050$) is equivalent to a 40% drop in consumption.

Table 2: Parameter estimates and standard error

Description	Symbol	Estimates	Std. Error
Annual discount factor	β	0.8657	0.0283
Coef of CRRA for private utility	γ	1.2348	0.0689
Consumption-leisure tradeoff	θ	0.2789	0.0154
Coef. on additive component of utility function	δ	0.3554	0.6202
Grandparent's utility gain from healthcare when ill	η_2	0.3491	0.0020
Utility gain if child is enrolled in high school	φ	0.0245	0.0079
Coef of CRRA for children's education	λ	2.6493	0.0078
Parent's guilt from low remittance	κ_1	0.0060	0.0004
Parent's guilt from not caring for grandparent	κ_2	0.0501	0.0004
Migrating parent's unemployment rate when the child is over age 18	pr^{ump}	0.4030	0.0416

5.2.2 Model fit

Table 3 shows the goodness of fit by contrasting the data moments with model estimates of the moments. In general, the model fits the data well.

The estimated model captures the decreasing time trend in the fraction of migrants as children grow up. The model slightly underestimates the fraction of migrants.[40] It also captures the decreasing school enrollment rates. When parents cannot fulfill both parts of the informal contract, their preference between financial support and elder care is also captured by the model estimates.

The model fails to capture the parents' decision to leave children behind when the children are between 6 and 11 years old. The model estimates on the fraction of migrants and the fraction of left behind children imply that all migrating parents bring their children when the children are in the primary school period. These parents are likely to be from wealthy

[40]In my model, the parent's generation only has one person, while in reality, a three-generational household is supported by two adults who can work in the urban labor market. To correct for this simplification, the urban wage rate I set in the model is the sum of the average wage of a male migrant and the average wage of a female migrant. However, not all couples in migrants' households move together, so the actual return to migration is lower than what I set in the model. Since I overestimate the financial return to migration, it is expected that I underestimate the fraction of migrants.

Table 3: Goodness of fit of model to the data

Moment			Data	Model Estimates
Migration (Data source: CFPS)				
Fraction of migrants by children's age				
Age 0 to 5			0.322	0.261
Age 6 to 11			0.243	0.211
Age 12 to 14			0.183	0.136
Age 15 to 17			0.171	0.206
Age 18 to 20			0.149	0.035
Informal contract (Data source: CHARLS)				
Conditioning on grandparents having provided childcare in the past, the fraction of parents who:				
Child's age	Parent provides			
	financial support	elder care		
Age 0 to 14	Yes	Yes	0.665	0.859
Age 0 to 14	Yes	No	0.059	0.000
Age 0 to 14	No	Yes	0.222	0.141
Age 15 to 20	Yes	Yes	0.647	0.757
Age 15 to 20	Yes	No	0.294	0.200
Age 15 to 20	No	Yes	0.038	0.043
Fraction of sick grandparents left behind by parents and children			0.083	0.021
Remittance and consumption (Data source: CFPS)				
Average annual remittance (RMB/year)			4,099	3,915
Fraction of grandparents receiving remittance			0.398	0.222
Average consumption per rural adult (RMB/year)			3,394	3,226
Children (Data source: CFPS)				
Fraction of left-behind children				
Age 0 to 5			0.257	0.261
Age 6 to 11			0.187	0.000
Age 12 to 14			0.099	0.136
Fraction of children enrolled in school				
Primary school			0.969	1.000
Middle school			0.898	0.871
High school			0.567	0.632
Average years of schooling of children at age 22 (years)			10.130	10.273
Grandparents' health (Data source: CHNS)				
Fraction of sick grandparent not receiving health care			0.548	0.652
Fraction of healthy grandparents by children's age				
Age 6 to 11			0.543	0.546
Age 12 to 14			0.430	0.394
Age 15 to 17			0.252	0.335
Fraction of sick grandparents by children's age				
Age 6 to 11			0.169	0.214
Age 12 to 14			0.115	0.172
Age 15 to 17			0.090	0.101

households in which the grandparents do not propose the contract, or from households in which the grandparents are unhealthy and thus are not eligible to propose the contract.

5.2.3 Sensitivity of parameter to moments

I provide an estimate for the sensitivity of the parameters to the 29 moments I fit the model. This statistic validates this model and its parameter estimates. The approach is from Andrews et al. (2017).

I first measure the Jacobian matrix G of the model moments $Q(\vec{\theta})$ with respect to the vector of parameters $\vec{\theta}$ at the parameter estimates presented in Table 2. Then I use the estimated Jacobian $\hat{G}$ and the weighting matrix W to calculate the sensitivity measure defined in equation (22).

$$\Lambda = (\hat{G}'W\hat{G})^{-1}\hat{G}'W \tag{22}$$

The estimate for the sensitivity matrix is reported in Table B16 of Appendix E. The results are consistent with the discussion in Section 5.1.3 and Appendix section E.4. Moreover, some elements of the sensitivity matrix imply that the household structure entangles all three generations. For example, the parent's and grandparent's utility gain from the child's high school enrollment (φ) is sensitive to the fraction of parents who fulfill the financial component of the informal contract when the child is in the compulsory education periods. The CRRA coefficient for children's education (λ) is also sensitive to the fraction of grandparents receiving transfers from the parents. The parent's utility cost of not fulfilling the financial component of the informal contract (κ_1) is sensitive to the fraction of left-behind children. Therefore, rural parents under binding credit constraints trade-off between fulfilling their duty to the grandparents and supporting their child's education. The informal contract and household finances link all family members together.

6 The value of intrahousehold cooperation

This section demonstrates the quantitative importance of the informal contract in welfare distribution within the household. Parents working in urban areas potentially raise welfare for their young children and elderly grandparents through intrahousehold cooperation. In this model, an informal contract, involving exchange of assets, time, and care between parents and grandparents, is only adopted if it benefits both parties, conditioning on their contractual space and the household's policy environment. Table A5 details the contract's benefits and costs for each generation, alongside the model's predictions for its overall effects.

Figure 4 illustrates the ex ante utility gains from the contract for each agent, categorized by household income levels and expressed in consumption equivalent scale. A utility gain is indicated by line segments above one. The figure indicates that higher productivity diminishes parents' welfare gains from the contract.

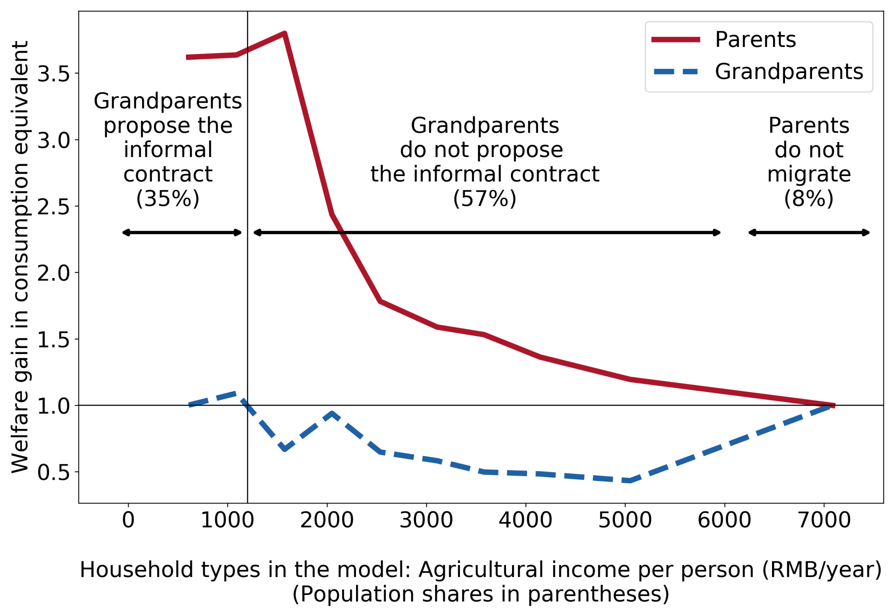


Figure 4: Parents' and grandparents' welfare gain from the informal contract by rural income

Note: The above figure presents the agents' welfare gain from the informal contract in consumption equivalent by household type. The red solid line and the blue dash line are $\frac{\text{welfare when household opts into the contract}}{\text{welfare when household does not opt into the contract}}$ for grandparent and parent, respectively. The vertical line is the household's agricultural endowment level that differentiates the grandparents' informal contract status.

While informal contracts allow parents to migrate alone to accumulate assets faster, these benefits last only as long as grandparents stay healthy. However, when grandparents fall ill, the contract either costs the parents in money or utility. As a result, the associated disutility from reneging imposes substantial costs on parents, enforces its continuation despite its limited commitment nature. My model shows that 92% of parents in three-generational rural households benefit from informal contracts under current policies,[41] with the contract's value increasing alongside rural-urban wage disparities. However, its value diminishes with higher remittances, private healthcare costs, and caregiving time for ill grandparents.

Grandparents' welfare gain from these contracts varies with their health and potential breach by the parent. The perceived benefits of the contracts increase as agricultural income drops, especially in 35% of rural households where annual income is below 1,200 RMB (CFPS 2010-2014), leading to potential welfare losses outweighing benefits for grandparents derived from additional income and elder care provided by their children. As a result, the value of informal contracts lessens with a high breaching rate by the parents. Grandparents' welfare gains from contracts depends on: (1) the extent to which the informal contract replaces benefits that would otherwise be provided by the social security system. With minimal crowd-out effects, grandparents gain more from cooperation within the household. Consequently, the poorest households (rural income = \$600) experience lesser welfare benefits from these contracts than those slightly wealthier (rural income = \$1,000). (2) current consumption levels, where additional income generates diminishing marginal utilities. And (3) the size of financial transfers.[42] Moreover, the desirability of contracts grows with increasing private healthcare costs and when family care significantly lowers grandparent's mortality.

The effect of the informal contract and parent's migration on rural children's education remains ambiguous, as summarized in Section 3.3 and evidenced by my model. Living with less educated grandparents often hampers children's education across all income levels.

[41] According to Figure 4, only the wealthiest household type among the 10 types in my model, representing 8% of all rural households (CFPS 2010-2014), consistently declines the informal contract, even if proposed by grandparents.

[42] Section 4.3 explains that transfer amounts increase with the grandparent's consumption level. When the grandparent is healthy, transfers are allocated to cover the child's consumption, calculated as a fixed proportion of the grandparent's consumption based on the OECD equivalence scale. When the grandparent is ill, the transfer covers the grandparent's consumption.

On the other hand, migrant parents, who are generally wealthier, can provide better for their children, leading to lower dropout rates due to relaxed financial constraints (see Figure A1a). Parental wealth positively correlates with agricultural income and thus diminishes the benefits of informal contracts. My model indicates that while intrahousehold cooperation reduces children's educational attainment by 0.07 year, the overall negative impact is small.

Figure 4 illustrates that parents benefit from informal contracts across all income levels, but wealthier grandparents do not. To further explore intrahousehold cooperation, I examine two hypothetical scenarios: one where all grandparents offer contracts, and one where none do. Apart from the wealthiest, all other income groups see consumption gains from these contracts (Figure A1). Despite a general correlation between agricultural income and consumption, only poorer households typically engage in such contracts, improving their welfare and living conditions in rural China.

7 Counterfactual experiments

The Chinese government pursues two main objectives with policies targeting rural households and rura-to-urban migrants. The first is to improve the welfare of rural populations, aiming to address issues such as the welfare of left-behind children and healthcare for the elderly. The second objective is to manage the composition of migrants in urban areas effectively. Migrating parents provide a vital labor supply, while migrating children consume urban educational resources, and all migrants utilize urban amenities.

This section evaluates an expanded health insurance subsidy and various policies focused on migrating children, then discuss the effects of these policies on the welfare and behavioral changes across generations within rural households and examine their impact on the migration patterns by generation. Table 4 reports the results of all counterfactual experiments. Because the effects of changing policies to household behaviors are monotonic, I present the results of two specific policy parameter values for each policy direction.

7.1 Expanded public healthcare coverage

The government may increase the public health insurance coverage to improve the welfare of rural grandparents. As of 2010, most rural residents are enrolled in the New Rural Cooperative Medical Care System (NRCMCS), which offers a 34% reimbursement rate. My first counterfactual experiment adjusts the parameter ρ^h in equation (13) from 10% to 80%, assessing the response of rural households. Table 4 illustrates the policy effects on children's average educational attainment, the proportion of migrating parents, healthcare spending by grandparents, and average consumption levels among migrants.

Effects on the targeted generation: The fraction of grandparents who spend on healthcare increases by up to 8%. They spend less when healthy in order to save for consumption on commodities and healthcare when falling ill.

Effects on other generations: In the new steady states, expanded policy reshapes intrahousehold dynamics through informal contracts, impacting all three generations.

More affordable healthcare lowers the grandparents' utility gain from the financial transfer from the parents, and thus the value of the informal contract to the grandparents falls. It reduces the grandparents' reliance on financial transfers from parents, making it more feasible for them to exit informal contracts, thus reducing overall contract adoption (see Figure 5). In addition, the need for parents to fulfill financial obligations under these contracts decreases. Grandparents benefit significantly from the policy as they become financially more self-reliant and less dependent on the parents' support.

On the contrary, the policy adversely affects parents. In households that opts out of the informal contracts, parents lose the option to leave children behind when migrating. Comparing the policy scenarios with public health coverage at 10% versus 80%, I see parents' migration reduces by 5% and consumption decreases by 15%.

Table 4: Counterfactual Experiment Results

	Baseline	Government coverage for healthcare ρ^h		Percentage reduction in childcare time ρ^{care}		Percentage of urban education subsidy ρ^{edu}	
		0.1	0.8	0.4	0.8	0.4	0.8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Migration behavior							
Fraction of parent migration	0.182	0.197	0.142	0.273	0.388	0.181	0.180
Fraction of children migration	0.063	0.035	0.142	0.090	0.182	0.063	0.063
Fraction of left-behind children	0.119	0.162	0	0.183	0.206	0.117	0.117
Grandparents' behavior							
Consumption when healthy	3,534	3,783	3,178	3,778	3,790	3,534	3,534
Consumption when ill	1,951	1,914	1,792	2,062	2,128	1,954	1,951
Fraction of grandparents spending on healthcare	0.637	0.592	0.675	0.650	0.651	0.637	0.637
Parents' behavior							
Consumption in rural area	3,351	3,496	3,005	3,881	4,404	3,346	3,381
Consumption in urban area (normalized to rural purchasing power)	2,257	2,328	1,965	2,251	2,255	2,292	2,347
Children's education							
Children's average education attainment	10.431	10.330	10.533	11.386	11.811	10.504	10.527

Note: This table reports the results from counterfactual experiments. The results presented cover the migration behavior of the parents and the children and the implied fraction of rural children left-behind, the grandparents' consumption on commodities and healthcare, the parent's consumption, and the children's average education attainment. Column 1 reports the simulated moments using the baseline model. In the baseline specification, the public health system reimburses 34% of healthcare costs ($\rho^h = 0.34$), while there are no reductions or subsidies for the time cost or education in urban areas when raising children, with $\rho^{care} = 0$ and $\rho^{edu} = 0$. Columns 2-3 report the results by changing ρ^h to 0.1 and 0.8. Columns 4-7 report the results of setting ρ^{care} and ρ^{edu} to 0.4 and 0.8, respectively.

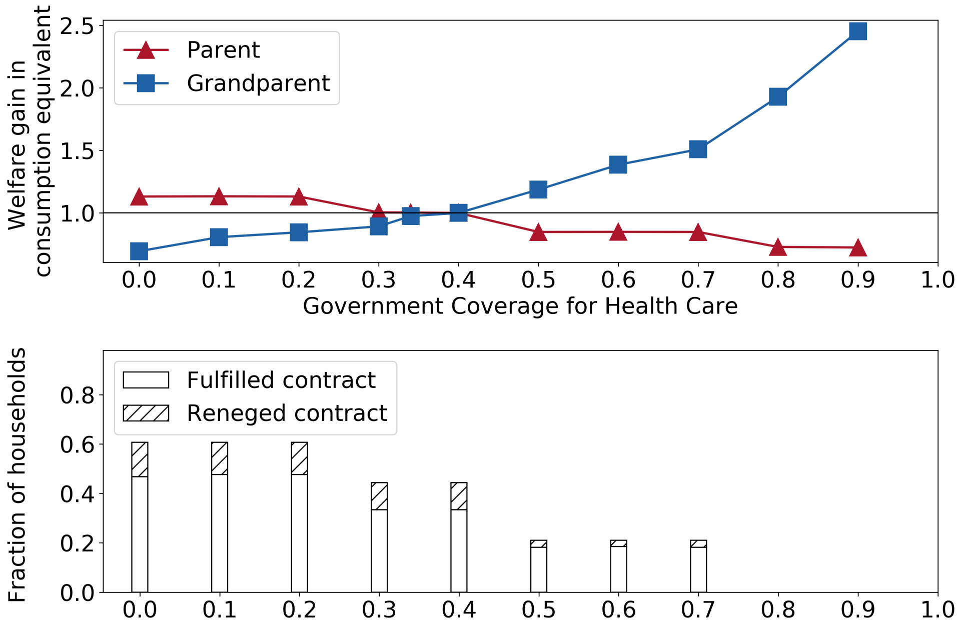


Figure 5: Public health policy effects on welfare and the household's contractual space

Note: The insurance coverage of the baseline model is 0.34. The consumption equivalent is calculated relative to the baseline model.

Public health policy effects on children are mixed. Tighter budget constraints from reduced parental migration lower investments in children's education. However, fewer children are left behind due to reduced reliance on informal contracts, potentially improving their educational outcomes. Overall, the policy appears to promote children's educational attainment by 0.2 year despite some adverse effects.

7.2 Policies for the children

As mentioned in Section 3.3, migrating parents often leave children behind due to high living and education costs, as well as insufficient childcare facilities. The model setup enables simulations of two policies targeted at rural children, assessing their effects on the three-generational household. The experiment results are reported in columns 4-7 of Table 4.

7.2.1 Childcare policy

Childcare policies aim to reduce the time costs associated with caring for migrating children.[43] The policy parameter ρ^{care} in equation (23) depicts the effect of expanded childcare facilities as the percentage reduction in childcare time cost, which in the baseline model is set to zero, indicating the absence of facilities for migrating children. Thus, parental leisure time in the model is defined as follows:

$$\begin{aligned} l_t^P \; = \; & T_{total} - T_{rural}\mathbb{1}_{urban_t^P=0} - T_{urban}\mathbb{1}_{urban_t^P=1} \\ & - \underbrace{(1-\rho^{care})}_{\text{(percentage reduction of childcare time cost)}} \times T_t^C \mathbb{1}_{urban_t^P=urban_t^C} - T^G \mathbb{1}_{urban_t^P=0}\mathbb{1}_{h_t=1} \end{aligned} \tag{23}$$

Effects on the children: The policy increases the children's average educational attainment by up to 1.4 years. It also promotes parents' migration from 18.2% to 38.8% and children's migration from 11.9% to 20.6%. The increased fraction of left-behind children indicates that the childcare policy makes parents easier to migrate the children when the grandparents fall ill, but would not alleviate the left-behind children when the grandparents are healthy.

Effects on other generations: The childcare policy yields considerable welfare benefits for parents (upper panel of Figure 6). The policy improves parents' flow utility when migrating with their children, without altering price levels, thus increasing the feasibility of family migration. This migration decision expands the parents' intertemporal budget, leading to higher urban and significantly improved rural consumption.

The policy mildly disadvantages grandparents through adjustments in informal contracts (lower panel of Figure 6). First, more grandparents are motivated to enter into these agreements. This stems from changes in parental behavior: once parents fulfill their part of the contract, they can migrate with their children. This agenda reduces the opportunity cost of

[43] In reality, despite some factories offering childcare facilities to attract migrants, many migrating children benefit little from these facilities due to financial, cultural, and transportation barriers. (Bland, Ben and Nicolle Liu. "China factories use childcare offer to lure migrant workers". Financial Times (2018); "Migrant workers and their children." China Labour Bulletin (2021).)

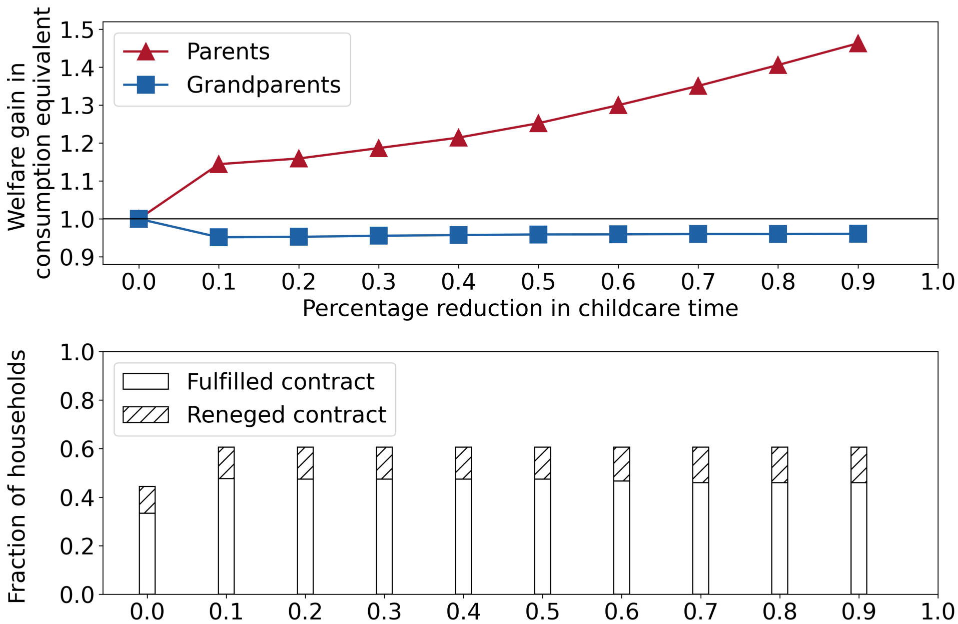


Figure 6: Childcare policy effects on welfare and the household's contractual space

Note: The upper panel of the above figure shows the effect of urban childcare policy on the welfare of the parents and grandparents. The welfare gains are presented in consumption equivalent term. The lower panel shows the policy effect on households' informal contract status. It presents the fraction of households with fulfilled contracts, and the fraction of households with reneged contracts. The statistics in both figures are calculated from the households with healthy grandparents when the children are born.

caring for ill grandparents and thus the likelihood of parents reneging on the contract. On the other hand, if grandparents fall ill before parents accumulate sufficient funds for education and contract obligations, parents might still opt to migrate with their children, leading to potential contract breaches. On average, while the increased participation in informal contracts provides the grandparents with more funds for commodities and healthcare, the ongoing risk of non-compliance on the private care part of the contract results in minor welfare losses for grandparents.

7.2.2 Urban education policy

Urban education policies intend to improve the welfare of migrating children by making education in cities more affordable and accessible. Recent initiatives have sought to integrate migrant children into urban schools by increasing government spending to subsidize urban tuition, and tightening regulations on low-quality private schools.[44] Despite these efforts, comprehensive policy evaluations remain scarce.

I introduce a parameter, ρ^{edu} in equation (14), to represent the percetage reduction of the urban education policy on the urban-rural tuition gap ($tuition_t^U - tuition_t^R$). I set $\rho^{edu} = 0$ in the baseline model,[45] and $\rho^{edu} = 1$ indicating equal tuition irrespective of the child's Hukou. Therefore, tuition costs are defined as follows:

$$\overline{tuition}_t = tuition_t^R + \underbrace{\left(1 - \rho^{edu}\right)}_{\text{(percentage reduction in urban-rural tuition gap)}} \times (tuition_t^U - tuition_t^R) \times urban_t^C. \tag{24}$$

Columns 6-7 of Table 4 show the effects of the policy on several outcomes:[46] children's average education attainment, the proportion of migrating parents, grandparents' healthcare spending, and migrants' average consumption levels.

Effects on the children: Children's average educational attainment is boosted by up to 0.1 year. The fraction of left-behind children is not affected, as migration rates for both parents and children increase simultaneously.

Effects on other generations: Changing urban education prices impacts parents' decision-making and subsequently influences grandparents' behavior through household dynamics. Initially, this price adjustment leads to a substitution effect within the migrating parents' budget constraint, incentivizing them to reallocate resources from expenditures on food and housing to children's education. Furthermore, while one might assume that lower urban tuition fees would incentivize parents to migrate with their children, particularly when

[44] Significant measures include the 13^{th} Five-Year Plan's promotion of public school admissions for migrant children and local government efforts to improve the quality of private urban schools (Central Compilation & Translation Press; Denton, Bryan. "One Target in Beijing's Migrant Crackdown: Schoolchildren," nytimes.com).

[45] The baseline model matches data moments corresponding to the policy environment between 2008 and 2014. The implementation of the current urban education policy happened after 2014, so I assume no subsidy in the baseline model.

[46] Note that the rural school tuition rates are not affected by this policy. I also examined a similar policy counterfactual that only lowers the tuition for compulsory education periods (i.e. exclude high school from the subsidy) and saw the same experiment outcomes.

grandparents are ill —potentially leading to more breaches of informal contracts —the model predicts otherwise. The lower panel of Figure A2 indicates that parents' likelihood of fulfilling or reneging on the informal contract remains unchanged, suggesting that educational costs are not the binding determinant in migration decisions, and that the guilt associated with reneging on the contract provide enough motivation to ensure grandparents' welfare. Overall, as shown in the upper panel of Figure A2, reducing the urban education costs provides only marginal welfare benefits to parents and negligible effects on grandparents' welfare.

7.3 Discussion on policy design

The government aims to improve the welfare of the rural population by alleviating the issue of left-behind children, promoting healthcare, and boosting daily consumption levels, while managing migrant composition for labor supply. The effects of these policies vary considerably significantly:

First, policy effects on welfare distribution within rural households vary in magnitudes and the way of re-allocation. Tuition reduction marginally improves welfare by at most 0.4%, whereas the childcare policy has much larger effects, increasing the welfare by up to 40%. The public health policy can boost grandparents' welfare by as much as 150%. These policies often lead to welfare trade-offs, particularly affecting budget and time constraints across generations. This suggests the necessity of using models accounting for intergenerational behaviors to make comprehensive predictions on policy effects.

Second, policies targeting rural elderly and children affect parental migration decisions and the demographic composition of migrants. Table A6 reports the flow of migration from the actual data and the predicted outcomes from counterfactual experiments. Childcare policies encourage migration with young children. Similarly, incentives for higher high school enrollment motivate parents to migrate for better educational opportunities for their children. Health policies promoting grandparent independence also prompt parents to migrate with their children, increasing migration among families with children in primary school. The expanded health policy detaches grandparents from the family unit. Consequently, parents are obliged to bring their children when they migrate, resulting in a notable migration uptick among parents of children aged 6 to 11.

8 Extension

An important discussion in a multi-generational household model is whether behaviors are motivated by altruism and/or exchange (Cox, 1987). I extend the model to analyze incentives behind intergenerational behaviors in Chinese households that face migration opportunities. In the model presented in section 4, the informal contract is an exchange between the grandparent and the parent. The grandparent provides service (care for the child) in exchange for the parent providing transfer and service (private care when the grandparent is ill)[47]. While the model incorporates the grandparent's and parent's altruism towards the child's educational outcomes, it does not account for altruism between these two agents.

Now I extend the model to include altruism between the grandparent and parent, re-estimate the parameters, and compare the model fit. The first extended model simultaneously accommodates both altruism and exchange motives. The flow utility functions are

$$U_t^{P,altruism} = u^P\left(\tilde{c}_t^P, l_t^P, enroll_t, g_t\right) + \alpha^P u^G\left(c_t^G, l_t^G, enroll_t, c_t^h\right) \tag{25}$$

and

$$U_t^{G,altruism} = u^G\left(c_t^G, l_t^G, enroll_t, c_t^h\right) + \alpha^G u^P\left(\tilde{c}_t^P, l_t^P, enroll_t, g_t\right). \tag{26}$$

in which U_t^P and U_t^G are defined in equations (3) and (5), with α^P and α^G quantifying the strength of altruism. All other model specifications remain unchanged. Identifying the altruistic parameters is challenging; therefore, I adopt $\alpha^P = 0.0027$ and $\alpha^G = 0.4781$ from Barczyk and Kredler (2018), which also emphasizes economic interactions between parent

[47] While the exchange motive model presented in Cox (1987) incorporate decisions of opting into the exchange, it does not specify the utility cost from an incomplete exchange. In this paper, I model the exit mechanism by introducing the disutility from reneging on the informal contract (equation (12)).

and child generations.[48]

The second extended model represents a pure altruism model that eliminates the exchange channel. In this model, the parent's behavior is independent of the grandparent's past actions, and there is no disutility associated with failing to fulfill the informal contract. Consequently, $g_t = 0$ in the flow utility function and $g_T = 0$ in the terminal condition, eliminating the need to estimate κ_1 and κ_2, the disutilities from reneging on the contract.

I re-estimate the parameters using GMM for the two extended models. Appendix Tables B18 and B19 report the parameter estimates and simulated moments of all three models. The pure exchange model most accurately matches 14 moments related to migration, intergenerational behaviors, grandparent's health status, as well as consumption.[49] In contrast, the pure altruism model and the dual-motive model do better job in matching only 4 moments.[50] The simulated moments on intergenerational behaviors differ across the models, demonstrating the appropriateness of selecting these moments to identify the underlying motives and parameter values associated with intrahousehold dynamics.

This analysis suggests that intrahousehold behaviors in rural Chinese households with migration opportunities are primarily driven by exchange motives. Altruism cannot adequately explain the observed patterns of outbound and return migration, child migration, private care for children and the elderly, and remittance decisions.[51] My finding aligns with the conclusions of Cox et al. (1998) and Kazianga (2006), who found that the motives for intrahousehold financial transfers are primarily exchange-based in developing countries such as Peru and Burkina Faso.

[48] Literature on Chinese households does not have reliable estimates for the parameters, so I use values from this paper which focus on U.S. households.

[49] These moments include: the parent's migration for children under age 12, the fractions of parents who provide either transfer or elder care when the grandparent is ill, the average annual remittance and consumption, the fractions of left-behind children, the fraction of ill grandparents by children's age, and the fraction of ill grandparent not consuming on healthcare.

[50] The pure altruism model best matches the parent's migration for children above age 15, the fraction of parents providing both transfer and elder care, and the fraction of grandparents receiving transfer. The dual-motive model best matches the parent's migration for children with age 12-14, the children's school enrollment rates, and their average education attainment.

[51] See Appendix E.5 for additional discussion on the parameter estimates and moments.

9 Conclusion

This paper studies the impact of internal migration in China on intergenerational behavior within rural households and explores how policies targeting migrants and their extended families influence the welfare of each generation. Using data from five micro datasets, findings show that migrating parents often leave children and grandparents behind, who then depend on remittances. The health status of grandparents significantly affects parental decisions to migrate or stay, impacting the care provided and the educational outcomes of children. I develop and estimate a three-generational rural households, in which migrating parents and left-behind grandparents form an informal contract covering child care, financial transfers, and elder care. Policies intended to improve the welfare of one generation inevitably influence the welfare of the remaining generations and affect the migration decisions of both parents and children. Counterfactual analyses suggest that reducing childcare costs can significantly boost educational outcomes and migration, while subsidies for living costs may increase overall consumption but intensify the left-behind issue. Further analysis confirms that the baseline model, highlighting exchange motives rather than altruism, is the most effective in capturing the intergenerational dynamics.

My research suggests that the government could influence the flow of rural-to-urban migration through policies targeting the left-behind family members. Over the past few years, the government primarily restricts urban settlement by demolishing urban villages and limiting migrants' access to public services, approaches that have sparked significant protests advocating for migrants' rights.[52] My counterfactual experiments indicate that altering the policy environment affecting migrants' families could shift the economic dynamics of migration. Such changes affect the associated costs and benefits, potentially prompting rural parents to voluntarily adjust their migration decisions. Furthermore, the inherent non-neutrality of exchange-motivated intrahousehold behaviors, as described in Cox (1987),

[52] Buckley, Chris. "Why Parts of Beijing Look Like a Devastated War Zone". The New York Times (2017); Kan, Karoline. "China's Migration Control Threatens Lives and Growth". Huffington Post (2017).

suggests that policies aimed at improving the welfare of left-behind generations will not be uniformly absorbed across household members. This ensures that the intended welfare improvements are not completely offset within the household.

In future work, the dynamic games in three-generational households can be studied in more depth. A more detailed model with heterogeneous informal contracts will allow us to better predict the effects of various public services on people with different age and household status, generating more comprehensive insights.

References

Alander, Jarmo T. 1992. "On optimal population size of genetic algorithms." In *1992 Proceedings Computer Systems and Software Engineering.* 65–70, IEEE.

Almås, Ingvild, Timothy KM Beatty, and Thomas F Crossley. 2018. "Lost in translation: What do Engel curves tell us about the cost of living?".

Altonji, Joseph G, Fumio Hayashi, and Laurence J Kotlikoff. 1997. "Parental altruism and inter vivos transfers: Theory and evidence." *Journal of Political Economy*, 105(6): 1121–1166.

An, Lei, Yu Qin, Jing Wu, and Wei You. 2024. "The local labor market effect of relaxing internal migration restrictions: evidence from China." *Journal of Labor Economics*, 42(1): 161–200.

Andreoni, James. 1989. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of political Economy*, 97(6): 1447–1458.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro. 2017. "Measuring the sensitivity of parameter estimates to estimation moments." *The Quarterly Journal of Economics*, 132(4): 1553–1592.

Baker, James Edward. 1985. "Adaptive selection methods for genetic algorithms." In *Proceedings of an International Conference on Genetic Algorithms and their applications.* 101–111, Hillsdale, New Jersey.

Barczyk, Daniel, and Matthias Kredler. 2018. "Evaluating long-term-care policy options, taking the family seriously." *The Review of Economic Studies*, 85(2): 766–809.

Barham, Bradford, and Stephen Boucher. 1998. "Migration, remittances, and inequality: estimating the net effects of migration on income distribution." *Journal of development economics*, 55(2): 307–331.

Bayer, Christian, and Falko Juessen. 2012. "On the dynamics of interstate migration: Migration costs and self-selection." *Review of Economic Dynamics*, 15(3): 377–401.

Becker, Gary S. 1974. "A theory of social interactions." *Journal of political economy*, 82(6): 1063–1093.

Biao, Xiang. 2007. "How far are the left-behind left behind? A preliminary study in rural China." *Population, Space and Place*, 13(3): 179–191.

de Brauw, Alan, J Edward Taylor, and Scott Rozelle. 1999. "The Impact of Migration and Remittances on Rural Incomes in China."

Buck, Nicholas, and Jacqueline Scott. 1993. "She's leaving home: But why? An analysis of young people leaving the parental home." *Journal of Marriage and the Family* 863–874.

Chang, Hongqin, Xiao-yuan Dong, and Fiona MacPhail. 2011. "Labor migration and time use patterns of the left-behind children and elderly in rural China." *World Development*, 39(12): 2199–2210.

Chen, Xuan, and Merril Silverstein. 2000. "Intergenerational social support and the psychological well-being of older parents in China." *Research on aging*, 22(1): 43–65.

Chen, Yuanyuan, and Shuaizhang Feng. 2013. "Access to public schools and the education of migrant children in China." *China Economic Review*, 26 75–88.

Costa, Dora L. 1999. "A house of her own: old age assistance and the living arrangements of older nonmarried women." *Journal of Public Economics*, 72(1): 39–59.

Cox, Donald. 1987. "Motives for private income transfers." *Journal of political economy*, 95(3): 508–546.

Cox, Donald, Zekeriya Eser, and Emmanuel Jimenez. 1998. "Motives for private transfers over the life cycle: An analytical framework and evidence for Peru." *Journal of Development Economics*, 55(1): 57–80.

De Brauw, Alan, and Ren Mu. 2011. "Migration and the overweight and underweight status of children in rural China." *Food Policy*, 36(1): 88–100.

De Brauw, Alan, and Scott Rozelle. 2008. "Migration and household investment in rural China." *China Economic Review*, 19(2): 320–335.

Démurger, Sylvie, and Hui Xu. 2011. "Return migrants: The rise of new entrepreneurs in rural China." *World Development*, 39(10): 1847–1861.

Deng, F, JH Lv, HL Wang, JM Gao, and ZL Zhou. 2017. "Expanding public health in China: an empirical analysis of healthcare inputs and outputs." *Public health*, 142 73–84.

Du, Yang, Albert Park, and Sangui Wang. 2005. "Migration and rural poverty in China." *Journal of comparative economics*, 33(4): 688–709.

Ermisch, John. 1999. "Prices, parents, and young people's household formation." *Journal of urban economics*, 45(1): 47–71.

Ermisch, John, and Pamela Di Salvo. 1997. "The economic determinants of young people's household formation." *Economica*, 64(256): 627–644.

Foster, Andrew D, and Mark R Rosenzweig. 2001. "Imperfect commitment, altruism, and the family: Evidence from transfer behavior in low-income rural areas." *Review of Economics and Statistics*, 83(3): 389–407.

Gai, Qingen, Naijia Guo, Bingjing Li, Qinghua Shi, Xiaodong Zhu et al. 2021. "Migration costs, sorting, and the agricultural productivity gap."

Gemici, Ahu. 2011. "Family migration and labor market outcomes." *Manuscript, New York University.*

Goldberg, David E, and Kalyanmoy Deb. 1991. "A comparative analysis of selection schemes used in genetic algorithms." In *Foundations of genetic algorithms*. 1: Elsevier, 69–93.

Goldscheider, Francis, and Calvin Goldscheider. 1999. *The changing transition to adulthood: Leaving and returning home*. 17: Sage Publications.

Görlach, Joseph-Simon. 2016. "Borrowing Constraints, Migrant Selection, and the Dynamics of Return and Repeat Migration." *Unpublished manuscript.*

Hamilton, Bruce W. 2001. "Using Engel's Law to estimate CPI bias." *American Economic Review*, 91(3): 619–630.

Hansen, Lars Peter. 1982. "Large sample properties of generalized method of moments estimators." *Econometrica: Journal of the Econometric Society* 1029–1054.

Holland, John H, and John H Miller. 1991. "Artificial adaptive agents in economic theory." *The American economic review*, 81(2): 365–370.

Hu, Zhan, and Xizhe Peng. 2015. "Household changes in contemporary China: an analysis based on the four recent censuses." *The Journal of Chinese Sociology*, 2(1): 2–9.

Huang, Ping, and Shaohua Zhan. 2005. "Internal migration in China: Linking it to development." *Migration, development and poverty reduction in Asia,* 65–84.

Jia, Zhaobao, and Wenhua Tian. 2010. "Loneliness of left-behind children: a cross-sectional survey in a sample of rural China." *Child: care, health and development*, 36(6): 812–817.

Kaplan, Greg. 2012. "Moving back home: Insurance against labor market risk." *Journal of Political Economy*, 120(3): 446–512.

Katz, Eliakim, and Oded Stark. 1986. "Labor migration and risk aversion in less developed countries." *Journal of labor Economics*, 4(1): 134–149.

Kazianga, Harounan. 2006. "Motives for household private transfers in Burkina Faso." *Journal of development economics*, 79(1): 73–117.

Kim, Sang Won, Marcelo Olivares, and Gabriel Y Weintraub. 2014. "Measuring the performance of large-scale combinatorial auctions: A structural estimation approach." *Management Science*, 60(5): 1180–1201.

Lagakos, David, Samuel Marshall, Ahmed Mushfiq Mobarak, Corey Vernot, and Michael E Waugh. 2020. "Migration costs and observational returns to migration in the developing world." *Journal of Monetary Economics*, 113 138–154.

Li, Bingqin. 2006. "Floating population or urban citizens? Status, social provision and circumstances of rural–urban migrants in China." *Social policy & administration*, 40(2): 174–195.

Li, Haizheng, and Steven Zahniser. 2002. "The determinants of temporary rural-to-urban migration in China." *Urban Studies*, 39(12): 2219–2235.

Li, Xiaoming, Liying Zhang, Xiaoyi Fang, Bonita Stanton, Qing Xiong, Danhua Lin, and Ambika Mathur. 2010. "Schooling of migrant children in China: Perspectives of school teachers." *Vulnerable Children and Youth Studies*, 5(1): 79–87.

Liao, Pei-Ju. 2013. "The one-child policy: A macroeconomic analysis." *Journal of Development Economics*, 101 49–62.

Liu, Jing, and Shi Li. 2011. "Changes in Consumption Inequality in China."

Lockwood, Lee M. 2018. "Incidental Bequests and the Choice to Self-Insure Late-Life Risks." *American Economic Review*.

Lu, Yao. 2012. "Education of children left behind in rural China." *Journal of Marriage and Family*, 74(2): 328–341.

Ma, Zhongdong. 2002. "Social-capital mobilization and income returns to entrepreneurship: the case of return migration in rural China." *Environment and Planning A*, 34(10): 1763–1784.

Manacorda, Marco, and Enrico Moretti. 2006. "Why do most Italian youths live with their parents? Intergenerational transfers and household structure." *Journal of the European Economic Association*, 4(4): 800–829.

Massey, Douglas S. 1990. "Social Structure, Household Strategies, and the Cumulative Causation of Migration." *Population Index*, 56(1): 3–26.

McElroy, Marjorie B. 1985. "The joint determination of household membership and market work: The case of young men." *Journal of labor Economics*, 3(3): 293–316.

Meyerhoefer, Chad D, and CJ Chen. 2011. "The effect of parental labor migration on children's educational progress in rural China." *Review of Economics of the Household*, 9(3): 379–396.

Mitchell, Melanie. 1998. *An introduction to genetic algorithms.*: MIT press.

Mommaerts, Corina. 2018. "Are coresidence and nursing homes substitutes? Evidence from Medicaid spend-down provisions." *Journal of health economics*, 59 125–138.

Morten, Melanie. 2016. "Temporary migration and endogenous risk sharing in village india."Technical report, National Bureau of Economic Research.

Mu, Ren, and Alan De Brauw. 2015. "Migration and young child nutrition: evidence from rural China." *Journal of Population Economics*, 28(3): 631–657.

Mu, Ren, and Dominique Van de Walle. 2009. *Left behind to farm? Women's labor re-allocation in rural China.*: The World Bank.

Mullan, Katrina, Pauline Grosjean, and Andreas Kontoleon. 2011. "Land tenure arrangements and rural–urban migration in China." *World Development*, 39(1): 123–133.

Müller, Armin. 2016. "Hukou and health insurance coverage for migrant workers." *Journal of Current Chinese Affairs*, 45(2): .

Nelder, John A, and Roger Mead. 1965. "A simplex method for function minimization." *The computer journal*, 7(4): 308–313.

Pezzin, Liliana E, Robert A Pollak, and Barbara S Schone. 2007. "Efficiency in family bargaining: Living arrangements and caregiving decisions of adult children and disabled elderly parents." *CESifo Economic Studies*, 53(1): 69–96.

Qin, Jiang, and Björn Albin. 2010. "The mental health of children left behind in rural China by migrating parents: A literature review." *Journal of Public Mental Health*, 9(3): 4–16.

Rong, Zhao, Liu Yang, Yan Yuan et al. 2012. "Labor migration choice and its impacts on households in rural China." In *Annual Meeting of Agricultural and Applied Economics Association. Seattle, Washington.*

Rosenzweig, Mark R, and Kenneth I Wolpin. 1993. "Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India." *Journal of political economy*, 101(2): 223–244.

Rozelle, Scott. 1994. "Rural industrialization and increasing inequality: Emerging patterns in China's reforming economy." *Journal of comparative economics*, 19(3): 362–391.

Shen, Jianfa. 1993. "Analysis of urban-rural population dynamics of China: A multiregional life table approach." *Environment and Planning A*, 25(2): 245–253.

Sieg, Holger, Chamna Yoon, and Jipeng Zhang. 2023. "The impact of local fiscal and migration policies on human capital accumulation and inequality in China." *International Economic Review*, 64(1): 57–93.

Silberschatz, Abraham, Peter B. Galvin, and Greg Gagne. 2012. *Operating System Concepts.*: Wiley Publishing, , 9th edition.

Song, Yan, and Yves Zenou. 2012. "Urban villages and housing values in China." *Regional Science and Urban Economics*, 42(3): 495–505.

Song, Yan, Yves Zenou, and Chengri Ding. 2008. "Let's not throw the baby out with the bath water: the role of urban villages in housing rural migrants in China." *Urban studies*, 45(2): 313–330.

Song, Yang. 2014. "What should economists know about the current Chinese hukou system?" *China Economic Review*, 29 200–212.

Stanhope, Stephen A, and Jason M Daida. 1998. "Optimal mutation and crossover rates for a genetic algorithm operating in a dynamic environment." In *International Conference on Evolutionary Programming.* 693–702, Springer.

Strauss, John, Hao Hong, Xiaoyan Lei, Lin Li, Albert Park, Li Yang, and Yaohui Zhao. 2012. "Healthcare and insurance among the elderly in China: evidence from the CHARLS pilot." In *Aging in Asia: Findings from new and emerging data initiatives.*: National Academies Press.

Su, Shaobing, Xiaolo Li, Danhua Lin, Xiaoyenan Xu, and Maoling Zhu. 2013. "Psychological adjustment among left-behind children in rural C hina: the role of parental migration and parent–child communication." *Child: care, health and development*, 39(2): 162–170.

Taylor, J Edward, Scott Rozelle, and Alan De Brauw. 2003. "Migration and incomes in source communities: A new economics of migration perspective from China." *Economic Development and Cultural Change*, 52(1): 75–101.

Thom, Kevin. 2010. "Repeated circular migration: Theory and evidence from undocumented migrants." *New York University, mimeo.*

Todaro, Michael P. 1969. "A model of labor migration and urban unemployment in less developed countries." *The American economic review*, 59(1): 138–148.

Tombe, Trevor, and Xiaodong Zhu. 2019. "Trade, migration, and productivity: A quantitative analysis of china." *American Economic Review*, 109(5): 1843–72.

Wang, Fei-Ling. 2004. "Reformed migration control and new targeted people: China's hukou system in the 2000s." *The China Quarterly*, 177 115–132.

Wang, Yin-Chi, Ping Wang, Chong Yip, Pei-Ju Liao et al. 2017. "Educational Choice, Rural-urban Migration and Economic Development: The Role of Zhaosheng in China." In *2017 Meeting Papers.*(738): , Society for Economic Dynamics.

Wen, Ming, and Danhua Lin. 2012. "Child development in rural China: Children left behind by their migrant parents and children of nonmigrant families." *Child development*, 83(1): 120–136.

Wiemers, Emily E, Vladislav Slanchev, Kathleen McGarry, and V Joseph Hotz. 2017. "Living arrangements of mothers and their adult children over the life course." *Research on aging*, 39(1): 111–134.

Wilhelm, Mark O. 1996. "Bequest behavior and the effect of heirs' earnings: Testing the altruistic model of bequests." *The American economic review* 874–892.

Yang, Juan, and Muyuan Qiu. 2016. "The impact of education on income inequality and intergenerational mobility." *China Economic Review*, 37 110–125.

Yang, Xiushi. 2000. "Determinants of migration intentions in Hubei province, China: individual versus family migration." *Environment and Planning A*, 32(5): 769–787.

Ye, Jingzhong, and Pan Lu. 2011. "Differentiated childhoods: impacts of rural labor migration on left-behind children in China." *The Journal of peasant studies*, 38(2): 355–377.

Yip, Winnie, SV Subramanian, Andrew D Mitchell, Dominic TS Lee, Jian Wang, and Ichiro Kawachi. 2007. "Does social capital enhance health and well-being? Evidence from rural China." *Social science & medicine*, 64(1): 35–49.

Yoon, Chamna. 2017. "Estimating a Dynamic Spatial Equilibrium Model to Evaluate the Welfare Implications of Regional Adjustment Processes: The Decline of the Rust Belt." *International Economic Review*, 58(2): 473–497.

Zhang, Jane. 2011. "Do girls in China compete just as much as boys? Evidence from an experiment that predicts educational choice."

Zhao, Yaohui. 1999. "Leaving the countryside: rural-to-urban migration decisions in China." *American Economic Review*, 89(2): 281–286.

Zhao, Yaohui. 2002. "Causes and consequences of return migration: recent evidence from China." *Journal of Comparative Economics*, 30(2): 376–394.

Zhao, Yaohui. 2003. "The role of migrant networks in labor migration: The case of China." *Contemporary Economic Policy*, 21(4): 500–511.

Zhao, Zhong. 2005. "Migration, labor market flexibility, and wage determination in China: A review." *The Developing Economies*, 43(2): 285–312.

Zhou, Shu, and Monit Cheung. 2017. "Hukou system effects on migrant children's education in China: Learning from past disparities." *International Social Work*, 60(6): 1327–1342.

Zhu, Nong. 2002. "The impacts of income gaps on migration decisions in China." *China Economic Review*, 13(2-3): 213–230.

Zhu, Nong, and Xubei Luo. 2010. "The impact of migration on rural poverty and inequality: a case study in China." *Agricultural Economics*, 41(2): 191–204.

Zhu, Xi, John Whalley, and Xiliang Zhao. 2014. "Intergenerational transfer, human capital and long-term growth in China under the one child policy." *Economic Modelling*, 40 275–283.

For Online Publication

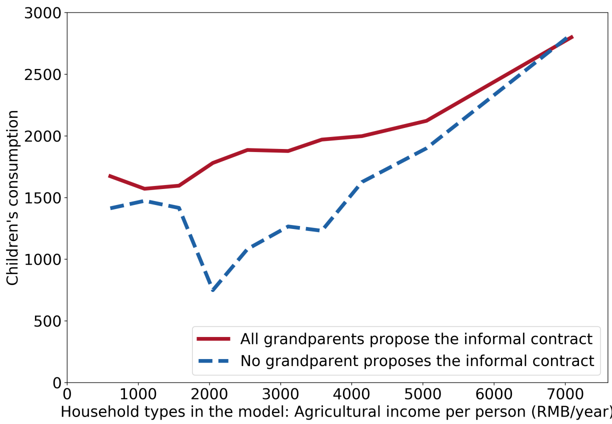
A Appendix A

A.1 Notations

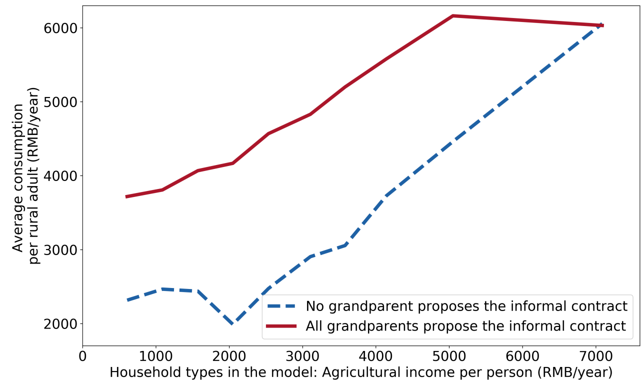
Table A1: Parameter definitions

Preference Parameters	
β	Annual discount factor
γ	CRRA on private utility
θ	Consumption-leisure trade-off
δ	Additive factor in utility functions
φ	Utility gain if child is enrolled in high school
λ	CRRA on child's education attainment
η_1	Grandparent's utility cost from illness
η_2	Grandparent's utility gain from healthcare
κ_1	Parent's guilt from low remittance
κ_2	Parent's guilt from no elder care
α^P	Parent's altruism towards the grandparent
α^G	Grandparent's altruism towards the parent
Assets, transfers, and consumption	
ψ	Price ratio between the rural and urban areas
ω	Grandparent's initial savings factor
B	Bequest
Tr_t	Transfer from parent to grandparent
c_t^P	Parent's daily consumption (nominal)
$\tilde{c}_t^P$	Parent's daily consumption (normalized)
c_t^G	Grandparent's daily consumption
A_j	Agricultural income per person per year for household type j
Time allocation	
T_{total}	Endowment of time
l_t^P	Parent's leisure
l_t^G	Grandparent's leisure
T_{rural}	Hours of labor supply in the rural area
T_{urban}	Hours of labor supply in the urban area
T_t^C	Time spent on childcare
T^G	Time spent on elder care
Health and healthcare	
c_t^h	Cost of healthcare
HC_t	Utility gain from health status
ρ^{death}	Effect of the parent's migration on grandparent's mortality risk
State variables	
X_t	Vector of all state variables
j	Household type
t	Age interval of the child
$contract^G$	Grandparent's contract status
$contract_t^P$	Parent's contract status
$urban_t^P$	Location of the parent
$urban_t^C$	Location of the child
s_t^G	Wealth of the grandparent
s_t^P	Wealth of the parent
h_t	Health status of the grandparent
g_t	Guilt of the parent
$enroll_t$	Child's school enrollment status
edu_t	Child's education attainment
Labor market	
w_t	Wage rate in the urban labor market
pr^{ump}	Migrating parent's unemployment rate when the child is over age 18
Child and education	
$tuition_t$	Tuition for the child
$p_t^{dropout}$	Child's dropout probability
c_t^{child}	Child's daily consumption
α	OECD equivalence scale
Government	
ρ^{care}	Childcare time reduction rate
$\rho^{consume}$	Child consumption subsidy rate
ρ^{edu}	Urban education subsidy rate
ρ^h	Government's health insurance coverage
c_{min}	Minimum consumption level
Estimation	
$\vec{\theta}$	Vector of structurally estimated param.
$Q(\vec{\theta})$	Model moments produced by $\vec{\theta}$
Q_0	Data moments
W	Weighting matrix
G	Jacobian matrix
Λ	Sensitivity matrix

A.2 Additional figures



(a) Children's consumption



(b) Adults' consumption

Figure A1: Adults' and children's consumption by informal contract status and income level

Note: The above figure show the effects of the informal contract on children's consumption levels in panel (a) and adult's consumption levels in panel (b) The consumption levels are estimated separately by the grandparents' informal contract status and their household's agricultural endowment level. In panel (a), urban consumption levels are normalized using the price ratio. Therefore, the average of children's consumption accounts for nominal rural consumption and normalized urban consumption. In panel (b), rural adults include parents who live in the rural area and all grandparents.

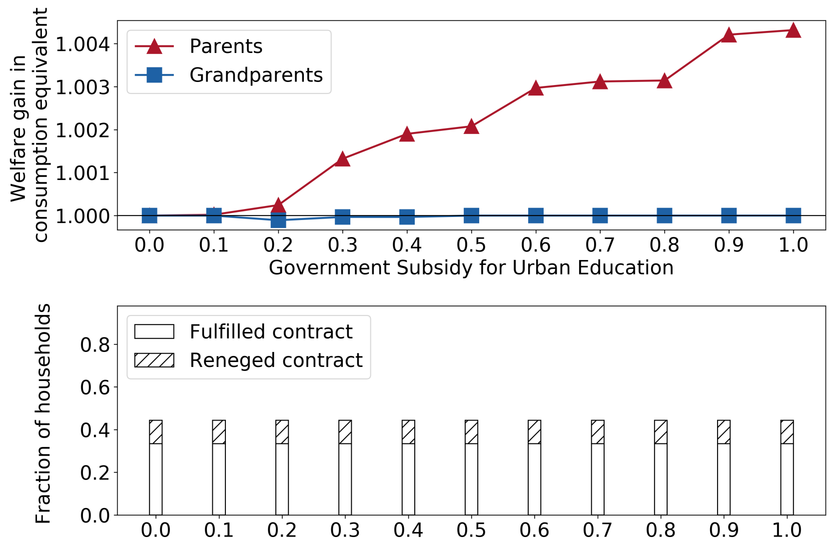


Figure A2: Education policy effects on welfare and the household's contractual space

Note: The upper panel of the above figure shows the effect of urban education policy on the welfare of the parents and grandparents. The welfare gains are presented in consumption equivalent term. The lower panel shows the policy effect on households' informal contract status. It presents the fraction of households with fulfilled contracts, and the fraction of households with reneged contracts. The statistics in both figures are calculated from the households with healthy grandparents when the children are born.

A.3 Additional tables

Table A2: Contribution of the data sources

Variables by topic	Dataset
Internal migration	
Movement	CFPS, RUMIC
Migrants' labor market	RUMIC
Household inter-generational behaviors	
Remittance	RUMIC, CHARLS
Daily care for ill grandparents	CHNS, CHARLS
Childcare	CFPS, CHNS
Grandparents' health status and medical costs	CHNS
Children's education and welfare	CFPS, RUMIC
Household wealth and consumption	CFPS, CHFS

Table A3: Economic factors faced by potential migrant

	Urban	Rural	Data sources
Nominal income (RMB/year)	23,428	2,702	CFPS
Hours at work (hours/week)	66	49	CHNS
Living expense (RMB/year)	16,464	2,389	RUMIC
Urban-rural price ratio	6.48	1	CHFS
Access to healthcare	Limited	Government pays 34%	(Deng et al., 2017)
Education cost for children (RMB/year)			
Kindergarten	2,023	1,066	
Primary school	1,576	765	CFPS
Middle school	1,643	993	
High school	3,589	3,164	

Note: The statistics in the urban economic environment are estimated on the sample of migrants, defined as people with rural Hukou living in the urban area.

Table A4: Effect of guardian on rural children's educational outcomes

Dependent variable:	School enrollment	Math score	Chinese score
Model:	Logit	OLS	OLS
Grandparent as primary care provider	0.18	0.06*	0.08
Guardian's years of schooling	0.22**	0.04***	0.02*
Guardian's years of schooling-squared	-0.006	0.002*	0.002***

Data source: CFPS, 2010-2014

Note: Restrict the sample to rural children attending rural schools. Regressions control for child's gender, age, and level of education. Parent is primary caretaker of the reference group.

Table A5: Benefits and costs of the informal contract to the parent, grandparent, and child

<table>
<tr><th></th><th>Benefits</th><th>Costs</th><th>Overall effect predicted by the model</th></tr>
<tr><td>Parent</td><td>• High income
(wage in the urban areas)
• Low consumption on children
(children live in rural area)
∗ Commodities are cheap
∗ Tuitions are subsidized
• More leisure time
(do not provide childcare)</td><td>• Fulfills the contract:
∗ Low income
(staying in rural area)
∗ Low consumption
(pay transfers, pay for grandparent's healthcare)
∗ Less leisure time
(elder care)
• Reneges on the contract:
∗ Guilt</td><td>All parents gain from the contract</td></tr>
<tr><td>Grandparent</td><td>• Additional income
(from remittance)
• Private care when ill
(parent fulfills contract)</td><td>• Income uncertainty
(parent reneges)
• Less leisure time
(provide childcare)
• Risk of not receiving elder care
(parent reneges)</td><td>Poor grandparents gain from the contract</td></tr>
<tr><td>Child</td><td>• Higher consumption
(from remittance)
• More funding for tuition
(migrant has more savings)</td><td>• Higher dropout probability
(left-behind children have poorer school performance)</td><td>Average education attainment drops by 0.07 years because of the contract</td></tr>
</table>

Table A6: Counterfactual Experiment Results on the Flow of Migration

Children's age	Actual Data	Government coverage for healthcare		Percentage reduction in childcare time	
		$\rho^h = 0.6$	$\rho^h = 0.9$	$\rho^{care} = 0.2$	$\rho^{care} = 0.6$
	(1)	(2)	(3)	(4)	(5)
0 - 5	0.329	0.124	0	0.398	0.656
6 - 11	0.268	0.349	0.473	0.074	0.074
12 - 14	0.188	0.064	0	0.274	0.275
15 - 17	0.182	0.112	0	0.386	0.685
18 - 20	0.150	0.001	0	0	0

Note: This table reports the predicted flow of migration from counterfactual experiments. The migration flows are summarized by the children's age ranges, consistent with the timeline of the model. Column 1 reports the fraction of migrant worker by children's age, estimated from the survey data (this should correspond to $\rho^h = 0.34$ and $\rho^{care} = 0$). Columns 2-5 report the model estimates of migration flow under different policy environment.

Appendix B

B Data Appendix

This appendix explains the pre-processing of the databases I used. Two aspects present challenges. First, the research question of my paper concerns three generations, as do the empirical evidence and moments presented in the paper. To link the information of these three generations, I match the individual level observations by family structure. Second, the timeline of the paper is pegged to the children's age. This requires that the children's age or age group is available in the datasets that I obtain time period-specific moments from.

B.1 China Family Panel Studies

In all of my empirical analyses, I restrict the sample to people with rural Hukou. To ensure a comprehensive dataset from the CFPS data, I match the community level, household level, and individual level surveys of a household by household ID.

I reshape the raw data into two different data structures. First, I construct an individual level panel data. The cleaned dataset I use for my analyses contains information on the following aspects:

- Hukou: the residential registration information in China, including the location (province level) and type (rural or urban).
- Household variables: annual agricultural and wage incomes, and family size.
- Individual variables:
  - demographics (gender, birth year, verified age, alive or dead)
  - migration behaviors (current residential location, reason of migration, and number of months per year in which the interviewee is away from home)
  - labor market information (annual individual income, annualized wage income)
  - education (actual education attainment, amount of education investment, enrollment status, current education level; scores on Chinese and math)
  - interaction with each of the interviewee's children (provide/receive money, help with housework, private care, and help with financial management)
  - health (for adults: condition of health, primary caregiver of the patient; for children: whether the child was breastfed, primary caregiver of the infants during the day and at night, whether the child lives with the parents)

I assume that rural people with children and between 25 and 50 years old are sampled from the *parent* generation. People between 55 and 75 years old are sampled from the living *grandparents*.

Second, I construct a single-thread three-generational household level panel data. Each observation contains information on a child, his/her father, mother, paternal grandparents and maternal grandparents, in a certain year.[53] For each person in this household, the data includes all the variables that the individual level panel data contains. In this way, the data is organized by the child's age, which is consistent with the structure of my model.

[53] Since many parents have more than one child, adults in the survey may be in multiple observations in the same year.

B.2 China Health and Nutrition Survey

The CHNS also has a multi-generational family structure, which allows me to match the grandparents' demographics, health condition, and healthcare behavior with their grandchildren's ages. The time allocation information is unique to this dataset among all five datasets I use.

B.3 China Health and Retirement Longitudinal Study

I construct an individual level panel data with detailed information on interactions between the parents and the grandparents from the CHARLS data. Each observation is unique to its grandparent's ID and year. All interviewees are sampled from people over 45 years old, who are asked about the interaction with each of their children. From each wave of the raw data, I collect the following information:

- Health condition of the grandparent: whether the grandparent is ill.
- Child and childcare: whether the grandparent provided childcare for each of his or her child's children
- Financial transfer: the presence of financial transfer activity, and the amount of transfers.
- Private care for ill grandparents: whether each of the children is providing private care for the ill grandparents.
- Migration of the parent (i.e. the interviewee's child): whether he or she lives in the rural area; whether the child's children are older than 16 years old.

Combining the five waves that are released, I obtain a panel data with a maximum length of 6 observations. The individual-level, elderly-oriented panel data allows me to find (1) the effect of childcare and migration on remittance and (2) the intertemporal correlation between childcare, financial support and private care when the grandparents are ill.

C Empirical Evidence Appendix

C.1 Urban labor market for rural migrants

Wage rate does not increase by experience. Table B1 shows the regression outputs on nominal wage and log wage of migrant workers in the urban area. I control for age, gender, and education background of the migrants.[54] The regression indicates that the effect of years of experience in the urban labor market on the migrant's wage is statistically significant, but the magnitude of the effect is very small. For example, the regression on nominal wage shows that one extra year of experience in the urban area increases the rural migrant's monthly wage income by 1.5%. Considering its small magnitude, I simplify the urban labor market in my model into a homogeneous market with fixed wage and hours.

Table B1: OLS regression of migrant's wage in the urban labor market

	Wage rate (RMB/month)	ln(wage)
Age	-8.52	-0.006***
Female	-414.9***	-0.188***
Years of schooling	37.3**	0.025***
Years of experience in the urban labor market	29.6***	0.011***
Constant	2214.5	7.545
Number of observations	3,045	3,045
R-squared	0.0159	0.0796

Data source: RUMIC 2009, migrant survey

Note: The sample restricts to full time workers with rural Hukou. It also restricts to people between 22 and 55 years old, and people receiving monthly wages higher than 500 RMB ($72) per month.

Urban wage drops for parents of college-aged children

When the household is in period 5, namely the period in which the child is between 18 and 21 years old, the parents are in their middle age and face increasing difficulty finding jobs in the urban area. Figure B1 shows that the wage rate drops sharply when the children of the migrants are between 18 and 25 years old.

I adopt the Todaro (1969)'s hypothesis by defining an unemployment rate for the migrants in period 5 and set the expected earnings in that period as the wage rate of the migrating parent. The unemployment rate, or job finding rate, for migrants is hard to measure or observe in micro-level data, so I incorporate the period 5 unemployment rate in the model as a parameter to be estimated. The difficulty is caused by a combination of several factors: First, unemployed migrants return to the rural area,[55] so the unemployment rate of rural migrants living in the urban area underestimates the actual unemployment rate of all rural-to-urban migrants. Second, migrants usually return to the rural area after being unemployed for some time, and the return migration movement complicates the following-up of individual-level surveys.' Third, many migrants are self-employed or take short-term jobs that keep them toggling between employment and unemployment, so the migrant's unemployment may not be a well-defined parameter, if I estimate it from the micro-level datasets.

[54] The RUMIC data has information on migrants' experience in the urban labor market, but does not have their household's agricultural income per person. On the other hand, CFPS data has all the variables in this regression except for working experience information. A wage regression on the CFPS data shows that, after controlling for age, gender, and education attainment, migrants' wages do not depend on the agricultural productivity of their rural households.

[55] Wong, Sue-Lin. "As China's economy slows, migrant workers head home." Reuters (2016).

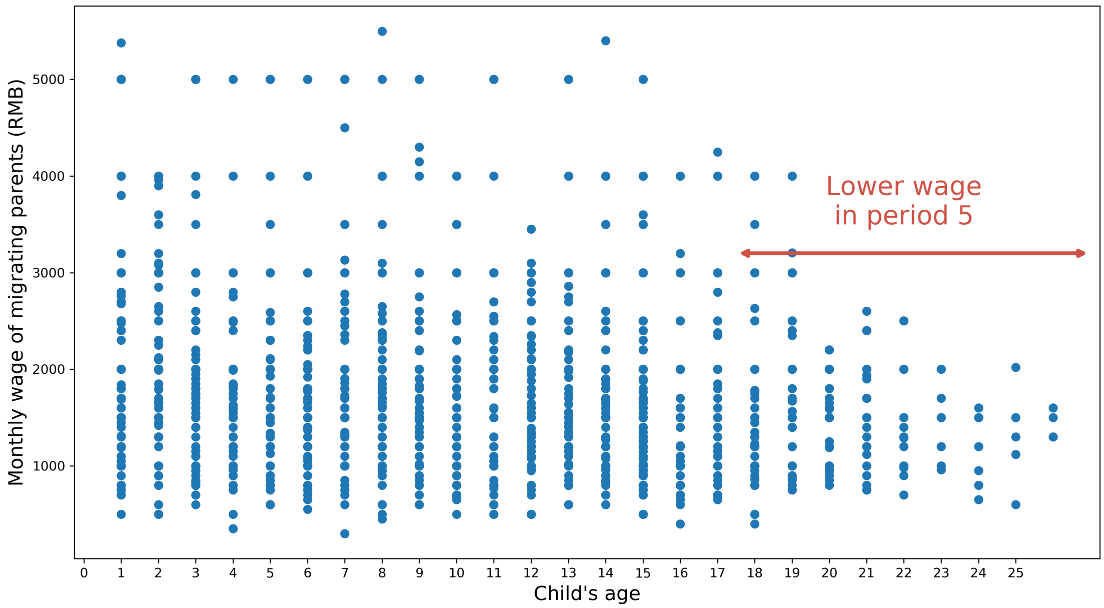


Figure B1: Distribution of migrating parent's wage by child's age

C.2 Rural household expenditure

Table B2 provides summary statistics of rural expenditures at the household level. The data source is CHFS 2012 survey, restricted to three-generational households located in the rural area who consume education and healthcare in order to focus on the subpopulation discussed in this paper. The consumption on food presented in the table is the sum of the amount spent on purchasing food and the market value of self-consumed produce from the household's farm. The consumption on housing includes the spending on renovation, expansion and new construction. The expenditure on social events includes cash and gifts to people outside the household during Chinese New Year, birthdays, weddings, funerals etc. The social events are either inevitable or unpredictable, so I do not model rural households' choice on social event spending.

Table B2: Expenditure of rural households

	Annual expenditure (RMB)	Percentage
Food and housing	15,048	38
Health care, net of insurance coverage	8,585	16
Education	4,762	14
Social events	5,327	12
Clothes and other commodities	2,092	6
Travel	1,945	5
Utility	1,403	4
Communication	1,238	4
Durable goods	502	1
Total	38,289	100

Note: Sample size is 684.

Note that the average expenditure by type and the average shares of each type of the expenditures are estimated separately on the same sample. For example, the average spending on education (4,762 RMB per year) is the average of education investment over all households. The average share of spending on education (14%) is the average of the share of education investment in each household. The household level heterogeneity, such as the varying tuition by education level attained, is the primary reason for the inconsistency between the rank of the amount of expenditure and the rank of the shares.

C.3 Temporary migration

Duration of rural-to-urban migration: As stated in the Section 3.1, rural-to-urban migration is mostly temporary for people without a college degree. In the RUMIC 2009 migrant's survey, a question asks "When did you first migrate out for work?". Based on the migrants' responses to this question, I recover the number of years that these current migrants have stayed in the urban areas by the time they are interviewed, and provide a histogram of their answers in Figure B2. It shows that 70% of the migrants have stayed in the urban areas for less than 10 years. Furthermore, it shows that almost all rural migrants ultimately return to the rural area.

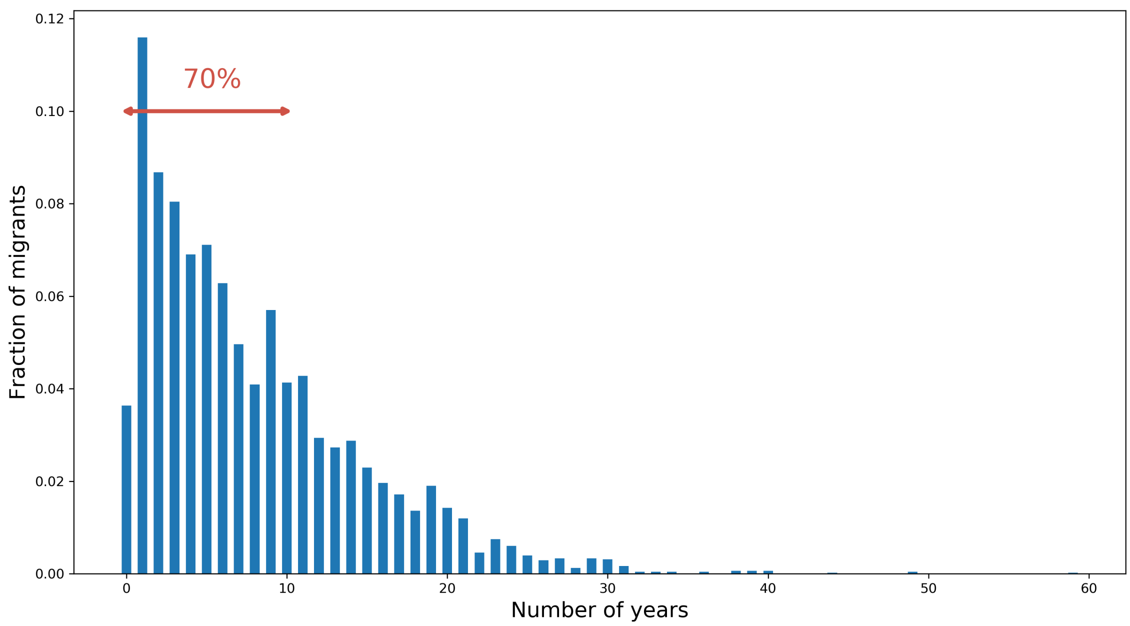


Figure B2: Length of stay in the urban area

C.4 Migration decisions

Determinants of migration decisions: As discussed in 3.3, the determinants of out and return migration decisions have been extensively studied in previous literature. Table B3 summarizes the relevant papers, the datasets they used, and their main findings.

My empirical analyses take the results from previous literature as given and further specify the household characteristics. I first use the CFPS 2010-2014 panel data to show that parents are more likely to migrate when the children are enrolled in school. The set of controls follows the findings from existing papers on this topic. The multi-generational household structure preserved by the CFPS data allows me to control for the parents' gender and age, children's gender and age, and the relationship between the children and their guardian. Second, I use the CHARLS 2013-2015 panel data to show that conditioning on grandparents being ill, parents are more likely to remain in the rural areas if the grandparents had provided childcare in the past. As a survey concentrated on the welfare and health of the elderly in China, the CHARLS data only contains information on the number of grandchildren

Table B3: Literature review on determinants of rural-to-urban migration decision in China

Paper	Data source	Findings
Zhao (1999)	Surveyed 7,410 individuals from 1,820 households in Sichuan province	age (-), edu (-), family size (+), land and case (-), economic development of home village (-)
Li and Zahniser (2002)	CHIP 1995	age(+) age-sq (-), ethnic minority (-), marital status (-), edu (+), edu-sq(-), family size (+), number of pre-school kids (-), farm income (-)
Yang (2000)	Surveyed 4,368 rural individuals in Hubei province	Migration decrease by wealth
Zhao (2002)	Surveyed 824 households from 6 provinces	return migration depends on age (+), education (+), family size (-)
Du et al. (2005)	CPMS 1997-2001	Migration decreases by wealth of the rural household

under 18 years old, and the gender and age of the parents. My regression analysis controls for all three aspects of characteristics on the extended family members.

Table B4: OLS regression on migration behavior

Dependent var: Parent migrates	(1)	(2)
Child is enrolled in school	0.053(0.013)***	
Grandparent		
Ill now, provided childcare		-0.04(0.02)**
Ill now, did not provide childcare		0.01(0.01)

Note: Standard errors are provided in parentheses.

Reason of staying in the rural areas: RUMIC 2008 and 2009 ask the rural residents about the "main reason for not migrating to find work". A tabulation of the responses is provided in Table B5. The sample restricts to married adults with children who are currently living in the rural areas. The main reason that middle-aged parents (between 20 and 45 years old) do not migrate is because they need to stay to take care of family members. The survey question indicates that the family members who need private care are mainly young children and ill grandparents. For rural residents over 50 years old, most stay in the rural area because their working ability is limited by their age, health condition, or disability. The table also shows that the rural people's belief in the labor market in the urban areas is very optimistic. Worries about the urban labor market condition rarely become the major obstacle to migration.

C.5 Financial transfer from parents to grandparents

As stated in Section 3.3, the act of sending remittance and the specific amount of the remittance both depend on the parent's migration experience, the grandparent's childcare behavior, and the grandparent's health condition. Here I provide the output of the regression analyses I ran on CHARLS data to justify the statement. In both regressions, I restrict the sample to rural households with children, and control for the gender and age of the parents and grandparents in the household. In column (2) of Table B6, the OLS regression on the amount of transfer, I further restrict the sample to households

Table B5: Reasons for staying in the rural areas by age group of adults

Age group	Taking care of family members	Age, illness, disability	Could not find work in urban areas	Agricultural production
20-24	0.61	0.05	0.16	0.18
25-29	0.53	0.05	0.13	0.29
30-34	0.51	0.04	0.12	0.32
35-39	0.50	0.04	0.08	0.38
40-44	0.42	0.11	0.07	0.40
45-49	0.29	0.21	0.08	0.42
50-54	0.20	0.46	0.03	0.29
55-59	0.18	0.59	0.03	0.20
60-64	0.13	0.73	0.01	0.13

Note: The cell count for every percentage statistic is over 100 people.

with positive transfer from the parents to grandparents. The significant and positive coefficients in the table imply the following facts:

1. Parents who are currently in the urban areas send more transfers.

2. Parents who are currently in the rural areas but have migrated before also send more transfers.

3. Grandparents who are currently caring for grandchildren receive more remittance.

4. Grandparents who are not taking care of grandchildren but have provided childcare in the past also receive more remittance.

5. Grandparents in poor health receive more remittance.

In my model, the informal contract between the parent and the grandparent implies all the above facts, especially #3 and #4. In addition, my model captures fact #1 and #5 by setting a positive lower bound on the amount of remittance when a parent migrates when a grandparent is ill.

Table B6: Remittances depend on migration, childcare, and grandparent's health

	Whether the parent sends transfer	Amount of transfer (RMB/year)
Parent's migration experience		
In the urban area now	0.12(0.01)***	550.10(62.91)***
Migrated before	0.11(0.01)***	450.67(64.12)***
Grandparent's childcare experience		
Taking care of children now	0.04(0.01)**	518.59(60.53)***
Took care of children before	0.04(0.01)**	297.63(60.35)***
Grandparent is ill	0.02(0.01)*	

Data source: CHARLS

Note: Regressions restrict to sample of households with children, and control for gender and age of the parents and grandparents. Regression on the amount of transfer restricts to households with positive transfer.

C.6 Left-behind children

Main reasons of leaving the children behind: RUMIC asks the migrants why their children do not live with them. It also asks the rural children's primary caretaker why the children do not live with their parents. Despite the small inconsistency between the answers from these two samples, the top three reasons for parents migrating without children are high education costs, high living costs, and the lack of childcare.

My model incorporates high education costs and evaluates a policy aimed at lowering urban education costs. The high living cost is modeled through the price ratio between rural and urban consumption, so that the welfare gain of $1 in the rural area has to be matched by $6.48 additional spending in the urban area. The lack of childcare is represented by the hours the guardian must spend on childcare. Given the long hours of labor supply in the urban areas, the marginal cost of the time spent on childcare is higher than the cost to a guardian living in the rural areas.

Table B7: Main reasons for leaving children behind

Reason	Rural survey	Urban survey
High cost of attending school or Kindergarten	0.1508	0.1093
High urban cost of living	0.2773	0.4079
Lack of childcare	0.2974	0.2141
No access to schools	0.0174	0.0083
Education in hometown is better	0.1415	0.0845
Other	0.1156	0.1759

Data source: RUMIC 2008 and 2009

Note: The sample for this question restricts to interviewees from migrant households. The sample size of the rural survey is 4,543. The sample size of the urban survey is 2,177.

Literature review: Left-behind children are an important consequence of temporary migration in China. The magnitude and the welfare of left-behind children have been studied by many economists and sociologists. Table B8 summarizes a selected set of papers on left-behind children in China.

Table B8: Literature review on left-behind children in China

Paper	Findings
	A: Magnitude of left-behind children
Jia and Tian (2010)	28.3% of rural children are left behind
	B: Educational outcomes
Wen and Lin (2012)	Left behind children are disadvantaged in health behavior and school engagement
Lu (2012)	Parent migration has no effect on children's education
Meyerhoefer and Chen (2011)	Migrants' children have higher education attainment
Chen and Feng (2013)	Access to public schools improves the quality of education for migrants' children
	C: Nutritional and psychological outcomes
Su et al. (2013)	Left behind children have poor psychological condition
De Brauw and Mu (2011)	No significant relationship between parents' migration and children's nutrition
Mu and De Brauw (2015)	No significant nutrition effect
Ye and Lu (2011)	Left-behind children receive low quality childcare

Guardian of rural children: Many rural households with three generations live together in the rural area. I use the RUMIC data to tabulate the primary caretaker of the rural children by the locations of the children and their parents. When children live with their parents, no matter where they live, many are cared for by their mother. Less than 25% of the rural children are primarily looked after by their grandparents. On the other hand, more than 70% of the left-behind children are cared for by their paternal grandparents. In my model, I let the parent be the guardian of all children who live with their parents, and let the grandparent be the guardian of those left behind by migrating parents.

Table B9: Guardian of rural children by parent's and child's migration status

Primary caretaker of the child	Children migrate with parents	Non-migrant's children	Left-behind children
Mother	0.3475	0.4951	0.1122
Father	0.0141	0.0463	0.0072
Maternal grandparents	0.0566	0.0321	0.0365
Paternal grandparents	0.1576	0.1968	0.7185
Day care	0.1111	0.0554	0.0556
Nanny	0.002	0.1735	0.0695
Other	0.3111	0.1735	0.0695
Sample sizes	495	18,943	1,943

Data source: CFPS 2010-2014

D Model Appendix

D.1 Agricultural income

In my model, I assume that rural income is a constant for each healthy adult member of the household. They key assumption of this setting is that rural income per capita does not decrease as family size increases. Here I provide a brief justification of this assumption. I run a regression of agricultural income per capita on the number of adults living in the household on 6,723 rural families in the CFPS 2010 data. The R-squared statistic of the regression is 0.0009, which indicates that the assumption of constant rural income at individual level is not unrealistic. Intuitively, this means that in the rural areas, the binding constraint in rural production is human capital instead of natural resources or production equipment. A returned migrant may expand agricultural production and increase the household's agricultural income.

D.2 Grandparent's initial wealth

My model starts when the child is born. I assume that the grandparents of the newborn have savings while the parents do not. The stock of savings of the grandparents depends on the household's agricultural productivity. Therefore, I use the CFPS 2010-2014 data to compute the ratio ω between assets $s_0^{G_j}$ and agricultural income per year A_j, where j indicates household type, categorized by rural income level.

$$s_0^{G_j} = \omega \times A_j$$

Among the 861 households with children under 2 years old, with a median of 4.14. Then I run a regression of ω on A_j. The R-squared statistic of the regression is 0.0023. Therefore, I assume a constant ratio between grandparent's initial wealth and the household's agricultural productivity, and this ratio $\omega = 4.14$.

D.3 Transition in health condition

In my model, I assume grandparents have no chance of recovery from illness. This assumption is supported by a tabulation of the transition of grandparent's health condition using the CFPS 2010-2014 data. Table B10 shows that only 2.76% of rural people over 55 years old may recover from illness. When I prolong the time difference to 4 years, the fraction of recovered grandparents increases slightly to 3.19%. My model makes a reasonable assumption by setting this recovery probability to zero.

Table B10: Change in health status of rural people over 55 years old

	Health condition (two years before)	
Health condition (current)	Healthy	Ill
Healthy	0.3740	0.0276
Ill	0.5211	0.8921
Dead	0.1049	0.0804

Data source: CFPS 2010-2014

Note: The statistics of the above table are obtained from a sample of 2,344 rural people over 55 years old. Note that the time difference between the two interviews of the CFPS panel data is 2 years.

D.4 Cost of healthcare

My model specification on health expenses has two components. First, the price of healthcare for a living grandparent is constant across household types. Second, once the grandparent dies, a one-time extra cost is charged. This extra cost is first deducted from the grandparent's savings. If her bequest is not enough to cover the cost, then the remaining debt is paid by the parent. All these charges are independent of the informal contract. Here I justify the settings of homogeneous costs and the prices of the two expenses.

Healthcare when the grandparent is ill: The intuition behind setting a homogeneous healthcare expense is that, in rural China, most people do not go to the hospital as long as they can bear their pain.[56] By the time the illness is causing unbearable pain, they are likely to be in very poor health. They go to the hospital and find that the price of treatments that can relieve the pain or illness is very high. At that stage, if they pay the high price, they may live; if they do not pay this fixed price or cannot afford it, they are unlikely to recover or live much longer afterwards.

The specific price of healthcare is estimated from the CHNS dataset. I restrict the sample to rural households with ill elderly, and restrict to the sample in which the ill elderly are alive in the next period. The median of the annual expense on healthcare is 3,305 RMB ($477). The magnitude is similar to the estimate from the public health literature (Strauss et al., 2012).

Healthcare in the last year of the grandparent's life: I observe a significant rise in healthcare expenditures right before and after the patient's death. In CHNS, the average annual total spending on healthcare in rural households with dying grandparents is 11,514 RMB ($1,662). Note that this is the total amount prior to any health insurance reimbursement. In the Chinese Longitudinal Healthy Longevity Survey (CLHLS), there is a direct question concerning the total out-of-pocket medical cost paid by the family in the last year of the elder person's life. The average spending, over 327 deceased elderly, is 8,756 RMB ($1,264). Because the spending recorded in the CLHLS data is post-reimbursement cost, I can recover the total cost $\frac{8756}{1-0.34} = 13267$, which is roughly consistent with the estimate from CHNS. Therefore, in my model, I assume that right after the grandparent's death, the household spends an extra 8209 RMB on healthcare (I use the CHNS estimate, so 11,514 - 3,305 = 8,209 RMB).

[56] Roberts, Dexter. "Chinaś Rural Poor Bear the Brunt of the Nation's Aging Crisis." Bloomberg (2017).

E Estimation Appendix

E.1 Externally estimated parameters

E.1.1 Distribution of agricultural income of rural households

Rural households have heterogeneous agricultural productivity. Appendix D.1 discusses why income levels are measured at the individual level instead of the household level. Here I describe the construction of the distribution of agricultural income in the CFPS data. First, I define the number of people who participate in rural production within each household. It is a count of rural people whose residential location is the same as the rural household's location, aged between 18 and 70. Then I assume the agricultural income of the household is shared equally among these participants. In the estimation, I discretize the continuous distribution of per capita agricultural income of the households into 10 types, i.e. $J = 10$.

Table B11: Agricultural income distribution

Agricultural income (RMB per person per year)	Fraction of households
500	0.1636
1,000	0.1824
1,500	0.1265
2,000	0.1037
2,500	0.0841
3,000	0.0878
3,500	0.0308
4,000	0.0510
5,000	0.0892
7,000	0.0809

Data source: CFPS 2010-2014

E.1.2 Rural-urban price ratio

Daily consumption in the model is defined as the total of food, clothes, commodities, and housing, so a ratio of the prices of a certain good, e.g. rice, is not representative, and a ratio of the total price of the same group of goods is not feasible since consumption bundles differ by location. Therefore, I estimate the price ratio from the Engel curves for food expenditure among the total daily consumption.

I use a rural-urban price ratio to standardize the daily consumption levels in the rural and urban area into comparable purchasing powers. Specifically, I assume that two households with the same fraction of expenditure on food shall have the same level of standard of living. I estimate the household-level Engel curves using the CHFS data on rural households (Figure B3 in Appendix C).

Let $F_{food} \in \{0, 1\}$ denote the fraction of food expenditure, and let $c^R(F_{food})$ and $c^U(F_{food})$ denote the total daily consumption corresponding to a given fraction F_{food}, respectively. The estimated price ratio is a factor ψ^* such that

$$\psi^* = \underset{\psi}{argmin} \int_0^1 \mid \psi \times c^R(F_{food}) - c^U(F_{food}) \mid \; dF_{food} \tag{27}$$

The estimated optimal value for ψ is 6.48.

Technically, the ratio is estimated in the following steps using the CHFS 2013 dataset.

1. Define the total expenditure of the household as the sum of
   - Annualized consumption on food, self-consumed food, utility, commodities, housekeepers, local transportation, telephone and internet services, and entertainment;
   - Annual consumption on clothes, housing, durable goods, education, traveling;
   - Annual transfers to other relatives outside the household;
   - Annual spending on gifts and social events.
2. Compute the fraction of total expenditure spent on food within each household
3. For each residential location, i.e. rural or urban, I estimate an Engel curve by a locally weighted scatter plot smoothing (LOWESS) regression. The predicted Engel curves are shown in Figure B3.
4. I find the minimum mean square estimator of ψ, the constant ratio between rural and urban consumption that minimizes the distance between the rural and urban Engel curves. The objective function used to estimate ψ is equation (27).

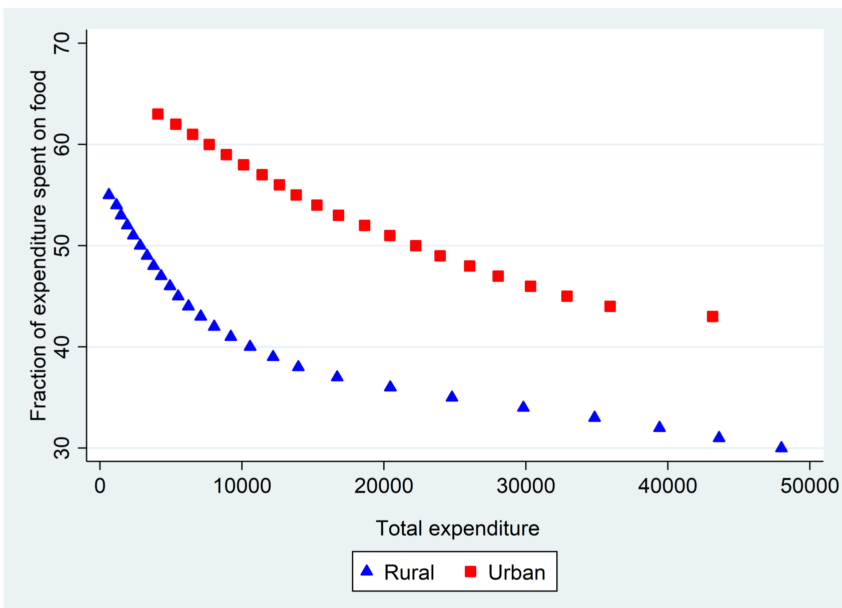


Figure B3: Estimated Engel curves

Furthermore, to check whether household composition affects the Engel curve outcomes, I run a robustness check. I run a regression of the household-level consumption on food on the number of children in the household. The coefficient on number of children is insignificant coefficient.

For additional reference, I provide the ratio between the price of rice in the urban versus rural area as a secondary reference for my estimate of the price ratio. In 2012, the price of rice in major cities in China was 5.52 RMB per kilogram (National Bureau of Statistics, February 2012), and the price at which the government bought rice from farmers was 1.21 RMB per kilogram (Ministry of Agriculture and Rural Affairs of China, February 2012). Therefore, the urban-rural price ratio for rice is 4.56.

E.1.3 Time allocation

The specific hours needed for work and private care activities are estimated from multiple sources.

Labor supply in the urban areas: I use the RUMIC 2009 migrant survey to estimate the average hours of labor supply for rural migrants. Among 2,739 observations, 90% work for more than 40 hours a week, and the average hours of labor supply is 66 hours (with a standard deviation of 19 hours).

Labor supply in the rural areas: I use the RUMIC 2009 rural survey to estimate the average hours of labor supply for rural residents who participate in agricultural production. The data contains 4,373 rural adults who spend time on farming. Their average hours of labor supply is 49 hours (with a standard deviation of 16 hours).

Time spent on private care: In my model, I incorporate three types of time-costly private cares. All three are estimated from the CHNS data. I combine the child's survey with the adult's survey to obtain the exact age of the child that the parent takes care of.

1. Childcare for children over 6 years old: 12 hours per week.
2. Childcare for children under 6 years old: 27 hours per week.
3. Elder care: 10 hours per week.

E.1.4 Cost of healthcare

The total costs of healthcare are measured from the CHNS data. The distribution is highly bimodal, with one peak at nearly zero spending and another peak at around 3,305 RMB ($483) per year. In addition, I observe that, on average, rural households spend an additional 8,209 RMB ($1,198) on healthcare in the year in which an elder family member dies. This is likely a combination of spending on some last-minute emergency procedure and funeral expenses but is nevertheless an unavoidable cost related to the grandparent.

E.1.5 Transition in grandparent's health status

Baseline transition probabilities: In this section I discuss the method to obtain the baseline transition probability of the grandparent's health condition by children's age group.

The probabilities have to be measured on the subsample of rural households above the poverty line that do not have migrants, in order to separate the effect of lack of money for healthcare and parent's migration from the change in health condition because of age. The grandparents have to be grouped by the child's age instead of the grandparent's own age, to fit the design of my model. The distribution of grandparent's health by children's age group estimated from the CFPS data is provided in Table B12.

Table B12: Distribution of grandparent's health condition by child's age group

	Grandparent's health condition		
Child's age	Healthy	Ill	Dead
0-5	0.6239	0.2847	0.0914
6-11	0.4569	0.2640	0.2791
12-14	0.3490	0.1911	0.4598
15-17	0.1959	0.1649	0.6395

Data source: CFPS 2010-2014

Because available panel datasets are short, the subsample in which I observe a change in health condition in follow-up interviews with the same person are very small. Therefore, I recover the transition probabilities from the above distribution.

- Pr(ill at $t+3$ | healthy at t) $= 0.15$
- Pr(dead at $t+3$ | ill at t) $= 0.65$

Effect parent's migration on grandparent's mortality: I use the CFPS data, and restrict the sample to grandparents in the rural area over 55 years old. The sample also restricts to grandparents who were ill 3 year before. I assume that parent's migration is equivalent to the scenario in which parents do not provide elder care. Table B13 provides the regression output of a Logit regression of grandparent's mortality on the grandparent's age, gender and an indicator for whether the parent provided elder care when the grandparents were ill. The regression indicates that the lack of elder care increases mortality risk. Using the baseline mortality rate of 0.65, I found that parent's migration when the grandparent is ill increases grandparent's mortality rate by 67%.

Table B13: Effect of parent's private care on grandparent's mortality

Dependent variable: grandparent alive now	Coefficient	Standard Error
Parent provided private care in the past	0.513**	0.195
Male	0.675***	0.158
Age	-0.083***	0.007
N	5,267	
Adjusted R-squared	0.081	

Data source: CFPS, 2010-2014

E.1.6 Education dropout probability

As stated in Section5.1.1, direct estimation of

$$Pr(dropout_t \mid tuition_t > 0, guardian, edu^{guardian}) \tag{28}$$

is not feasible due to data limitation.

I take advantage of the independence between the effect of the guardian's education on the child's education and the relationship between the guardian and the child, and estimate the following two sets of probabilities in CFPS data:

- $Pr(dropout_t \mid tuition_t > 0, edu^{guardian})$: Table B14
- $Pr(edu^{guardian} \mid guardian)$: Table B15

The two sets of probabilities are presented in Tables B14 and B15. The matrix multiplication of these two marginal probability distribution results in the desired joint distribution.

Table B14: Children's dropout probabilities by guardian's education attainment

	Guardian's education				
Child's education level upon dropping out	Illiterate	Primary school	Middle school	High school	College or above
Illiterate	0.0100	0	0	0	0
Primary school	0.1268	0.0387	0.0242	0.0155	0
Middle school	0.0546	0.0506	0.0312	0.0553	0.0800
High school	0.1324	0.1556	0.1510	0.1216	0.15

Data source: CFPS 2010-2014

Table B15: Distribution of parent and grandparent's education attainment

	Parent	Grandparent
Illiterate	0.2111	0.6684
Primary school	0.2714	0.2209
Middle school	0.3463	0.0845
High school	0.0947	0.0221
College or above	0.0768	0.0040

Data source: CFPS 2010-2014

Passing probability in the National College Entrance Exam: I use national public information from the 2004 National College Entrance Exam. In that year, 7.23 million people participated in the exam. The national failure rate was 39%. Therefore, 4.41 million people entered college in that year. Among the participants, 3.98 million were rural residents. Among the admitted students, 27% were rural students. Therefore, 1.19 million rural students entered college. The overall passing rate for rural children is $\frac{1.19}{4.41} = 0.2698$.

E.2 Sensitivity of the parameters to the moments

The Jacobian matrix $\hat{G}$ is measured following the equation on page 23 of Kim et al. (2014). Table B16 reports the estimated Λ, following the formula in equation (22). The tables show that the 10 internally estimated parameters are identified by the moments.

The calculation of the sensitivity matrix relies on the Jacobian matrix, estimated by taking partial derivative of each moment with respect to each parameter locally around the point estimate. This implies that the sensitivity is a measure for the contribution of each moment in locally identifying each parameter. Some moments, while may not directly contribute to the identification from a local search perspective, vary as I change the parameter values with a larger magnitude. For this reason, I keep the moments in the model estimation even if Table B16 shows that they do not locally identify any parameter.

Table B16: Sensitivity Matrix

Name of the Moment			β	γ	θ	δ	η_2	ϕ	λ	κ_1	κ_2	pr^{ump}
Conditioning on grandparents provided childcare in the past, the fraction of parents who:												
Child's age	Parent provides											
	transfer	elder care										
Age 0 to 14	Yes	Yes	0.022	-0.423	0.096	-0.011	0.047	0.065	0.043	0.003	0.004	-0.168
Age 0 to 14	Yes	No	0.006	-0.108	0.024	-0.003	0.012	0.089	0.020	-0.001	-0.002	-0.136
Age 0 to 14	No	Yes	-0.037	0.683	-0.151	0.019	-0.086	-0.206	-0.078	-0.005	-0.002	0.376
Age 15 to 20	Yes	Yes	-0.002	0.063	-0.013	0.001	-0.008	-0.047	-0.008	0.020	0.010	0.081
Age 15 to 20	Yes	No	-0.004	0.022	-0.004	0.002	-0.010	0.044	0.001	-0.014	-0.019	-0.091
Age 15 to 20	No	Yes	-0.003	0.066	-0.014	0.001	-0.009	-0.048	-0.008	-0.009	0.010	0.083
Frac. of migrants, child age 0-5			0.083	-1.724	0.385	-0.038	0.073	0.168	0.201	-0.006	-0.006	-1.179
Frac. of migrants, child age 6-11			-0.001	0.075	-0.018	-0.001	0.031	0.036	-0.015	0.006	0.004	0.152
Frac. of migrants, child age 12-14			0.002	-0.061	0.014	0.000	-0.014	0.000	0.012	-0.003	-0.002	-0.108
Frac. of migrants, child age 15-17			0.022	-0.390	0.087	-0.011	0.010	-0.062	0.009	-0.003	0.013	0.427
Frac. of migrants, child age 18-20			0.015	-0.406	0.091	-0.006	0.006	0.105	0.085	-0.003	-0.017	-1.011
Frac. of left-behind children age 0-5			0.083	-1.724	0.385	-0.038	0.073	0.168	0.201	-0.006	-0.006	-1.179
Frac. of left-behind children age 6-11			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frac. of left-behind children age 12-14			0.001	-0.039	0.009	0.000	-0.016	-0.019	0.008	-0.003	-0.002	-0.080
Frac. of children enrolled in primary school			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frac. of children enrolled in middle school			0.001	0.035	-0.008	-0.001	0.010	-0.436	-0.047	0.000	-0.001	0.255
Frac. of children enrolled in high school			-0.001	-0.007	0.002	0.001	-0.002	0.078	0.073	0.000	0.000	-0.433
Avg. annual remittance			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frac. of grandparents receiving remittance			0.061	-1.276	0.285	-0.028	0.051	0.120	0.150	-0.005	-0.005	-0.883
Avg. years of schooling of children at age 22			0.000	0.003	-0.001	0.000	0.001	-0.027	0.002	0.000	0.000	-0.013
Frac. of ill grandparent not receiving health care			0.223	-5.975	1.333	-0.088	0.002	-0.083	0.666	0.000	0.015	-3.504
Frac. of ill grandparents left behind			0.000	0.021	-0.005	0.000	0.009	0.010	-0.004	0.002	0.001	0.042
Frac. of ill grandparents, child age 0-5			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frac. of healthy grandparents, child age 0-5			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frac. of ill grandparents, child age 6-11			0.003	-0.060	0.013	-0.001	0.003	0.006	0.007	0.000	0.000	-0.041
Frac. of healthy grandparents, child age 6-11			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Frac. of ill grandparents, child age 12-14			-0.007	0.135	-0.030	0.003	-0.008	-0.016	-0.015	0.000	0.000	0.084
Frac. of healthy grandparents, child age 12-14			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Avg. consumption per rural adult			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

E.3 Weighting matrix

The weighting matrix is a diagonal matrix. When the moment is a fraction, the weight is 1. When the moments are numerical levels of consumption or transfer, the weight is the reciprocal of the variance of the data moment. Specifically, the weight for the average annual remittance is 1.6×10^{-7} (the mean[sd] of the data moment are 4,099[2,515]), the weight for the average for children's years of schooling is 0.027 (the mean[sd] are 10.130[6.135]), and the average annual consumption per rural adult is 1.1×10^{-7} (the mean[sd] are 3394[3053]).

E.4 Identification of the Internally Estimated Parameters

In this section, I explain the main source of identification in the estimation of each parameter. Table B17 lists the 11 structurally estimated parameters and their key identification sources.

Table B17: Parameters estimated in the structural model

Description	Symbol	Main sources of identification
Annual discount factor	β	Consumption behaviors Migration movement Contract decisions
CRRA for private utility	γ	
Consumption-leisure trade-off	θ	
Coefficient on additive component of utility function	δ	
Grandparent's utility gain from healthcare when ill	η_2	Consumption on healthcare
Utility gain if child is enrolled in high school	φ	Education enrollment rates Average years of schooling
CRRA for children's education attainment	λ	
Parent's guilt from low remittance	κ_1	Frac. of parents fulfilling each component of the contract
Parent's guilt from not taking care of grandparent	κ_2	
Migrating parent's unemployment rate when the child is over age 18	pr^{ump}	Migration movement

The discount factor: (β), as detailed in equations (17) and (19), influences future utility values. Parents with higher discount factors tend to prioritize remaining in rural areas to uphold informal contracts, rather than migrating for greater income. Similarly, grandparents with high discount factors often propose contracts exchanging early leisure for later benefits like increased consumption and healthcare access. The identification of the discount factor is based on the proportion of migrants over time and parents' adherence to informal contract components.

The CRRA coefficient: (γ in equation (3)) impacts the marginal utility of additional consumption and the incentive to migrate for higher income. Parents with a higher γ value gain less utility from increased urban wages and are less inclined to breach informal contracts. Their relative indifference to consumption decreases also motivates investments in their children's education, likely boosting school enrollment rates. Conversely, grandparents with a high γ value are less sensitive to consumption changes, causing them to allocate more towards healthcare. The identification of the CRRA coefficient utilizes data on migration rates, school enrollment, healthcare spending by grandparents, and parental adherence to informal care commitments.

Consumption-leisure trade-off: (θ in equations (3) and (5)) shapes parental decisions between urban and rural labor markets, migration choices for their children and care decisions for grandparents. A lower θ value corresponds with a reduced likelihood of parents leaving children behind and influences their balance of financial support against elder care in informal contracts. Grandparents with a lower θ tolerate lesser utility loss from childcare, prompting them to favor informal contracts that exchange childcare for additional income for daily needs and healthcare. Thus, θ is identified using data on the proportions of migrating parents, left-behind children—particularly in the first period—and parental compliance with informal contract components.

Children's education parameters: (φ and λ in equations (3), (15), and (16)) directly influence parental decisions on tuition payments and child migration. Parents who value their children's education may migrate for higher income to support their schooling financially. Children perform better in school when they reside with their parents, thus parents who prioritize educational outcomes are less likely to leave their children behind. Consequently, parents are less inclined to leave their children behind when they prioritize educational outcomes, potentially reneging on both financial and caregiving aspects of informal contracts Additionally, grandparents gain more utility if they live longer while their grandchildren are in school, with φ and λ affecting this utility increase. Notably, φ specifically enhances utility from high school enrollment, while λ impacts returns at all educational levels. These parameters are identified using data on children's enrollment rates, educational attainment by age 22, migration patterns of families, fulfillment of informal contracts, and grandparents' healthcare expenditures.

Guilt: associated with reneging on financial and care obligations in the informal contract (κ_1 and κ_2 in equation (12)) influences parental behavior towards ill grandparents. High κ_1 values lead parents to increase financial transfers during grandparental illness. High κ_2 values encourage parents to remain in rural areas to provide elder care. These parameters are identified using the proportion of parents who fulfill each contractual component.

Utility gain from healthcare for ill grandparents: (η_2 in equation (6)) impacts their medical consumption and decision-making related to informal contracts. Grandparents prioritizing healthcare are more inclined to invest in medical services.As healthcare concern increases, the perceived value of the informal contract also rises, prompting more grandparents to propose these agreements, leading to more children being left behind. This, in turn, results in increased financial support from parents to grandparents. To identify η_2, I analyze the proportion of grandparents spending on healthcare, their health status over time, and the frequency of receiving financial transfers from parents.

Additive component of utility function coefficient: (δ in equation (3) and (5)) impacts agents' sensitivity to child education, informal contract adherence, and healthcare. Higher δ values correlate with increased household cooperation and greater parental fulfillment of contractual obligations. This coefficient is identified through the proportions of migrants and migrating children, average remittances to grandparents, and rural consumption levels.

Unemployment rate in period 5: (pr^{ump}) affects the expected return of migration during that period, directly affecting migration decisions. A high unemployment rate reduces parent migration in period 5. If the grandparents are ill in period 5, they are more likely to receive care from non-migrating parents, potentially extending their lifespan. This situation indirectly increases parents' urgency to migrate before period 5, leading to a higher likelihood of reneging on the care component of the informal contract.

E.5 Estimation Results of the Extended Model

The baseline model, presented in section 4, characterizes intergenerational behaviors through an exchange of private care and financial transfers between the parent and the grandparent. This model is referred to as the Exchange model. The first extended model is called the Exchange+Altruism model, and the second extended model is called the Altruism model. Section 8 details differences in the setup of these models.

Table B18 below reports the internally estimated parameter values, and Table B19 reports the model moments corresponding to the three sets of parameters. The estimates of the Exchange+Altruism model are very similar to those of the Exchange model. This demonstrates the limited extent to which altruism explains the intrahousehold behaviors that this model aims to capture. The Exchange+Altruism model's fit is substantially poorer than that of the Exchange model. The fractions of migrating parents are higher than the data moments in period 0 (when children are aged 0 to 5) and period 3 (when children are aged 15 to 17), but are lower in other periods. The Exchange+Altruism model overestimates the fraction of left-behind children in periods 0 and 2, as well as the fractions of parents who provide financial transfers and elder care at the same time. This indicates that altruistic grandparents and parents would be predisposed to provide more financial support and services for each other than what is observed in reality.

By setting the guilt parameter g_t to zero, the Altruism model implies that all intrahousehold behaviors are driven by altruistic motives. This model further overestimates the extent of parents' out-migration in periods 0 and 3-4, as well as the number of children left behind. This occurs because both agents gain utility from the child's educational outcomes, a factor that the altruism model particularly emphasizes. As a result, the moments predicted by the Altruism model overshoot the child's school enrollment in middle school and high school, consequently inflating the average years of schooling. Additionally, the model overestimates the proportion of ill grandparents not receiving healthcare, even though most parents who received childcare help from grandparents provide financial transfers. This is also due to the household's high valuation of children's education and, as implied by the neutrality result in altruistic household models (Cox, 1987), a substantial share of household assets is allocated to education rather than to healthcare for grandparents.

Table B18: Internally Estimated Parameters from the Three Models

Description	Symbol	Exchange	Exchange and Altruism	Altruism
Annual discount factor	β	0.8657	0.8661	0.8751
Coef of CRRA for private utility	γ	1.2348	1.2348	2.5575
Consumption-leisure trade-off	θ	0.2789	0.2789	0.2822
Additive factor in utility	z	0.3554	0.3554	2.1315
Grandparent's utility gain from healthcare when ill	η_2	0.3491	0.3491	0.3195
Utility for children being enrolled for high school	φ	0.0245	0.0245	0.0512
Coef of CRRA for children's education	λ	2.6493	2.6943	3.6901
Parent's guilt from low remittance	κ_1	0.0060	0.0061	-
Parent's guilt from not taking care of grandparent	κ_2	0.0501	0.0551	-
Parent's unemployment probability when the child is over age 18	pr^{ump}	0.5970	0.5995	0.5007

Table B19: Compare the Goodness of fit of Three Models to the Data

Name of the Moment			Data Moments	Model Moments		
				Exchange	Exchange and Altruism	Altruism
Migration						
Fraction of migrants by children's age						
Age 0 to 5			0.322	0.261	0.435	0.565
Age 6 to 11			0.243	0.211	0.116	0.116
Age 12 to 14			0.183	0.136	0.186	0.025
Age 15 to 17			0.171	0.206	0.283	0.197
Age 18 to 20			0.149	0.035	0.035	0.234
Informal contract						
Conditioning on grandparents provided childcare in the past, the fraction of parents who:						
Child's age	Parent provides transfer	Parent provides elder care				
Age 0 to 14	Yes	Yes	0.665	0.859	0.860	0.655
Age 0 to 14	Yes	No	0.059	0	0	0
Age 0 to 14	No	Yes	0.222	0.141	0.140	0.346
Age 15 to 20	Yes	Yes	0.647	0.757	0.750	0.769
Age 15 to 20	Yes	No	0.294	0.200	0.107	0.146
Age 15 to 20	No	Yes	0.038	0.043	0.143	0.085
Fraction of ill grandparents left behind by parents and children			0.083	0.021	0	0
Remittance and consumption						
Average annual remittance (RMB/year)			4099	3915	3910	3852
Fraction of grandparents receiving remittance			0.398	0.222	0.361	0.341
Average consumption per rural adult (RMB/year)			3394	3226	3620	3653
Children						
Fraction of left-behind children						
Age 0 to 5			0.257	0.261	0.435	0.565
Age 6 to 11			0.187	0	0	0
Age 12 to 14			0.099	0.136	0.186	0
Fraction of children enrolled in school						
Primary school			0.969	1.000	1.000	1.000
Middle school			0.898	0.871	0.879	0.999
High school			0.567	0.606	0.591	0.953
Average years of edu of children at age 22 (years)			10.130	10.273	10.170	11.577
Grandparent's health						
Fraction of ill grandparent not getting health care			0.548	0.647	0.720	0.786
Fraction of ill grandparent by children's age						
Age 6 to 11			0.169	0.214	0.214	0.215
Age 12 to 14			0.115	0.172	0.178	0.178
Age 15 to 17			0.090	0.101	0.108	0.122

F Computation Appendix

In the Generalized Method of Moments estimation process, the computational tasks are in 2 folds: first, for a specific set of parameters, one needs to calculate the model moments; second, one needs an algorithm to search for the optimal set of parameters to minimize the distance between the model moments and the data moments.

In this appendix, sections F.1 explains my approach to solve the first task. Section F.2 shows that this approach yields a consistent result as the standard backward induction approach as in dynamic programming. In section F.3, I explain my solution to the second task.

F.1 Solving the model

Given a set of parameter values $\vec{\theta} = \{\gamma, \theta, \eta_2, \varphi, \lambda, \kappa_1, \kappa_2, pr^{ump}\}$ and the initial states on household wealth and grandparent's health, $\{A_j, h_0\}$, I describe the algorithm it takes to obtain the set of model moments corresponding to $\vec{\theta}$ as follows.

1. Obtain $\left\{contract^G, \{X_t, choice_t^P, choice_t^G\}_{t=1}^5 | A_j, h_0\right\}$, the set of all feasible paths of choices and states:
   (a) Recall that the choice variables include (1) a pair of migration decisions for the parent and the child (3 choices), (2) a binary education investment decision for the child (2 choices), (3) a binary healthcare decision of the grandparent (2 choices), and (4) a pair of continuous daily consumption decisions of the grandparent's household and the parent's household. I discretize the daily consumption decisions by putting them on a grid. In my estimation specification, I chose grids with 300 RMB/year ($44) spacing, which corresponds to 9 and 15 consumption grids for agents living in the rural and urban areas, respectively.
   (b) With completely discretized choice set, the set is finite and there are up to 3,240 distinct choices per period per household.
   (c) On the other hand, the set that state of the household takes value on is also finite. The state variables include (1) a pair of contract status of the two agents (3 states), (2) a pair of location status of the parent and the child (3 states), (3) the wealth of the two agents (same grids as the consumption, up to 51×41 states), (4) the health status of the grandparent (3 states), (5) the level of guilt of the parent (4 states), (6) the education attainment of the child (5 states), and (7) the education enrollment status of the child (2 states). There are up to 2.3 million distinct states per period per household, although the actual state space would be much smaller as it is limited by many model constraints.
   (d) I build the set of paths by appending the set of feasible choices to a given state, and then the set of possible states to a given choice. The number of distinct paths on the resulting tree diagram ranges between 1.6×10^6 and 3.6×10^{14}, depending on the initial wealth and the health of the grandparent.
2. Compute the probability measure of each path: The probability of a path depends not only on states but also on the history of choices. For each complete profile of states and choices in every period in the model, I am able to compute the conditional probability of a state and time X_t given its past states X_{t-1} and choices $choice_{t-1}$.
3. Identify the optimal path of choices given each state:

(a) I start from period T, at which the value functions of the two agents are deterministic and can be computed from the complete path that leads to this terminal state. Therefore, I obtain V_T^P and V_T^G for each path defined by $\{contract^G, \{X_t, choice_t^P, choice_t^G\}_{t=1}^5\}$.

(b) Using backward induction, following equations (17) - (19), the optimal choice at period t conditioning on X_t can be obtained by comparing the averages of the value functions with respect to the probability measure of the paths.

(c) When the optimal choices of complete paths are found for each profile of states, the probability distribution over these optimal choices are also assigned. Therefore, the set of complete profiles of choice and state variables in all time periods and their corresponding probabilities is the solution of the model, with respect to a specific pair of A_j and h_0.

4. Compute the marginal probabilities for model moment estimation: In my model, the time periods vary in lengths. So conditioning on observing a household, the probability that this household is in period t is $Pr(t) = \frac{length(t)}{20}$. Therefore, I can compute the marginal probability of observing a household in a certain time period, wealth level, initial grandparent's health condition as follow

$$\begin{aligned} &Pr\left\{X_t, choice_t^P, choice_t^G, A_j, h_0\right\} \\ =&Pr\{contract^G, \left\{X_t, choice_t^P, choice_t^G\right\}_{t=1}^5 |A_j, h_0\} \\ &\times Pr(A_j) \times Pr(h_0) \times Pr(t) \end{aligned} \tag{29}$$

The marginal probabilities are used to compute the model moments, and the calculation is straightforward.

F.2 Equivalence between dynamic programming and my approach

The approach I use to solve the model, as described in section F.1, conceptually employs a discretized forward induction algorithm. In this subsection, I demonstrate that this algorithm is consistent with the dynamic programming approach with an example.

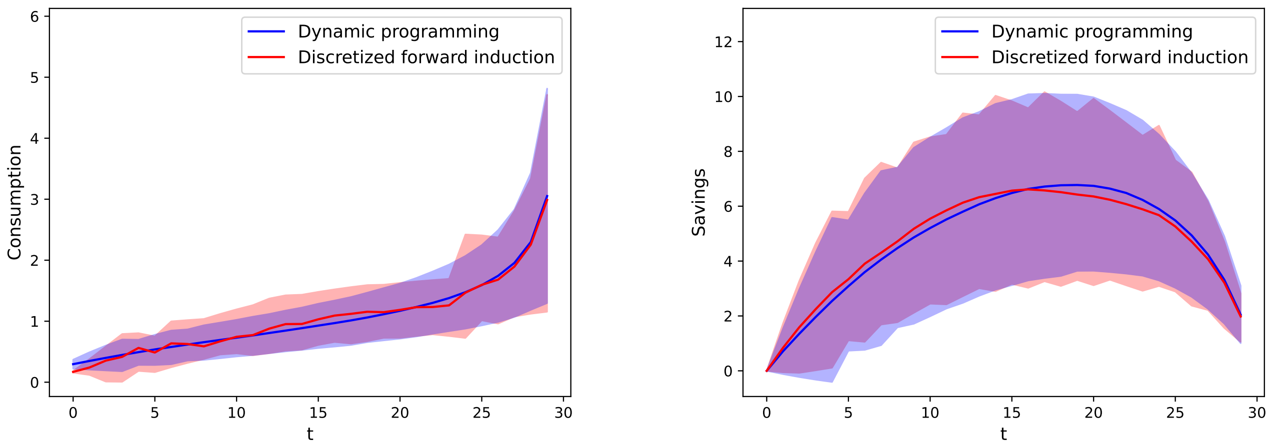
Consider a simple lifecycle model featuring stochastic income and a borrowing constraint.

$$\begin{aligned} &\max_{(c_1,\dots,c_T)} \sum_{t=1}^{T} \beta^{t-1} u(c_t) \\ &\text{s.t.} \quad a_{t+1} = (1+r)(a_t + y_t - c_t) \quad \text{for } t = 1, \dots, T \\ &\qquad\; a_{T+1} \geq 0 \end{aligned} \tag{30}$$

in which the utility function is CRRA, i.e. $u(c_t) = \frac{c^{(1-\gamma)}}{(1-\gamma)}$, and income y_t is uncertain.

Calibration: Consider a model with the number of time periods set at $T = 30$, an interest rate of $r = 0.01$, a discount factor of $\beta = 0.98$, a coefficient of relative risk aversion $\gamma = 10$, an initial asset value $a_0 = 0$, and an uncertain income $y_t = 0.1$ with a probability of 0.5 and $y_t = 0.9$ with a probability of 0.5. In both approaches, I use 50 grid points for the asset a_t for all t, and the grids take equal steps in logs. For the dynamic programming approach, I utilize the Piecewise Cubic Hermite Interpolating Polynomial (*pchip*) interpolation method. The forward induction is fully discretized, and thus does not need interpolation. I simulate 10,000 observations to generate the model moments.

Simulation result: Figures B4a and B4b present a comparison of the model moments estimated using two computational approaches. The lines and shaded areas in blue represent the estimated average and 95% confidence band from the dynamic programming approach, while those in red correspond to the discretized forward induction approach. Although results from the forward induction approach appear more discretized than those from the dynamic programming approach, the distributions of consumption c_t and assets a_t remain remarkably consistent. Practically speaking, to guarantee that both approaches yield consistent outputs, at least 30 grid points should be used for continuous state and choice variables. In the estimation of the model in this paper, I use 50 grids for savings s_t^P and s_t^G.



(a) Distribution of consumption by t

(b) Distribution of assets by t

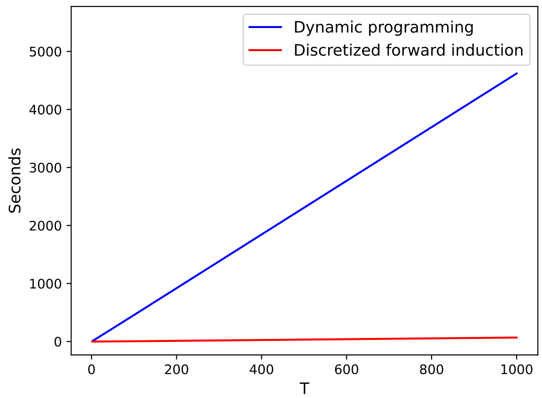
Figure B4: Simulation Results from Dynamic Programming versus Discretized Forward Induction

Note: This figure displays the estimates for the average values and the 95% confidence bands of c_t and a_t. The blue line and shaded area represent the estimates obtained from a standard dynamic programming approach using backward induction. The estimates from the discretized forward induction approach are represented by the red line and shaded area.

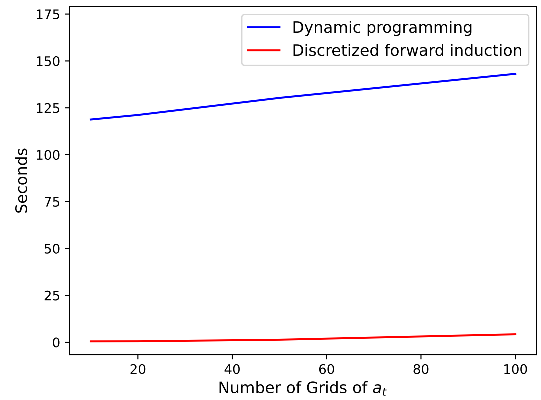
Compare the speed of the approaches: The model developed in this paper involves two agents, a path-dependent informal contract, and multiple sources of uncertainty. To estimate this model efficiently, I employ the forward induction approach, which significantly accelerates the computation compared to the dynamic programming approach. First, discretization saves on the time required for interpolation, albeit at the expense of accuracy. For instance, the model moments and predictions exhibit sudden variations, as demonstrated in Figure A1a. Second, and more crucially, forward induction facilitates parallel computation within a single iteration (i.e., one set of parameter values), in contrast to dynamic programming, which requires sequential computation from $t = T$ backward to $t = 1$. With abundant access to high-performance computers, the time savings can be substantial.[57]

To demonstrate the difference in computational speed, I estimate the lifecycle model defined in equation 30 using both approaches on the same 64-core computer. Figure B5a illustrates how the computational time, measured in seconds, increases as T, the total number of time periods in the model, grows. Figure B5b displays the increase in time costs as the number of grid points for a_t rises, while Figure B5c demonstrates how time costs rise as the number of distinct values for uncertain income, y_t, increases. Across all scenarios of increased model complexity, the dynamic programming approach consistently requires significantly more time than the discretized forward induction approach, with time costs escalating more rapidly as complexity intensifies.

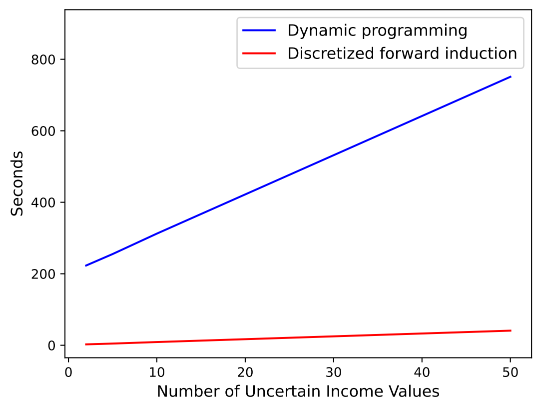
[57]The parallel computation method is explained in section F.3.



(a) Time Cost by the Total of Time Periods (T)



(b) Time Cost by the Number of Grids for a_t



(c) Time Cost by the Number of Distinct Values for Uncertain Income (y_t)

Figure B5: Time Cost Comparison by Computational Approaches

Note: This figure illustrates the time costs associated with each computational approach and model specification, expressed in seconds. The blue line represents the estimates obtained from a standard dynamic programming approach using backward induction. The estimates from the discretized forward induction approach are represented by the red line.

F.3 Computation algorithm

F.3.1 Genetic Algorithm

The Genetic Algorithm belongs to the family of Evolutionary Algorithms. The methodology is inspired by the process of natural selection. And it takes the concept of selection, crossover and mutation from biological evolution.

The basic idea of the algorithm is to test the performance of a set of model specifications (e.g. parameter values), called a *generation*, select the ones with better performance, and use them to create the next generation. When generating new model specifications from a relatively small set of "surviving" models, the algorithm allows for pairwise crossover and mutation at single parameter level. As the algorithm loops over generations, the models with better performance will be kept in the population, and those with worse performance will be left out of selection. The exchange and mutation of the genes in the models with good performances induces improvements upon the existing models. As the genes in the surviving population evolve over generations, the best performing model will be the only model specification left in the process. In the end, this intelligent grid search algorithm converges when the population is homogeneous.

The terminology, parameters and functions used in the environment of Genetic Algorithm are defined as follows:

- Individual $\vec{v_i}$: the set of parameters that defines one specific model. In the context of my paper, it refers to the set of parameters to be structurally estimated.
- Population P: the set of individuals that the algorithm tests in one iteration.
- Survivor: an individual that is included in the population of the next generation and is used to generate individuals that are created in the next generation.
- Set of survivors S: the set of all survivors from a generation.
- Population size (N): The number of individuals in a population. So $|P| = N$.
- Survival rate (p_s): The fraction of survivors in a generation. So $|S| = N \times p_s$
- Fitness function: the objective function that is used to judge the performance of various model. The function value of individual $\vec{v_i}$, denoted by $score(\vec{v_i})$, is defined in equation (31).

$$score(\vec{v_i}) = [(Q_0 - Q(\vec{v_i}))' W^{-1} (Q_0 - Q(\vec{v_i}))]^{-1} \tag{31}$$

  It is the reciprocal of the objective function in equation (21). Therefore, a larger fitness value indicates better model fit.

The Genetic algorithm I use in the structural estimation takes the following steps:

1. Initialize the first generation: randomly choose parameter values from their reasonable domains until a set of N individuals is obtained. Denote this population by P_0.
2. Collect the fitness function values for all of the individuals in this generations, i.e. $\{\vec{v_i}, score(\vec{v_i})\}_{i=1}^{N}$.
3. The set of survivors S_{k+1} from the one generation P_k to the next generation P_{k+1} is constructed based on two selection rules:
   (a) Initialize $S_{k+1} = \emptyset$.
   (b) Elitism Selection: sort the individuals by scores, select the best $N \times \frac{p_s}{2}$ individuals, and appended them to S_{k+1}.
   (c) Roulette Wheel Selection: Let $P_{k,nonelite}$ denote the set of the remaining $N \times (1 - \frac{p_s}{2})$ individuals of the current population. Another $N \times \frac{p_s}{2}$ individuals are randomly drawn from $P_{k,nonelite}$. The probability that a particular individual $\vec{v_i}$ is chosen is

$$pr(\vec{v_i} \in S_{k+1}) = \frac{score(\vec{v_i})}{\sum_{\vec{v_i} \in P_{k,nonelite}} score} \tag{32}$$

   The survivors are added to S_{k+1} until $|S_{k+1}| = N \times p_s$
4. Create the next generation P_{k+1} from the survivors S_{k+1}.
   (a) Initialize $P_{k+1} = S_{k+1}$.
   (b) While $|P_{k+1}| < N$:

i. Randomly choose (with replacement) two individuals from S_{k+1}, denote them by $\vec{v_1}$ and $\vec{v_2}$. They will generate two new individuals $\vec{v_1}'$ and $\vec{v_2}'$ after crossover and mutation.

ii. Crossover on the two individuals. For each component in the two individuals, they may exchange the value of that component with probability $pr_{crossover}$. Specifically, the crossover affects component j of the new individuals in the following way:

$$\begin{cases} \vec{v_1}'_j = \vec{v_2}_j \text{ and } \vec{v_2}'_j = \vec{v_1}_j & \text{with } \; pr_{crossover} \\ \vec{v_1}'_j = \vec{v_1}_j \text{ and } \vec{v_2}'_j = \vec{v_2}_j & \text{with } \; 1 - pr_{crossover} \end{cases} \tag{33}$$

The crossover happens on all the components of the individuals independently.

iii. Mutation on each of the two individuals. For each component in $\vec{v_i}'$, it may be affected by a random shock with probability $pr_{mutation}$. Specifically, the mutation affects component j of the individual in the following way:

$$\begin{cases} \vec{v_1}''_j \sim N(\vec{v_1}'_j, \sigma_j^2) & \text{with } \; pr_{mutation} \\ \vec{v_1}''_j = \vec{v_1}'_j & \text{with } \; 1 - pr_{mutation} \end{cases} \tag{34}$$

where σ_j^2 is the exogenously chosen mutation variance of component j. The mutation happens on all the components of the individuals independently.

iv. The resulting new individuals $\vec{v_1}''$ and $\vec{v_2}''$ are appended into $|P_{k+1}|$.

v. Repeat steps 4-b-i to 4-b-iv.

5. Repeat step 2 to 4 until the algorithm converges. Convergence is defined as

$$\max \{score(\vec{v_i})\}_{\vec{v_i} \in P_k} - \min \{score(\vec{v_i})\}_{\vec{v_i} \in P_k} < c \tag{35}$$

where c is a constant tolerance parameter.

Specifically, in my algorithm, I set $N = 100$, $p_s = 0.3$, $pr_{crossover} = 0.5$ and $pr_{mutation} = 0.6$.

F.3.2 Distributed System

In addition to the parallel programming enabled by the Genetic Algorithm, I decompose one iteration of my model further to allocate the computing tasks at a finer scale.

The complication of the model structure leads to long running time, so I build a distributed system that makes full use of multiple high-performance computers at the same time. Although the system design faced some practical challenges, the resulting system accelerated the estimation procedure tremendously, and allowed me to estimate my model without making further simplifications.

Decompose one iteration: The initial condition of the households in my model differ by the household type and the grandparent's health condition. I allow for 10 household types. And the health condition has three categories, healthy, ill or dead. Since only healthy grandparents have the opportunity to enter the informal contract, their initial choice on the contract divide their paths of choices and states into two subsets without overlap. In this way, a single iteration, i.e. the computational task for evaluating one set of parameters, may be divided into 40 sub-tasks.

These sub-tasks are computationally independent, in the way that the cache from one sub-task is not useful for another sub-task. Therefore, it is feasible and efficient to let the 40 sub-tasks run concurrently. And the computational task of my structural estimation is organized and distributed by setting sub-task as the smallest unit of task assignment.

Distributed system: In the Genetic Algorithm, each generation has 70 new individuals (= population size - number of survivors from the last period). When the new parameter values are generated, a set of 2800 sub-tasks are created at once. Processing them sequentially takes a long time, so I consider a distributed system that allows my program to run on multiple multi-core high performance computers (HPC) at the same time. However, in practice, I have access to 10 72-cores HPCs, which means the number of sub-tasks is much larger than the number of CPUs in the system. An allocation problem arises when I assign the sub-tasks to computers.

I use a 1-core machine (the "leader") to control 10 HPCs (the "workers"), each with 72 cores. The leader is responsible for operating the Genetic Algorithm, and divide the computation task of the new generation into sub-tasks. The workers with spared CPUs requests new sub-tasks from the leader, and assign each sub-task to a single core.

Practical challenges: Concurrent computation faces two major types of practical challenges. First, the system may result in context switching overhead problem, when the CPUs on a HPC are overflowed by the amount of tasks allocated to this machine. The problem refers to the scenario in which, because it is energy and time-consuming for a CPU to switch between computing tasks, the management and storage of the task information negatively affect the operating system and application performance (Silberschatz et al., 2012). To increase efficiency of the program, the objective is to match one-on-one the number of tasks with the number of CPUs whenever possible. In my program, each HPC controls the size of the sub-tasks they are capable of processing given the number of available CPUs to alleviate the switching overhead problem.

Second, large sub-tasks that run at the end of a section of concurrently ran sub-tasks may cause long tail of latency. The problem refers to the scenario when a set of sub-tasks have to be all completed for the program to proceed, while the last few large unfinished sub-tasks of the set may keep most of the CPUs unoccupied, and thus causes a waste of time and energy. In my model, the choice set is much larger for households who adopt the contract than those who do not, and also for wealthier households. So each iteration has a few large sub-tasks (usually $5 \sim 10$), while the rest of the sub-tasks are much less computationally demanding. Therefore, the workers in my program request large sub-tasks first to reduce the length of the length of the latency at the end of each section.

Performance of the system: Using the computation system described as above, one iteration of the Genetic Algorithm (70 iterations of the model) takes about 35 minutes to run. While its running, the CPU usage of the system in a consecutive 3 hours is presented in B6. It shows that the distributed system with careful allocation design (1) achieves full usage of all the 720 CPUs most of the time and (2) eliminates most of the context switching overhead problems.

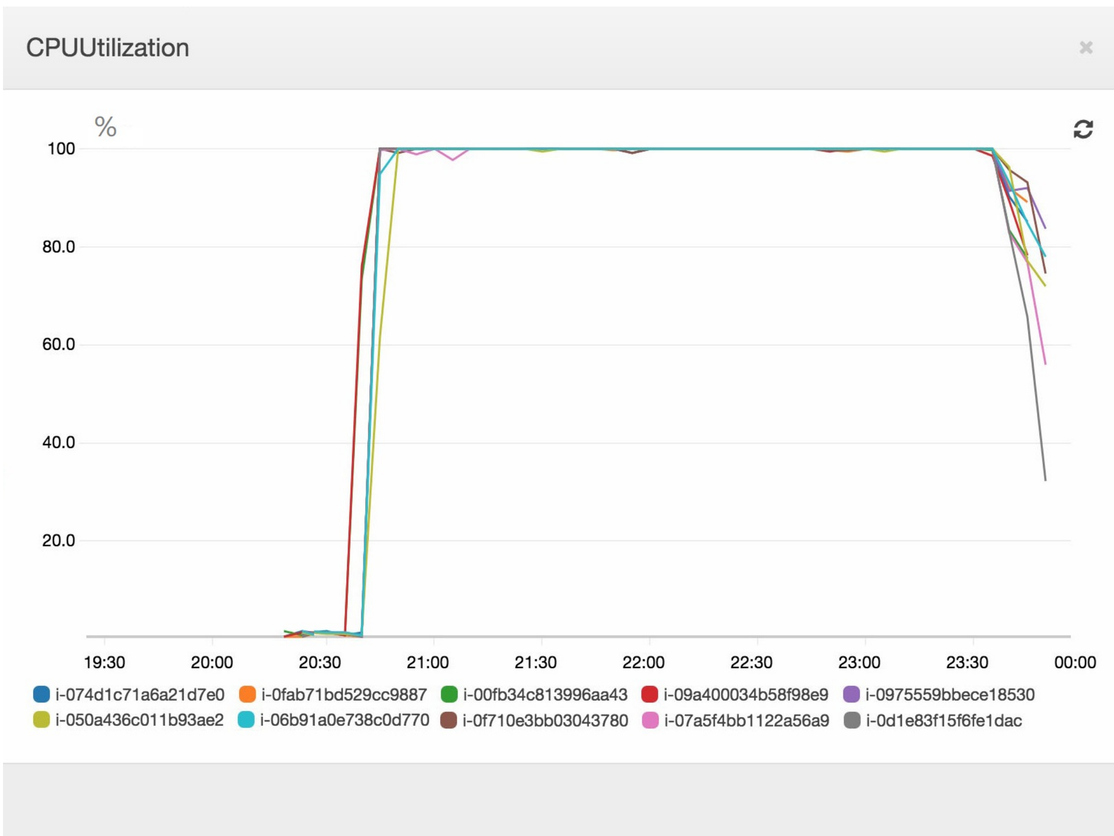


Figure B6: CPU usage of 10 72-core high performance computers controlled by a one-core machine on the Amazon Web Services