Internal migration and extended families in China

Ling Zhong*

November 26, 2018

Click here for the latest version

Abstract

In this paper, I study how rural-to-urban migration in China affects households' inter-generational behavior, and the effects of policies targeting migrant households on their welfare. Internal migration from rural to urban areas can have significant welfare effects on migrants and their extended families. In China, migration is often temporary, and most family members of migrant workers are left behind. In these households, many left-behind grandparents look after the children of migrating parents. However, the behavioral and welfare effects of government policies directed towards rural households with potential migrants remain unknown.

Using five Chinese data sets on the migration patterns, education choices, financial transfers, and health of multi-generational families, I first present a rich set of stylized facts about migration and household behavior. The evidence shows that in many rural households, parents migrate to urban areas for work when healthy grandparents are able to provide childcare. When the grandparents are sick, migrating parents return to the rural area to provide elder care and pay for their parents' healthcare. With the facts as a guide, I develop and estimate a structural model of the behavior of migrants and their families. The model features an informal limited-commitment contract over child care, financial transfers, and elder care. Parents and grandparents play a sequential game by choosing migration status, informal contract status, remittances, children's education, and grandparents' healthcare. The estimates suggest that poor households adopt the informal contract so that rural consumption, education, and healthcare are funded by the migrants' remittance.

I then use the model to evaluate the effects of a set of hypothetical government policies. An urban education subsidy promotes children's education, increases the migrants' consumption in the urban area, and does not affect the grandparents' welfare. But it does not alleviate the problem of children left behind as the government had hoped. An improved insurance coverage that lowers out-of-pocket healthcare costs would reduce the grandparents' demand for the informal contract. It would generate a welfare gain to the grandparents, discourage parents' migration, and increase children's education. The policy counterfactual outcomes imply that policies intended to improve the welfare of one family member can affect the welfare, consumption behavior and migration decisions of all three generations through intra-household cooperation. The design of these policies should account for intra-household responses.

*Department of Economics, Yale University, New Haven, CT, 06511, USA. E-mail: ling.zhong@yale.edu. I thank Joseph Altonji, Emily Oster and Cormac O'Dea for their invaluable support and guidance. Thanks also to Elyse Adamic, Isaiah Andrews, Christian Dustmann, Ran Gu, John Eric Humphries, Louise Laage, Costas Meghir, Yusuke Narita, Jesse Shapiro, Joseph Shapiro, James Thomas, Jaya Wen, Xiaoyi Zhang, Qingyou Zhao and seminar participants at Brown University and Yale University for their helpful comments. All remaining errors are my own.

1 Introduction

Internal migration, especially rural-to-urban migration, is an important economic phenomenon in many countries (Lucas, 1997; Greenwood, 1997). In China, the focus of this paper, millions of rural people migrate to major cities in search of higher paying jobs. Migrant workers coming from rural areas account for as much as 41% of the urban labor force[1]. However, migration is often temporary. Additionally, because the cost of living in urban areas is significantly higher than in rural areas, migrants often choose to leave their family members behind. A substantial share of this labor supply is lost in the return migration movement for those over 35 years of age, and the frequent turnover in employment caused by temporary migration results in a loss of productivity.

The current situation forces rural households into adverse circumstances. They benefit from the migrants' increased remittance to fund health care and education, but suffer from a lack of private care due to the family's physical separation. Parents do not wish to be separated from their children, but they cannot afford the high education and living cost for children in the urban areas, and they do not have time to look after the children[2]. Many left-behind children also suffer from poorer educational, psychological, and nutritional conditions than those who live with their parents, resulting in a lower quality of life in their most formative years[3]. Grandparents have better living conditions due to the migrants' remittances, but they report a lower level of life satisfaction as migrating parents are less likely to fulfill their elder care responsibility (Luo, 2009; Zhuo and Liang, 2015). Ultimately, while these circumstances provide households with greater financial resources, their impact can be felt across multiple generations of a household.

Families with migrants attract much attention from policy-makers and researchers, because they account for 40% of all Chinese rural households, and migration induces a welfare trade-off for left-behind family members. The intra-household welfare distribution of migrants' households is inefficient from the government's perspective, because the Chinese government is particularly concerned with the generational welfare outlook for children and the elderly[4].

The Chinese government wants to improve the welfare of children and the elderly who are left-behind while maintaining some control over the size of migration flows. However, recently implemented policies addressing one generation of the household have either been ineffective or have unintended externalities.

[1] China Family Panel Survey, 2010-2014

[2] China Women's Federation Children's Work Department. "Research report on left behind and migrant children" (2011).

[3] Zhuang, Pinghui. "Revisiting Chinas village of left-behind children after suicides that shocked nation". South China Morning Post (2018).

[4] Chinese State Council. "Report on strengthening rural left-behind children care and protection work." (2016).

For example, in order to alleviate the problem of left-behind children, in 2017 the Chinese government began closing the lower quality migrants' schools and opening public schools to migrants' children in the urban area. Nevertheless, not many parents migrate with their children due to the high living and education costs[5].

In this paper, I address three main questions. First, how do extended family characteristics (for example, the wealth of the household, the enrollment status of the children, and the health condition of the grandparents) affect the rural parent's migration decisions? Second, what are the welfare gains from intra-household cooperation? And third, what policies could improve the welfare of the grandparent, parent and child in the presence of migration and the intra-household cooperation? To do so, I develop and estimate a structural model of rural Chinese households. The model features an informal contract that the parents and the grandparents may enter to enforce a set of financial and daily care interactions. The estimated model enables a deeper understanding of the interplay between the rural parents' migration behavior and the welfare of their children and grandparents. It also allows for counterfactual experiments that illustrate the rural households' reactions to various policies.

To guide my analysis, I begin by documenting a set of facts concerning migration and the migrants' extended families. I analyze five different nationally representative micro-level datasets on rural China. My analyses suggest that rural and urban areas have very different income levels, labor supply, living expenses, children's education costs, and access to healthcare. A parent's migration behavior depends on the education of the child and the health status of the grandparent in his household, in addition to the economic factors. Children left behind by migrating parents have worse educational outcomes because their guardians are less educated. Grandparents left behind benefit from remittances from the migrants. The amount of the remittance depends on the grandparent's decision of whether or not to care for the migrant's child.

These facts guide the development of my structural model. I assume that parents and grandparents play a dynamic sequential game involving migration, care, consumption, and transfer decisions. The parent may migrate with or without his child while providing remittance and private care to the grandparent. He values private consumption and his child's educational outcomes while considering whether to take on the responsibility of elder care. The grandparent chooses whether she is willing to take care of the child if the parent migrates. She values consumption, health, and the grandchild's education. The child is assumed to be passive. Her educational outcome depends on whether her parents pay for school

[5] Liang, Lim Yan. "China's migrant children dilemma". The Straits Times (2018).

and who cares for her on a day-to-day basis. The model incorporates an informal contract stating that if the grandparent takes care of the left-behind grandchild, the migrating parent pays for the grandparent's medical treatment and will return to the rural area if and when the grandparent is sick. The contract is one of limited commitment for the parent, who is subject to a disutility if he fails to fulfill the contract requirements.

Households are heterogeneous in their productivity in the rural area. They face two sources of idiosyncratic risks: First, children may drop out of school. The dropout rate depends on the education level of the child and her guardian. Second, the grandparent faces illness and mortality risks. Her health transition probabilities depend on her own age and healthcare consumption decision, in addition to the parent's migration decision when the grandparent is sick.

I estimate the parent's and grandparent's preference parameters in the model using the Generalized Method of Moments. I use a broad set of moments generated using my five datasets on migration movements, commitment towards the informal contract, remittance and consumption amounts, children's education outcomes, and the grandparents' health status and healthcare.

The model estimates show that both the parent and grandparent care about the child's education. When the grandparent is sick, the parent feels guilty if he does not stay in the rural area to take care of the grandparent, or if he does not pay for the grandparent's spending on commodities and healthcare. The model indicates that the informal contract involves a trade-off for every generation in the family. The parent benefits from not having to pay for the child's spending in the expensive urban area. But if he fulfills the contract, he suffers the financial cost of paying for healthcare and a time cost of providing elder care. The grandparent gains from additional income from the migrant's remittance, but she sacrifices leisure time for childcare, and suffers from the risk of being abandoned by the parent when she is in poor health. The child's education is more likely to be paid for if her parent migrates, but she is also more likely to drop out of school because of poor school performances. My model estimates suggest that poor households adopt the informal contract. Grandparents propose the contract to parents when it enables them to afford healthcare. Parents accept the contract when they want to migrate. In households with moderate flow income levels, parents can obtain welfare gains from the informal contract, while grandparents cannot. Wealthy rural families do not have an incentive to migrate because, for them, the rural-urban income gap is small while urban prices are high.

I use the estimated model to evaluate several counterfactual policies toward migrants and their extended families. The first policy I test is an education subsidy that lowers the tuition for migrant's children in the urban area. While such a subsidy has been informally implemented by the local government in some cities[6], the effects have not been evaluated empirically. My experiment shows that the policy increases the education attainment of rural children who migrate, while the children's migration behavior does not change. Parents gain from higher urban consumption levels and their children's increased education attainment. In addition, grandparents are not affected by the policy, because the disutility of reneging on the informal contract motivates the parents to stay in the rural areas when grandparents are sick.

Chinese policy makers have also considered an expanded health insurance policy that increases the government's share in the co-pay system[7]. My model indicates that this policy generates a welfare trade-off between grandparents and parents. Grandparents receive a welfare gain from the policy, as more of them are able to afford healthcare. The lower price of healthcare makes grandparents less concerned about their budget, and reduces the pecuniary value of the informal contract to the grandparents. Some grandparents opt out of the informal contract and thus gain utility from not bearing transfer and private care uncertainty. The policy affects parents' welfare through the contract. As fewer grandparents are willing to provide childcare, more parents must migrate with their children. As a result, the fraction of migrants decreases by 5%. Migrating parents lower their consumption level to afford commodities, housing, and education of their children. More children live with their parents and benefit from having higher-educated guardians.

My paper is the first to formalize a three-generational household model with migration. It contributes to three literatures.

It contributes primarily to the literature on structural models on migration. My paper reveals many intra-household effects that only a three-generational model could capture. Among existing works on migration, the structural models focus either on households with one or two generations (Barham and Boucher, 1998; Thom, 2010; Gemici, 2011; Görlach, 2016; Morten, 2016), or on the migrants' labor market (Bayer and Juessen, 2012; Yoon, 2017). Other theoretical and reduced form empirical works on households with migrants in developing countries include, for example, Todaro (1969), Hoddinott (1994) and Guriev and Vakulenko (2015).

My paper is also related to the literature on household models with inter-generational behaviors My contribution incorporates the interplay among all three generations and highlights the migration decisions. Previous works on the elderly and their adult children focus on the financial transfers and private care (Becker, 1974; Altonji et al., 1997; Costa, 1999; Pezzin et al., 2007; Mommaerts, 2015;

[6] Chi, Ma. "China pushing for education equality". China Daily (2017).
[7] "Reform and Innovation for Better Rural Health Services in China". The World Bank (2015).

Wiemers et al., 2017). The literature on the relationship between parents and their young children concerns the consumption smoothing of the children (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). I incorporate some of the key intergenerational features of the existing models, and add the migration opportunity and a third generation.

Finally, my paper also relates to the literature on internal migration in China. One series of the papers studies the migrants' labor market and living condition in cities (Zhao, 2003; Song et al., 2008; Song and Zenou, 2012). Significantly more research on internal migration concerns migration behavior (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 migrant's household (Brauw et al., 1999; Zhao, 1999; Zhu, 2002; Biao, 2007; Mu and Van de Walle, 2009; Qin and Albin, 2010; Chang et al., 2011). My paper builds on the empirical findings of my own analyses and the prevailing literature to construct a structural model on the Chinese rural-to-urban migrants and their extended families. A three-generational household model with a sequential game is a complicated system. The structural model is a natural tool to describe and analyze the mechanism of the family. Furthermore, it provides a way to evaluate counterfactual government policies.

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 of two active decision makers and a passive child, with a migration option and an informal contract option. Section 5 discusses the estimation procedure and the model estimates. Section 7 uses the model estimates to evaluate the role played by the informal contract and the effects of an education policy and a healthcare policy on the rural household and rural-to-urban migration. Section 8 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 outside the model. 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. I also provide additional details on the data processing in Appendix C.

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 and 2014. 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, which allows me to identify rural households with migrants. Second, the CFPS captures various household arrangements including childcare and remittance. Third, it documents family sizes and agricultural income, which allows me to recover the income distribution in rural China.

Rural-Urban Migration in China (RUMIC) The RUMIC data targets the internal migrants with a survey specifically designed for rural-to-urban migrants. It was intended to be a longitudinal study starting from 2008. As following the migrants proved difficult in practice, however, only two waves were released. Thus, I treat it as a cross-sectional data. Among the three subsamples (rural residents, urban residents, rural migrants in urban areas) and over two years, I use the 2009 urban migrant survey the most. It contains 5,426 individuals over 15 years old. The data provides information on migration movements, migrants' labor market condition, remittance and left-behind children.

China Health and Nutrition Survey (CHNS) One of the longest-running panel 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 the time allocation information in the data 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 information about the financial transfers and daily care from the elderly to their grandchildren and from the adult children to the elderly. I take advantage of this strength of the data to find correlations among various inter-generational 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 inter-generational 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 an 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 1. 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.

Table 1: 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 sick 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

3 Facts about migration and extended families

In this section I highlight facts about rural-to-urban migration in China (Section 3.1), Chinese rural households (Section 3.2), and how these rural households behave under the massive migration movement (Section 3.3). The facts are based on empirical work 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 inter-generational behaviors in rural Chinese households with migrants.

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 from the mid-1990s (Zhao, 2005). Currently, this policy change is used as a means to balance the excess rural labor supply, caused by increased agricultural productivity, with the excess urban labor demand, driven by the thriving manufacturing industry. However, the migration had to be temporary[8]. 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. Table 2 presents a list of location-specific economic factors that are key to the migrants. In choosing between rural and urban settings, a potential migrant is implicitly deciding between two very different sets of jobs and expenses.

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).

[8] 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.

Table 2: 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.

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

Expenditure concentrates on food, healthcare and education. Table 11 in Appendix D disaggregates the annual expenditure of rural households by spending category. The table 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 sick. The 2005 Chinese Longitudinal Healthy Longevity Survey reports that 63% of sick grandparents are taken care of by their adult children[9]. The 2015 National 1% Sampling Survey reports that 46% of sick grandparents financially depend on their children.

Children who are taken care of by grandparents have poorer educational outcomes. Table 3 reports the regression output on children's school enrollment status and course performances. These significantly depend on the years of schooling completed by the child's primary caretaker. The relationship between the children and their primary caretaker, however, does not have a significant effect. Therefore, children have a worse educational outcome if they are taken care of by their grandparents because rural

[9] 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 "sickness" 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, "sickness" means not able to provide childcare, and requiring private care in the grandparent's daily life.

grandparents are less educated than the children's parents[10] (Appendix D).

Table 3: 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.

3.3 Rural households with migrants

The opportunity for migration allows rural parents to earn more, and the rural household structure extends the effect of the parent's migration decision onto the welfare of the grandparents and children in the family. 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 migrant's wealth and demographics, children's education and grandparent's 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 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). Third, when grandparents are sick, 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 D, 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 key age variable 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 depends on migration, childcare, and

[10] 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.

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 parent's migration, grandparents' childcare, and health status. First, the financial transfer between rural parents and grandparents are 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 15 in Appendix D). Parents with migration experience are more generous, even after they return to the rural area. Parents' financial transfer behavior is also positively affected by the grandparents' childcare experience and sickness.

Left-behind children suffer from being apart from 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 three most popular reasons are high cost of living, lack of childcare, and high tuition (Table 16 in Appendix D). 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 remains unclear. My paper models two main effects of migration on the 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 correlation between parent's migration behavior and grandparents' childcare behavior 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 children if the parent chooses to migrate to the urban area while leaving the children behind. It also obliges the migrating parent to send remittance to the grandparent and come back to the rural area when the grandparent is sick, 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 the 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[11]. Thus, migrants have no escape from the consequences of their reputation in their home villages.

3.5 Summary of facts on migration and rural households

The facts I document in this section aim to align the key features of the household structure of the migrants and the social and economic environments they face. I highlight 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:
    - Pecuniary: the parent with migration experience provides financial support to the left-behind extended family;
    - Non-pecuniary: the grandparent trades with the parent by providing childcare in exchange for daily care when she is sick.

4 Model

The model I propose in this section has the following features: First, it describes the sequential game played between a parent and grandparent of a typical three-generational rural Chinese household while the

[11] 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 D.4.

children are growing up. Second, the model emphasizes the interactions between parent and grandparent in terms of monetary transfers and exchange of private care. It further organizes these inter-generational behaviors into an informal contract with limited commitment.

Each household consists of an active grandparent, an active parent, and a passive child. The agents in the household make choices in turn. Before the model's start time, the grandparent decides whether or not to offer the informal contract. In every time period after a child's birth, the parent moves first to choose migration status, informal contract status, financial transfer to the grandparent, consumption, and children's education. Then the grandparent chooses daily consumption and spending on healthcare.

The parent maximizes his intertemporal utility by making migration, consumption, child's education, and remittance decisions. His utilities depend on (a) his own daily consumption and leisure, (b) his child's educational outcomes, and (c) his commitment to the informal contract. The grandparent maximizes her intertemporal utility by making home production (i.e. whether or not she is willing to provide childcare), daily consumption, and medical consumption decisions. Her utilities depend on (a) her own consumption and leisure, (b) her grandchild's educational outcomes, and (c) her health status and spending on healthcare. The household faces the grandparent's illness and mortality risks as well as the child's school drop out risk. Households differ by agricultural endowment, represented by heterogeneous rural income per capita. Every healthy adult who lives in a rural area earns this income. The government's policy instruments includes public health insurance for the grandparents, rural and urban tuition for rural children, and the social security system for people under the poverty line.

Section 4.1 discusses the household structure, and the three generations separately. Section 4.2 sets up the household's economic and social environment. Section 4.3 explains the inter-generational behaviors and the informal contract in the household. Section 4.4 describes the government policies affecting rural households with migrants. Section 4.5 assembles all aspects of the model, so as to define the optimization problems of the grandparents and parents.

4.1 Household structure, utility and budget constraints

Demographics Time is discrete and the model starts at the birth of the child. Time is denoted by the child's age and is divided into five periods by the child's potential education level. The periods vary in length. Period 1 and 2 each cover 6 years and represent the child's pre-school and primary school-aged periods. Period 3 and 4 each cover 3 years and represent the child's middle school and high school-aged periods. Period 5 is 4 years long and corresponds to the age interval when the child might be in college, which is the last period when she may financially depend on her rural family.

The fundamental unit in the model is a household consisting of three people: a grandparent, a parent and a child. Households are categorized into J types, indexed by j. Households of the same type have a homogeneous level of agricultural income per capita per year, A_j[12]. Household types are determined before the start of the model and they 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:

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

- $\left(contract^G, contract_t^P\right) \in \{(0,0), (1,0), (1,1)\}$ 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.
- $h_t \in \{healthy = 0, sick = 1, death = 2\}$ 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 the education attained by the child, represented by years of schooling completed by the end of the current period.

Child The children are passive in this model. Their consumptions are determined by their parents. The child's role in the household is simplified into her education production function. The child consumes two goods: education at price $tuition_t$ and other goods with a total consumption of c_t^C. When the child lives with the parent, the parent is her guardian. When she is left behind, the grandparent is her guardian. The parent pays tuition as the necessary condition for the child to be enrolled in school. The guardian pays for her consumption of other goods.

[12] 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 D. 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 decreases in the last several decades due to the One Child Policy and internal migration.

Exiting education is irreversible. The child must be enrolled to have a chance of completing the current level of education. Enrolled children may drop out with probability $p_t^{dropout}$. Therefore, the two education-related variables, namely $(enroll_t, edu_t)$, evolve from $(enroll_{t-1}, edu_{t-1})$ in the following manner:

- If the child was not enrolled in the last period, i.e. $enroll_{t-1} = 0$, then
  - she cannot be enrolled in school in the current period: $enroll_t = 0$, and
  - her education attainment remains the same as the last the period: $edu_t = edu_{t-1}$.
- If the child was enrolled in the last period, i.e. $enroll_{t-1} = 1$, then in the current period, her access to school depends on whether the parent is willing to pay for tuition.
  - If the parent does not pay for the tuition ($tuition_t = 0$), then the child's education is terminated from this period, with $(enroll_t, edu_t) = (0, edu_{t-1})$.
  - If the parent pays for the tuition ($tuition_t > 0$), then the child is enrolled in school: $enroll_t = 1$. In this case, the child faces a positive drop out probability.
    * With probability $p_t^{dropout}$, the child does not complete the t^{th} level of education: $edu_t = edu_{t-1}$.
    * With probability $1 - p_t^{dropout}$, the child completes the t^{th} level of education: $edu_t = edu_{t-1} + 1$.

The formulation of the child's education production function reflects several effects of the parent's decision on the child's educational outcomes. First, the model requires continuous enrollment to receive more education, rendering the parent's education investment an intertemporal decision. Second, setting tuition payment as a distinct consumption choice echoes the empirical findings that education expenditure is a major consumption category in rural households. Finally, I specify the drop out probability to depend on the child's level of education, and on the guardian's educational background.

The child's consumption of other goods is a portion of the household's expenditure on housing, food and commodities. Instead of allowing the child's consumption level to be an independent choice variable, I set the fraction of the child's daily consumption in the household-level total consumption to a constant as specified in the OECD equivalence scale for China (Liu and Li, 2011). Note that the child belongs to her guardian's household, and thus the guardian's choice of his or her own consumption determines the child's consumption. Therefore, the guardian spends

$$c_t^{\text{child}} = \alpha \times c_t^{\text{guardian}} \tag{2}$$

on the child's daily consumption, where α is the equivalence scale parameter.

Parent's preferences and budget constraint The parent has time-separable utility preferences. His flow utility depends on his private utility and two factors representing his altruism towards the child and the grandparent. He faces an isoelastic private utility function with two inputs, consumption c_t^P and leisure l_t^P, and two parameters, the coefficient of relative risk aversion (CRRA) γ and the consumption-leisure tradeoff θ. The parent gains φ of his private utility if the child is enrolled in high school[13], and loses g_t proportion of his private utility if he reneges on the informal contract. The altruistic terms have two effects on the parent's utility. First, they are amplifiers on the parent's private utility, so they multiply the parent's utility gain from consumption and leisure. Second, they cause another fixed amount of utility gain/loss, so I add them to the term involving parent's private utility with a factor δ. The addition component represents the non-substitutability of components other than daily consumption and leisure. When $\delta = 0$, parents' utility gain from children's education and loss from contract breach are proportional to their private utility. To account for the much higher price level in the urban area, consumption is normalized by the price ratio between the urban and rural area before it enters the parent's utility function. i.e. $\tilde{c}_t^P = c_t^P \times \frac{p^{rural}}{p^{urban}}$ if $urban_t^P = 1$. The flow utility function of the parent is defined as follow[14]:

$$\begin{aligned} U_t^P &= u^P\left(\ \tilde{c}_t^P, l_t^P, enroll_t, g_t\ \right) \\ &= \underbrace{\frac{\left(\tilde{c^P}^{\ \theta}\ {l^P}^{\ 1-\theta}\right)^{1-\gamma} - 1}{1-\gamma}}_{\text{(private utility)}} \times \underbrace{\left(1 + \varphi \times \mathbb{1}_{enroll_t=1} \times \mathbb{1}_{t=4}\right)}_{\text{(child's education)}} \ \times \ \underbrace{(1 - g_t)}_{\text{(contract with grandparent)}} \\ &\quad + \delta \times \left(1 + \varphi \times \mathbb{1}_{enroll_t=1} \times \mathbb{1}_{t=4}\right) \times (1 - g_t) \end{aligned} \tag{3}$$

in which c^P is the individual level consumption of the parent.

The parent earns either a constant flow income A_j from agricultural production if he stays in the rural area or a constant wage w_t from the urban labor market if he migrates to the urban area. He also receives a bequest B from the grandparent right after the grandparent's death. Note that, as a way to

[13] 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 costly and non-compulsory education period should result in a welfare gain.

[14] I define the terminal utility for the parent later in equation (17).

incentivize contract commitment, the parent does not receive any bequest if he reneges on the contract, or if he does not enter the contract. The income is spent on daily consumption of himself c_t^P and his co-residing child $c_t^P \times \alpha \times \mathbb{1}_{urban_t^P=urban_t^C}$. It is also spent on the child's education $tuition_t$ as well as the transfer Tr_t to the grandparent. The wealth of the parent is denoted by s_t^P. The intertemporal budget constraint of the parent is

$$
\begin{aligned}
s_{t+1}^P = (1+r) \Big(s_t^P + & \underbrace{A_j \times \mathbb{1}_{urban_t^P=0} + w_t \times \mathbb{1}_{urban_t^P=1}}_{\text{(labor income)}} + \underbrace{B \times \mathbb{1}_{g_t=0}}_{\text{(bequest)}} \\
& - \underbrace{c_t^P \times (1 + \alpha \times \mathbb{1}_{urban_t^P=urban_t^C})}_{\text{(daily consumption)}} - \underbrace{tuition_t}_{\text{(child's education)}} - \underbrace{Tr_t}_{\text{(transfer)}} \Big).
\end{aligned} \tag{4}
$$

Grandparent's preferences and budget constraint The grandparent, with similar preferences as the parent, values her health status and consumption on healthcare. She gains utility from private consumption c_t^G and leisure l_t^G through an isoelastic utility function with a same CRRA parameter γ and comsumption-leisure trade-off θ. Her utility is reduced by η_1 when she is sick, but is restored by η_2 if she spends c_t^h on healthcare to receive medical treatment[15]. She is altruistic toward her grandchild's education and gains the same φ of her utility if the grandchild is enrolled in high school. As in the case of the construction of the parent's utility, the child's education and the grandparent's health affect the grandparent's gain from her private utility, and also cause a fixed welfare gain/loss with factor δ[16]. The flow utility function of the grandparent is defined as follows:

$$
\begin{aligned}
U_t^G &= u^G \left(\, c_t^G, l_t^G, enroll_t, c_t^h \, \right) \\
&= \underbrace{\frac{\left(c^{G\ \theta} \ l^{G\ 1-\theta} \right)^{1-\gamma} - 1}{1-\gamma}}_{\text{(private utility)}} \times \underbrace{(1 - \varphi \times \mathbb{1}_{enroll_t=0})}_{\text{(child's education)}} \times \underbrace{HC_t}_{\text{(health care)}} \\
&\quad + \delta \times (1 - \varphi \times \mathbb{1}_{enroll_t=0}) \times HC_t
\end{aligned} \tag{5}
$$

in which HC_t, the factor on health status and healthcare, is 1 if the grandparent is healthy, and is smaller than 1 and depends on whether she spends on healthcare if she is sick. The specific definition of the

[15] The consumption on healthcare is discretized into a homogeneous annual cost paid by the grandparent, i.e. $c_t^h \in \{0, \bar{c}^h\}$ 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.

[16] 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.

healthcare factor is in equation (6).

$$HC_t = \begin{cases} 1 & \text{if } h_t = 0 \quad \text{(healthy)} \\ 1 \quad \underbrace{- \ \eta_1}_{\text{(disutility of illness)}} \quad + \underbrace{\eta_2 \times \mathbb{1}_{c_t^h > 0}}_{\text{(utility of health care)}} & \text{if } h_t = 1 \quad \text{(sick)} \end{cases} \tag{6}$$

The grandparent receives remittance Tr_t from the parents and has a constant flow income from agricultural production A_j when she is healthy. She lives on savings s_t^G and the parent's financial support when she is sick. The grandparent spends on daily consumption of herself c_t^G and her co-residing grandchild $c_t^G \times \alpha \times \mathbb{1}_{urban_t^P \neq urban_t^C}$ as well as her healthcare c_t^h if they are sick. The grandparent's intertemporal budget constraint is defined as follows:

$$\begin{aligned} s_{t+1}^G = (1+r) \Big(s_t^G + & \underbrace{A_j \times \mathbb{1}_{h_t=0}}_{\text{(agricultural income when healthy)}} + \underbrace{Tr_t}_{\text{(transfer from parent)}} \\ & - \underbrace{c_t^G \times (1 + \alpha \times \mathbb{1}_{urban_t^P \neq urban_t^C})}_{\text{(daily consumption)}} - \underbrace{c_t^h}_{\text{(spending on healthcare)}} \Big) \end{aligned} \tag{7}$$

4.2 Economic and social environment

Labor market and earnings The parent may farm on land assigned to the household, or migrate to enter the urban labor market. The rural income is heterogeneous and indexed by household type j, while the hours are homogeneous across types. The urban labor market offers a constant wage $\bar{w}$ to all migrants whenever they are employed. I assume that rural parents who migrate to the urban area are employed full-time in period 1 through 4. In period 5, a migrant may be unemployed with probability pr^{ump}. Therefore, the expected income of a migrant is defined as the expected wage rate $w_t = \bar{w} \times (1 - pr^{ump} \times \mathbb{1}_{t=5})$. 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 D.1).

The parent's migration decision $urban_t^P \in \{0, 1\}$ is equivalent to a choice between two pairs of wage and hour: (A_j, T_{rural}) and (w_t, T_{urban}). The grandparent is in the rural labor market when she is healthy and exits the labor force as soon as she becomes sick.

Daily consumption Rural and urban areas in China have segregated goods markets. I use a price ratio, $\frac{p^{urban}}{p^{rural}}$, between the two markets to convert nominal consumption to standard of living. Additional details of the definition and estimation procedure of this ratio are in Section 5 and Appendix D.

Tuition As discussed in Section 3.2, the tuition amounts differ by location and level of schooling. For each time period and location, the education investment choice is binary between a homogeneous tuition and zero, i.e. $tuition_t \in \left\{\overline{tuition_t}(urban_t^C), 0\right\}$.

Healthcare I assume that the grandparent is the only consumer of healthcare. She does not consume healthcare when she is healthy. When she becomes sick, she decides whether or not she is willing to pay a homogeneous annual cost c_t^h to be treated for her illness[17], i.e. $c_t^h \in \left\{\bar{c}_t{}^h, 0\right\}$. To account for the higher costs of intensive medical treatment in the last year of the grandparent's life, a lump sum cost is imposed on every household immediately after the grandparent's death[18]. This amount is deducted from the grandparent's bequest before it is transferred to the parent. If the bequest is not enough to cover the payment, however, the remainder is paid for by the parent.

Leisure The leisure time of the parent and grandparent is defined as the number of hours net of labor supply, child care, and private care for the sick grandparent. As noted in Section 3.1, more hours are committed to work in the urban area than in the rural area. Childcare takes more time in period 1, when the child is between 0 and 5 years old, than in later periods, and takes no time when the child leaves home in period 5, when she goes to college or enters the labor market. Private care from the parent to sick grandparents takes a fixed number of hours per week. I denote the weekly endowment of time by T_{total}, the hours spent on working in rural and urban areas by T_{rural} and T_{urban}, the hours spent on childcare by T_t^C, and the hours spent on private care for sick grandparent by T^G. The respective leisure times of the parent and grandparent are specified in equation (8).

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

Uncertainty There are two sources of uncertainty in the model. First, the health status of the

[17] I provide supporting evidence of the assumption in Appendix E.4.

[18] 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 constraints of the parents and grandparents for clarity. I provide supporting evidence of this specification in Appendix E.4.

grandparent changes over time, with transition probabilities depending on the grandparent's age, proxied by t, and whether the parent fulfills the contract by providing daily care when the grandparent is sick. The model assumes no recovery from sickness[19] and people are assumed not to die without first being sick for a period (i.e. no sudden death), except for the last period. Let ρ^{death} denote the probit model coefficient for the effect of the absence of the parent's private care on the grandparent's mortality risk, conditioning on the presence of the informal contract. The transition probabilities between health status are defined as follows:

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

Second, if the child is enrolled in school, she may not be able to graduate from her current stage of education. As the empirical facts in Section 3.3 show, the child's school performance depends on the child's education level and whether she is left behind. The dropout probability is a function of time period t and both her and her parent's migration status:

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

4.3 Informal contract and the sequential game

The informal contract between the grandparent and the parent is a trade of the grandparent's childcare provision for the parent's financial transfer and elder care. Only healthy grandparents can propose the informal contract. If the grandparent proposes the contract, she is obliged to take care of the child when the parent migrates without the child for as long as she is healthy[20]. If the parent accepts the contract, he is obliged to send remittance when he is in the urban area, and to provide financial support and private care when the grandparent is sick. If the parent reneges on the contract, he suffers from guilt. This subsection describes the inter-generational behaviors related to the informal contract in detail.

Sequential game played between the parent and grandparent The series of the two agents' behaviors around the informal contract forms a sequential game. Table 4 summarizes the timeline and

[19] I provide supporting evidence of this assumption in Appendix E.3.

[20] The grandparent cannot take care of the child when she is sick. Thus, the parent either returns to the rural area to live with the child, or migrates to the urban area, taking the child with him.

presents the choice and state variables in each period of the model.

Table 4: Timeline of the model

Stage of the household	Child's Birth	Three generation 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

Here I describe the sequence of the actions in detail:

1. When the child is born, the grandparent moves first by deciding whether or not to propose the contract.

2. If the grandparent proposes the contract, the parent may accept the contract at any time as the child grows up, signaled by his decision to migrate without the child.

3. In every subsequent period, the parent moves first by making decisions regarding migration, consumption, and transfer. The grandparent then observes the transfer she receives, and moves next by making decisions regarding daily consumption and healthcare.

4. When the parent migrates, 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 sick, the parent must stay in or return to the rural area to take care of the grandparent and pays for the grandparent's total expenditure.

Contract The contract holds only if both agents accept it. It is not reversible —when both agents accept the contract, the contract stands and will hold until the end of the model. The grandparent has full commitment to the contract, with $contract^G \in \{0, 1\}$ in which 1 means propose, while the parent has limited commitment. The indicator for the parent's contract status depends on the grandparent's contract status and evolves with t as defined in equation (11).

$$contract_t^P = \begin{cases} 1 \quad \text{if} \quad \underbrace{contract^G = 1}_{\text{(grandparent proposes)}} \quad \text{and} \quad \underbrace{urban_t^P \neq urban_t^C}_{\text{(parent leaves child behind)}} \\ 1 \quad \text{if} \quad \underbrace{contract^G \times contract_{t-1}^P = 1}_{\text{(contract held in the previous period)}} \\ 0 \quad \text{otherwise} \end{cases} \tag{11}$$

Transfer A parent within the contract is obliged to pay positive remittance in two scenarios: The first scenario is when the parent migrates while leaving the child behind to be taken care of by the grandparent. In this case, the remittance must cover the child's daily consumption[21]. The second scenario is when the grandparent is sick. In this case, the remittance must cover the grandparent's daily and medical consumptions. The amounts of the transfers are summarized in equation (12).

$$\begin{cases} Tr_t = \underbrace{\alpha \times c_t^G}_{\text{(covers child's consumption)}} & \text{if } \underbrace{urban_t^P \neq urban_t^C}_{\text{(parent leaves the child behind)}} \\ Tr_t = \underbrace{(c_t^G + \bar{c}_t^h)}_{\text{(covers grandparent's consumption)}} & \text{if } \underbrace{h_t = 1}_{\text{(grandparent is sick)}} \\ Tr_t = 0 & \text{otherwise} \end{cases} \tag{12}$$

In addition, a parent who migrates when the grandparent is sick sends a remittance that covers her daily consumption, regardless of the informal contract status[22].

$$Tr_t = c_t^G \qquad \text{if } contract_t^P = 0 \text{ and } h_t = 1 \text{ and } urban_t^P = 1 \tag{13}$$

Guilt If the parent reneges on any component of the contract, he suffers from guilt. The guilt accounts for two obligations specified in the contract separately. Let parameter κ_1 denote the guilt from reneging on the financial component of the contract, i.e. the inequalities specified in equation (12), and parameter κ_2 denote the guilt from reneging on the care component. The level of guilt is cumulative in

[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.

[22] When the grandparent is sick, the parent has to migrate with his child.

that guilt from a new breach of the contract adds to the flow levels, and guilt from past breaches persists.

$$g_t = \begin{cases} max\left\{\underbrace{g_{t-1}}_{\text{(last period's guilt)}}, \kappa_1 \times \mathbb{1}_{\underbrace{Tr_t < \bar{c}^h_t + c^G_t}_{\text{(low transfer)}}} + \kappa_2 \times \mathbb{1}_{\underbrace{urban^P_t = 1}_{\text{(no elder care)}}}\right\} & \text{if } \underbrace{\text{contract}^P_t = 1 \text{ and } h_t = 1}_{\text{(contract holds, grandparent is sick)}} \\ g_{t-1} \quad \text{(guilt persists)} & \text{if contract}^P_t = 1 \text{ and } h_t \neq 1 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

4.4 Government

The government has three policy channels that may affect the decisions depicted in the model.

Health insurance Most Chinese citizens with rural Hukou were covered by the New Rural Cooperative Medical Care System (NRCMCS) by 2010[23]. The government gradually adjusted the coverage to increase its copay fraction. I model the government's policy of medical insurance coverage by specifying that the government pays a fixed fraction ρ^h of the total medical cost $\tilde{c}^h$ when the grandparent is sick. Therefore, the homogeneous private cost of healthcare $\bar{c}^h_t$ is:

$$\bar{c}^h_t = \left(1 - \rho^h\right) \times \tilde{c}^h \tag{15}$$

Education subsidy Education, especially compulsory education (including primary and middle school), is cheaper if children attend school in their Hukou location. For this reason, migrants must pay higher tuition if they bring their children to the urban area and do not wish to terminate the child's education. In addition, urban schools who admit rural children are under-regulated[24] and more expensive than rural public schools. Therefore, the education cost of a child with migration status $urban^C_t$ is

$$\overline{tuition}_t = tuition^R_t \times (1 - urban^C_t) + tuition^U_t \times urban^C_t \tag{16}$$

where $tuition^R_t$ and $tuition^U_t$ represent rural and urban tuition by education level indexed by t.

Rural social security system The Chinese government has a Rural Minimum Living Security (*Dibao*) system[25]. The system subsidizes rural people below the poverty line to ensure their annual income reaches 1,210 RMB ($174.5)[26]. In my model, I denote the minimum consumption level specified

[23] "Enrollment Rate for Three Basic Medical Insurance was More than 95% in 2014." CNR.cn.

[24] Hernndez, Javier C. and Iris Zhao. "One Target in Beijing's Migrant Crackdown: Schoolchildren". New York Times (2017).

[25] "China's new approach to beating poverty." The Economist (2017).

[26] "Annual report of statistics of social services", Ministry of Civil Affairs of the People's Republic of China (2009).

in the rural security system by c_{min}. If a rural resident's total resources are below the poverty line (i.e. $s_t + A_j < c_{min}$), then I assume that s/he spends all of his/her savings ($s_{t+1} = 0$) on housing, food, and commodities but spends nothing on education or healthcare. The government then increases his/her consumption level to the poverty line ($c_t = c_{min}$). 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.

4.5 State variables, choice variables and the maximization problem

I specify the initial and terminal conditions of the model, and then describe the maximization problem of the parent and the grandparent to formulate the structure of the model.

Initial condition and terminal condition The model starts when the child is born. I assume that parents have zero savings, while the grandparents have positive initial wealth. I assume that grandparent's initial wealth is a linear function of A_j, the agricultural income per capita of this household. Therefore, $s_0^P = 0$ and $s_0^G = \omega A_j$. The grandparent may be healthy, sick, or dead when the child is born, with probability distribution $pr(h_0 = 0), pr(h_0 = 1), pr(h_0 = 2)$.

The model ends when the child is 22 years old, right after college graduation (for those who go to college). All children exit education at the end of the model with final outcome edu_T. I assume all parents live in the rural area thereafter. I also assume that, at the end of the model, all grandparents are dead[27].

I define the parent's terminal utility as the present discounted value of the parent's flow utility for the next 20 years. I assume that he completely smoothes consumption throughout those years, i.e. ${c_j}_T^P$ satisifies $\sum_{i=0}^{19} \frac{{c_j}_T^P}{(1+r)^i} = s_T^P + \sum_{i=0}^{19} \frac{A_j}{(1+r)^i}$. I also assume that he only spends time on agricultural production, i.e. $l_T^P = T_{total} - T_{rural}$. His utility gain from the child's school enrollment is an isoelastic utility function of the child's education outcome at $t = T$. Equation (17) defines the terminal utility of the parent.

$$U_{j,T}^P = \sum_{i=0}^{19} \beta^i \Big(\frac{\left({c_j}_T^{P\ \theta}\ {l_T}^{\ 1-\theta}\right)^{1-\gamma}}{1-\gamma} \times \frac{edu_T^{1-\lambda}}{1-\lambda} \times \ (1-g_T) \Big) \tag{17}$$

I also define the grandparent's terminal utility as the isoelastic utility gain from the child's education outcome at time of the grandparent's death. The flow utility of the grandparent after his death are set to zero[28].

[27] I provides the distribution of health status of the grandparents by children's age in Appendix F.1.4 (my analysis, CFPS 2010-2014). In addition, the life expectancy is 58.99 for rural men and 72.46 for rural women (Shen, 1993).

[28] The factor δ is difficult to identify in my model given the data available. So I set $\delta = 1$ in my estimation.

$$U^G_{t_{death}} = \delta \frac{edu^{1-\lambda}_{t_{death}}}{1-\lambda} \tag{18}$$

Parent's optimization problem The parent's choice variables are his migration decision for himself ($urban^P_t$) and his child ($urban^C_t$), and the daily consumption for himself (c^P_t) and, implicitly, for his co-residing child through the equivalence scale (c^C_t)[29], the child's tuition ($tuition_t$), and the remittance to the grandparent (Tr_t). These variables are summarized as $choice^P_t = \{urban^P_t, urban^C_t, c^P_t, tuition_t, Tr_t\}$.

The parent's maximization problem is presented in equation (19), subject to the state variables X_t defined in equation (1), the definitions of the flow utility function in equation (3), the budget constraint of the parent in equation (4), l^P_t in equation (8), g_t in equation (14), $enroll_t$ in Section (4.1), in addition to the probability distribution over future states as specified in equation (9) and (10).

$$V^P_t\left(X_t \left| j \right.\right) = \max_{choice^P_t}\left(u^P\left(c^P_t, l^P_t, enroll_t, g_t\right) + \beta E_t\left[V^P_{t+1}\left|X_t, j, choice^P_t\right.\right]\right) \tag{19}$$

Grandparent's optimization problem The grandparent's choice variables are the one-time contract decision ($contract^G$), private consumption for himself (c^G_t) and implicitly for the co-residing child through the equivalence scale (c^C_t)[30] and consumption on healthcare (c^h_t). These variables are summarized as $choice^G_0 = \{contract^G\}$ and $choice^G_t = \{c^G_t, c^h_t\}$ for $t > 0$.

The grandparent's maximization problem is presented in equation (20) and (21), subject to the state variables X_t defined in equation (1), the definitions of the flow utility function in equation (5), HC_t in equation (6), the budget constraint of the grandparent in equation (7), l^G_t in equation (8), $enroll_t$ in Section 4.1, in addition to the probability distribution over future states as specified in equation (9) and (10).

Equation (20) below states that when the child is born i.e. $t = 0$, the grandparent chooses whether to propose the contract by comparing the expected value function of the two futures, conditioning on her household type j. Equation (21) states that while the child is growing up, in each period, the grandparent chooses c^G_t and c^h_t to maximize her value function, conditional on her household type, the state variables including her previous contract choice, and the parent's choices in the same period.

[29] Once c^P_t is chosen, c^C_t is determined by the equivalence scale. Therefore, c^C_t is not a separate choice variable for the parent.

[30] Similar to its role in the parent's optimization problem, c^C_t is a separate choice variable for the grandparent.

$$t = 0: \ V_0^G(j) = \max_{contract^G} \Big\{ E_0 \left[V_1^G \,\middle|\, contract^G = 0, j \right], \quad E_0 \left[V_1^G \,\middle|\, contract^G = 1, j \right] \Big\} \tag{20}$$

$$t > 0: \ V_t^G\left(X_t \mid j, choice_t^P\right) = \max_{choice_t^G} \Big(u^G\left(c_t^G, l_t^G, enroll_t, HC_t \,\middle|\, choice_t^P\right) + \beta E_t\left[V_{t+1}^G \,\middle|\, X_t, j, choice_t^P, choice_t^G \right] \Big) \tag{21}$$

The maximization problems described in equation (3) - (21) is computationally demanding. I solve the problem using high performance computation and memory intensive computers through the algorithms discussed in Section 5.1.4 and in more detail in Appendix G.

5 Estimation and results

This section describes the structural estimation process and its 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. Section 5.1.1 describes the estimation procedure in detail.

In the second step, I estimate 9 preference parameters and one heterogeneity parameter using the Generalized Method of Moments. Section 5.1.2 provides the objective function of the structural estimation. Section 5.1.3 discusses the parameters I structurally estimate, the moments that I match in the data, and how the moments identify the parameters. Section 5.1.4 introduces my computational strategy.

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 F, I provide additional information on the estimation of the parameters.

Distribution of agricultural income I obtain the distribution of agricultural income per capita from 3,509 rural households in the CFPS 2010 data. I divide the household-level agricultural income by the number of adults living in the rural area to calculate the per person income. The distribution of per capita agricultural income is a left-skewed distribution, and 90% of the households' income level falls between 500 RMB/year and 10,000 RMB/year (roughly \$73 and \$1,459). I discretize the sample by income level into 10 groups with comparable sizes, indexed by j, and the fraction of group j in the data is used as the probability of a household being type j, i.e. with agricultural income per capita A_j.

Price ratio and subsistence consumption level Due to the market segregation between urban and rural goods and income inequalities, a price ratio between the two areas is required to transform nominal consumption into purchasing power (or standard of living). However, an estimate of the price ratio for consumption net of education and healthcare is not available. I combine the consumption of food, clothes, commodities, and housing to define daily consumption in the model. Thus, a ratio of the prices of a certain good, e.g. rice, is not representative, and a ratio of the total price of a same group of goods is not feasible since consumption bundles differ by location. To overcome this problem, I estimate the price ratio from the Engel curves for food expenditure among the total daily consumption (Hamilton, 2001; Almås et al., 2018)[31].

Time allocation I estimate average hours at work in the rural and urban areas from the RUMIC 2009 data, which is convenient for computing statistics based on the migration status of people with rural Hukou. On average, a migrant works for 66 hours per week in the urban area, and a non-migrant works for 50 hours in the rural area.

The average number of hours needed to care for a child under 15 years old is 12 hours per week, estimated from the CHNS data. The average number of hours needed to care for a sick grandparent is also 12 hours per week and is estimated from the same data.

I compute the number of hours of leisure time by assuming an endowment of 12 hours per day, and subtract the hours taken by each individual labor and home production activity from the endowment.

[31] 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 14 in Appendix D).

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

$$\psi^* = \underset{\psi}{argmin} \int_0^1 \mid \psi \times c^R(F_{food}) - c^U(F_{food}) \mid \; dF_{food} \tag{22}$$

The estimated optimal value for ψ is 6.48.

In addition, 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). In the same year, 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.

Cost of healthcare Initially, patients pay the total cost of healthcare; then, the government reimburses a fraction of their expenses. I obtained the government's co-pay fraction from Deng et al. (2017). 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. However, 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.

Baseline probabilities on health status The five datasets available are either cross-sectional or short panel datasets with 2 or 3 interviews. This means there are a very limited number of observations of a transition in health status. Moreover, because my model is organized by childrens age groups, I use these groups as the basis to estimate the health-related probabilities of the elderly. To overcome these difficulties, and to make better use of the CFPS data, I implemented a two-step estimation procedure, the only dataset that has informative variables on all three generations. 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 sick[32]. I estimate the distribution of the health status of the grandparents by the children's age group on this sample. Second, I recover the health transition probabilities from the static distribution. The details of the estimation procedure is explained in Appendix F.1.4.

In addition, I estimate the effect of the parent's migration on the grandparent's mortality from CFPS. I run a logit regression of the grandparent's survival based on whether the parents took care of them in the past. The analysis suggests that lack of adult children's private care increases the grandparent's mortality rate by 1.67. I present the regression results in Appendix F.

Dropout probability by education level and guardian The sample size is too small to allow the dropout probability to depend flexibly on the child's level of education, the grandparent's role as guardian, and the guardian's education. Taking advantage of the findings from Section 3.2, I assume that the effect of the guardian on the dropout probability can be separated into two parts: the effect of the education level of the guardian on the child's dropout probability, and the guardian-specific education attainment distributions. Both of these effects can be measured from the CFPS data with convincing sample sizes. Let $pr_{t_c,i}^{dropout}$ denote the dropout probability for a specific education level t_c and guardian $i \in \{\text{parent}, \text{grandparent}\}$, and let edu_g denote the education level of the guardian. The

[32]In order to eliminate the effects of poor financial condition or lack of private care on mortality.

dropout probability based on the child's education level and the guardian's can be computed following equation (23).

$$pr^{dropout}_{t_c,i} := \sum_{t=0}^{16} pr(dropout\,|t_c, edu_g = t) \times pr(edu_g = t\,|i) \tag{23}$$

Note that, to separate involuntary dropouts caused by poor economic conditions from dropouts caused by poor performance or lack of incentive to continue education[33], when estimating the right-hand side probabilities in equation (23), I exclude children who claimed that they dropped out because of economic difficulties.

Risk in college admission In period 5, if the parents decide to invest in the children's college education, the children have an opportunity to take the National College Entrance Exam (*Gaokao*). 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. The estimated probability of being admitted by any college conditioning on taking the exam is 27%[34].

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 (24):

$$\hat{\theta}_{GMM} = \arg\min_{\vec{\theta}} \; (Q_0 - Q(\vec{\theta}))' W^{-1} (Q_0 - Q(\vec{\theta})) \tag{24}$$

5.1.3 Parameters, moments and identification

In this section, I discuss the parameters estimated using moments implied by the model. Table 5 lists the 11 structurally estimated parameters and their key sources of identification. For each parameter, I summarize the behaviors that the parameter influences, and discuss the moments that capture those behaviors[35]. At the end of this section, I introduce the weighting matrix I used in the GMM estimation.

[33] Sociology research argues that migrants' children are more eager to quit school earlier to become a migrant worker as they see their parents as examples (China Women's Federation Children's Work Department, 2011).

[34] 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.

[35] A complete list of moments is provided later in Section 5.2.2.

Table 5: Parameters estimated in the structural model

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

Discount factor (β in equation (19) and (21)) affects the value of future utilities. A parent with a high discount factor is more likely to prefer staying in the rural area (to fulfill the informal contract) over reneging on the contract (to migrate for higher income and consumption). A grandparent with a high discount factor is more likely to propose the contract to trade her leisure time in the first period (because of childcare) for higher consumption and the opportunity to receive healthcare in the future. I use the fraction of migrants in each period and the parents' decision of whether to fulfill each component of the informal contract to identify the discount factor.

CRRA coefficient (γ in equation (3)) influences the marginal return to additional consumption. It affects the incentive to migrate for higher income. A parent with a high value of γ has lower utility gain from working in the urban area, so he would be less likely to renege on the informal contract. His insensitivity to a consumption drop would also encourage investments in children's education. Consequently, a high value of γ should result in a higher rate of enrollment in school. On the other hand, a grandparent with a high CRRA coefficient is relatively insensitive to change in consumption. In this sense, she would be more likely to spend on healthcare. I use the fraction of migrants, the fraction of left-behind children, the children's school enrollment rates, the fraction of grandparents spending on healthcare, and the fraction of parents fulfilling the private care obligation of the informal contract to identify the CRRA coefficient.

Consumption-leisure tradeoff (θ in equation (3) and equation (5)) influences the parent's choice between an urban and rural labor market, his migration decision for his child, and his private care decision for the grandparent. If θ is small, parents are less likely to leave their children behind. The value of θ also affects how parents with an informal contract trade off between financial support and elder care for sick grandparents. Grandparents with low value of θ experience a smaller utility loss from time spent on childcare. They would be more willing to propose the informal contract, which trades their time spent on childcare for additional income that they can use for daily consumption and healthcare. Therefore, I identify parameter θ by the fraction of migrating parents and the fraction of left-behind children, especially those in the first period. The fraction of parents fulfilling each component of the informal contract also contributes to identifying this parameter.

Children's education parameters (φ in equation (3) and λ in equation (17) and (18)) influence the parent's incentive to pay for tuition and his decision regarding the child's migration. Parents with a utility gain from their children's education are more likely to migrate when the children are enrolled, in order to ensure financial support for the children's future education. The children's school performance is better if their migration status is consistent with their parents', so parents are less likely to leave the children behind if they care more about the children's educational outcome. A higher return to the children's education increases the parents' likelihood of reneging on both the financial component and the care component of the informal contract. In addition, grandparents receive a higher utility gain in their terminal condition if they live longer while the children are enrolled in school. The two education parameters affect the magnitude of this utility gain. As a result, the return to spending on healthcare increases as φ increases and as λ decreases. Note that while λ affects the returns to children's education at all levels, parameter φ only changes the utility gain from high school enrollment. Therefore, I identify the two education parameters by the children's education enrollment rate, the children's average education attainment as of 22 years old. The fraction of migrants over time, the fraction of migrating children, the fraction of parents fulfilling each component of the informal contract, and the fraction of grandparents who spend on healthcare also contribute to the identification of the two parameters.

Guilt from reneging on the financial and the care-related terms in the informal contract (κ_1 and κ_2 in equation (14)) influences the parent's behavior when the grandparent is sick. When κ_1 is high, parents participating in the informal contract send more financial transfers when the grandparents are sick. When κ_2 is high, more parents stay in the rural area to provide elder care to the sick grandparents. I use the fraction of parents fulfilling each component of the informal contract to identify these two parameters.

Utility gain from healthcare when the grandparent is sick (η_2 in equation (6)) influences her medical consumption, and affects her choice regarding the informal contract. When grandparents care about healthcare, they are more likely to spend on healthcare. Furthermore, the value of the informal contract to grandparents increases as healthcare becomes a greater concern. More grandparents propose the contract, and thus, more children are left behind. Consequently, more grandparents receive financial support from parents. I use the fraction of grandparents who spend on healthcare, the distribution of grandparents' health status over time, and the fraction of grandparents receiving financial transfers from parents to identify the parameter on healthcare.

Coefficient on additive component of utility function (δ in equation (3) and (5)) influences the agents' sensitivity to the child's education, informal contract status, and healthcare. The higher the value of δ, the more households are willing to cooperate, and the more parents fulfill the contract Therefore, the additive factor is identified by the fraction of migrants, the fraction of migrating children, the average amount of money remitted to grandparents, and the average rural consumption level.

The unemployment rate in period 5 (pr^{ump}) affects the expected return of migration in that period, and thus affects migration behavior. When the unemployment rate is high, fewer parents migrate in period 5. If the grandparents are sick in period 5, they are more likely to be cared for by parents and thus live longer. As an indirect effect, they are more eager to migrate before period 5, and thus are more likely to renege on the care component of the informal contract.

Weighting matrix in GMM estimation I use an identity weighting matrix for the 26 moments on various fractions. I give smaller weights on the three moments on the average of rural consumption, the average of financial transfers to grandparents, and the average education attainment of rural children, in order to standardize them to the same scale as the fractions. The small weights for the three quantities are determined from the standard errors from data estimates[36].

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 following vector:

[36] I do not use the optimal weighting matrix. Most moments I match in the data are nearly perfectly estimated without substantial sampling errors. 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 of the moments with equal weights, so I assign higher weights to moments estimated on smaller samples than their weights in the optimal weighting matrix.

$$\left\{j, contract^G, \left\{X_t, choice_t^P, choice_t^G\right\}_{t=1}^5\right\}$$

Finding 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[37]. 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[38]. I use various techniques to improve efficiency. The resulting algorithm takes 17.164 hours for a one-core computer to complete one iteration[39].

Second, the objective function, as specified in equation (24) is not globally concave, which disqualifies gradient methods in the estimation, as the derivative-based search would easily fall into a local optimum that is not the global optimum. A grid search would be the most desirable approach for this type of objective function, but is not feasible 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[40]. My second technique is, within each group of estimation in the Genetic Algorithm, to divide each iteration into 40 non-overlapping optimal sub-tasks[41], 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, each

[37] 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.

[38] 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.

[39] (a) Techniques I applied include 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.

[40] 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 literature on the Mathematical Theory of Computation (Alander, 1992; Stanhope and Daida, 1998).

[41] An iteration can be divided into 10 different tasks by household type j. In each household, the grandparent has 3 possibilities of her initial health status. If the grandparent is healthy when the model starts, she has two options on the informal contract: accept or reject. A sick or dead grandparent, meanwhile, cannot enter the contract. Each combination of household type, initial health status, and the grandparent's choice on the contract defines a sub-task. Therefore, each iteration has 40 non-overlapping sub-tasks.

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 high performance computers for most of the time by centralized resource allocation and the efficient assignment of sub-tasks to computers. Figure 15 in Appendix G provides an example of the system's performance, illustrated by the CPU usage of 10 72-core high performance computers controlled by a one-core machine on Amazon Web Services.

5.2 Estimation Results

5.2.1 Parameter estimates

Table 6: Parameter estimates and standard deviation

Description	Symbol	Estimates	Standard 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
Coefficient on additive component of utility function	δ	0.3554	0.6202
Grandparent's utility gain from healthcare when sick	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 taking care of 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

Table 6 gives the estimates of the parameters. My estimates for the standard errors follow Lockwood (2018)'s approach. Prior works on the risk aversion 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 $\gamma = 1.2348$ is in line with existing literature.

The discount factor, $\beta = 0.8657$, 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 $\eta_2 = 0.3491$ indicates that grandparent gains from healthcare, which provides an incentive for the informal contract. The consumption equivalent of the utility gain from healthcare is a 32-fold increase in daily consumption. This gain is huge and implies that the grandparent would spend on healthcare whenever she can afford it[42]. The estimates of parameters governing the parent's and grandparent's altruism toward the child's education ($\lambda = 2.6493$) reveals a high marginal return to additional years of schooling. High school enrollment ($\varphi = 0.0245$) 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 ($\kappa_1 = 0.0060$) is equivalent to a 6% decrease in consumption. The guilt from reneging on the private care component of the contract ($\kappa_2 = 0.0501$) is equivalent to a 40% drop in consumption.

5.2.2 Model fit

Table 7 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 fraction of migrants as children grow up. The model slightly underestimates the fraction of migrants[43]. It also captures the decreasing school enrollment rates. When parents cannot fulfill both parts of the informal contract, their preference between the financial support and the 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 fraction of migrants and 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 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 to. 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

[42] Healthcare costs far more than the annual rural income for most grandparents (Appendix E.4). The consumption equivalent of the welfare gain from healthcare is supposed to be large as well.

[43] 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 7: Goodness of fit of model to the data

Moment			Data	Model Estimates
Migration				
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				
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				
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				
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				
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

parameters $\vec{\theta}$ at the parameter estimates presented in Table 6. Then I use the estimated Jacobian $\hat{G}$ and the weighting matrix W to calculate the sensitivity measure defined in equation (25).

$$\Lambda = (\hat{G}'W\hat{G})^{-1}\hat{G}'W \tag{25}$$

The results are consistent with the discussion in Section 5.1.3. In addition, 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. And 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 intra-household cooperation

Rural parents receive higher income from working in the urban area. Consequently, young children and elder grandparents may share the welfare gain from the parents' migration through intra-household cooperation. In this section, I demonstrate the quantitative importance of the informal contract for the welfare distribution within rural households.

In my model, the informal contract distributes the parent's gain from migration to the children and grandparents. The nature of the informal contract is a trade between the parents and grandparents involving assets, time, and private care. Entering the contract is voluntary, so it could only be adopted by a household if it benefits both parties (that is if it is in the contractual space of both agents). In this subsection, I discuss how the contractual space is affected by the household status, and the policy environment.

Table 8 lists each source of benefits and costs of the informal contract to each generation. In the table, I provide a brief explanation for including each factor in parentheses. I also record the model prediction of the overall effects of the informal contract to the three generations.

Figure 1 provides the parent's and grandparent's ex ante gain from the informal contract, by the household's agricultural income level, in consumption equivalent scale. It visualizes the expected lifetime

Table 8: Benefits and costs of the informal contract to the parent, grandparent, and child

	Benefits	Costs	Overall effect predicted by the model
Parent	• 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)	• 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	All parents gain from the contract
Grandparent	• Additional income (from remittance) • Private care when sick (parent fulfills contract)	• Income uncertainty (parent reneges) • Less leisure time (provide childcare) • Risk of not receiving elder care (parent reneges)	Poor grandparents gain from the contract
Child	• Higher consumption (from remittance) • More funding for tuition (migrant has more savings)	• Higher dropout probability (left-behind children have poorer school performance)	Average education attainment drops by 0.07 years because of the contract

utility gain of each agent from the informal contract. Line segment above one indicates that accepting the contract generates a welfare gain. The figure shows that the parent's welfare gain from cooperating with the grandparent decreases as the agricultural productivity rises.

In households that adopt the informal contract, migrating parents accumulated more assets rapidly by migrating without children. They only benefit from the contract as long as the grandparents are healthy. When the grandparents are sick, the informal contract either costs the parents in money or utility. Although the informal contract is of limited commitment from the parents' perspective, the disutility from reneging ensures that the contract embeds positive costs to the parents. My model estimates suggest that, in the current policy environment, 92% of parents from three-generational rural households could receive welfare gains from the informal contract if the grandparents propose the contract[44]. The value of the contract to parents increases as the rural-urban wage and the consumption price gaps rise. It decreases as the amount of remittances and private cost of healthcare rise, and as the time cost of caring for sick grandparents increases.

[44] In my model specification, households are categorized into 10 types by agricultural productivity. Figure 1 suggests that only parents from the wealthiest household type do not accept the informal contract even if the grandparents do. The share of the wealthiest household type in all rural households is 8% (my analysis, CFPS 2010-2014).

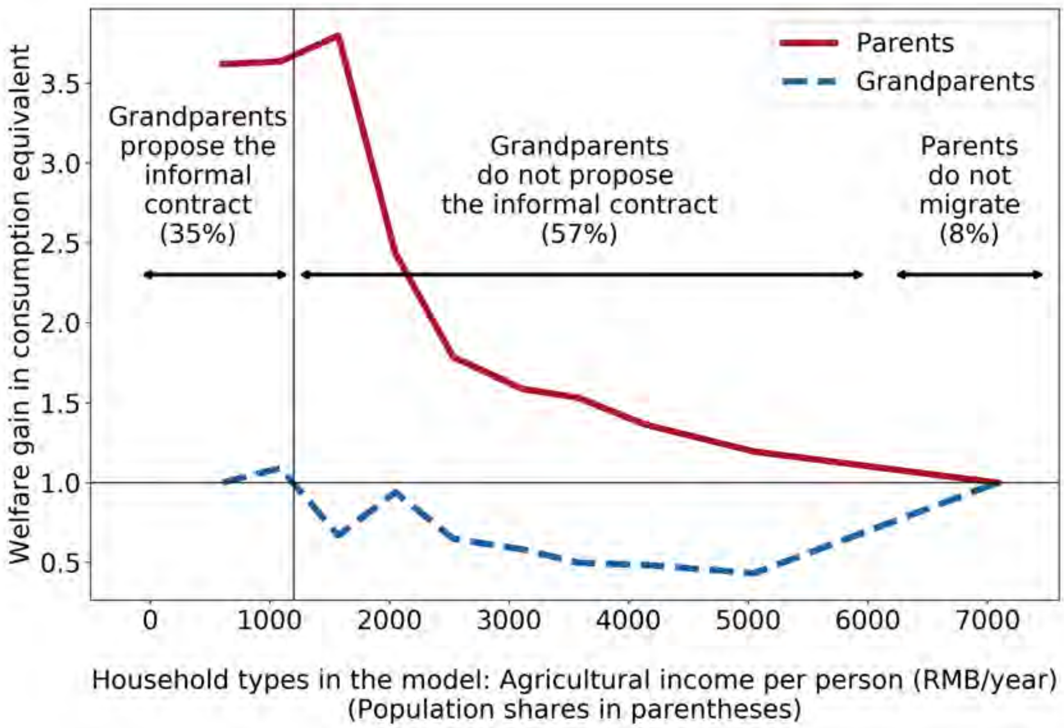


Figure 1: Parents' and grandparents' welfare gain from the informal contract by agricultural income

Note: The above figure presents the agents' welfare gain from the informal contract in consumption equivalent by household type. The vertical line is the household's agricultural endowment level that differentiates the grandparents' informal contract status.

Grandparents experience different welfare trade-offs raised by the informal contract depending on their health. The major source of costs is the possible large welfare loss caused by parents reneging on the contract. My model suggests that the welfare gain generated by the informal contract to the grandparents drops as a household's agricultural productivity rises. Among all rural households, 35% of their agricultural income per capita is below 1,200 RMB per year (my analysis, CFPS 2010-2014). For grandparents from these households, the potential welfare loss and uncertainty offsets the gain from the additional source of income and elder care that their adult children give. The value of the informal contract is low when parent's contract reneging probability is high. The magnitude of grandparent's utility gain from financial transfer depends on three factors. First, it depends on the extent to which the informal contract replaces benefits that would otherwise be provided by the social security system. Grandparents benefit more from intra-household cooperation when the crowd out effect is small. This is why the welfare gain of the informal contract is smaller for the poorest group (rural income = \$600) than for the second poorest (rural income = \$1,000). Second, the magnitude of the grandparents' welfare gain also depends on the consumption level, because of diminishing marginal utility to additional income.

And third, it depends on the amount of financial transfer[45]. On the other hand, the attractiveness of the contract is also related to the health condition and healthcare of the grandparent. It increases as the private cost of healthcare rises, and as the effect of private care from family members on the grandparent's mortality becomes more significant.

The effect of the informal contract and parent's migration on rural children is ambiguous from prior studies I summarized in Section 3.3 or from my model. On the one hand, living with less educated grandparents has a negative effect on children's education, and this effect is homogeneous with respect to the household's agricultural income level. On the other hand, because migrants are wealthier, they can spend more money on their children. The children have higher consumption levels (Figure 7 in Appendix B), and they are less likely to quit school because the household is financially constrained. Parents are also wealthier if their household's agricultural income level is high. Thus, the value of the informal contract to rural children decreases as the household's rural income rises. Intra-household cooperation decreases the average education attainment of the children by 0.07 year. My model estimates suggest that the overall effect of the intra-household cooperation on children's education attainment is negative but small.

As shown in Figure 1, all parents could benefit from the informal contract, regardless of income level. Wealthier grandparents, by contrast, could not. I examine two hypothetical scenarios to illustrate the effect of intra-household cooperation on rural consumption levels: one scenario in which all grandparents propose the informal contract to parents, and another in which none of them propose the contract. Figure 2 presents the average of individual-level consumption in the rural area by household's income level in these two scenarios.

Besides the wealthiest group of households in which parents choose not to migrate, all other households' consumption levels rise because of the informal contract. The strong correlation between agricultural income and consumption level of the household persists when all households adopt the informal contract. In my model, however, only poor households accept the contract. So the intra-household cooperation improves the welfare of poor families and increases their consumption level. It implies that the informal contract improves the living condition of the poor in rural China.

[45] Equation (12) in Section 4.3 implies that the transfer is higher when the grandparent's consumption level is high. When the grandparent is healthy, the transfer covers the child's consumption, which is a fixed proportion of the grandparent's consumption according to the OECD equivalence scale. When the grandparent is sick, the transfer covers the grandparent's consumption.

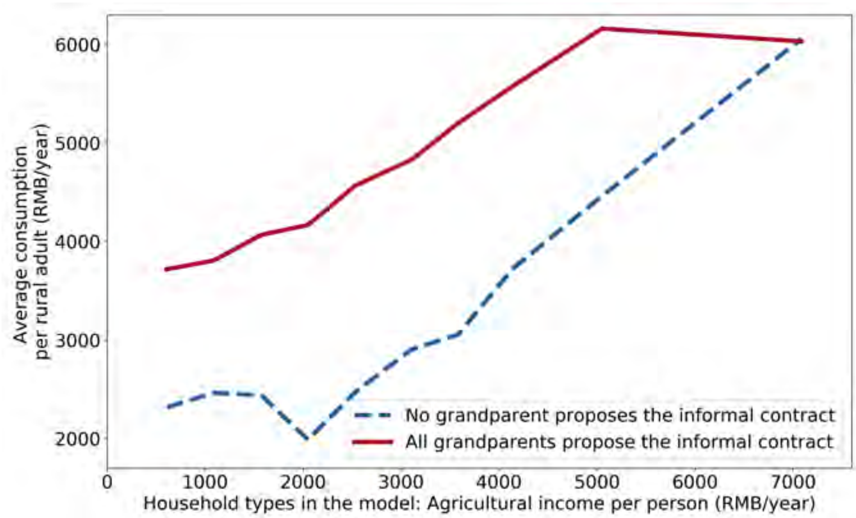


Figure 2: Average consumption by informal contract status and income level

Note: The above figure shows the average consumption level of rural adults by their household type and the grandparent's informal contract status. Rural adults include parents who live in the rural area and all grandparents.

7 Counterfactual experiments

The Chinese government has two objectives in policies targeting rural households and rural-to-urban migrants. The first objective to improve rural people's welfare; the second is to control the fraction of rural migrants in the urban areas.

In this section, I evaluate an urban education subsidy and an expanded health insurance subsidy. The discussion of policy effects will involve the welfare and behavioral change of each generation of a rural household, as well as the impact on the flow of migrants.

7.1 Urban Education Policy

Policy makers in China regard improved accessibility to urban education as a major tool to improve rural children's welfare and alleviate the problem of left-behind children. In the past few years, the government has initiated and implemented various policies in this direction[46], but policy evaluations are lacking. The direct effect of these policy reforms on rural households with migrants is that they face lower urban tuition when making migration, informal contract, consumption, and education decisions. In this subsection, I explore the effect of a policy that lowers the prices of urban schools to rural price levels.

[46] Policy proposals include increasing government spending on schools for migrants' children, subsidizing the migrants to lower urban tuition for their children, or enrolling the children in the urban public education system. As of 2018, the central government announced a future policy of promoting public urban school admission among migrants' children (The 13^{th} Five-Year Plan for Economic and Social Development of the People's Republic of China (2016-2020), Central Complication & Translation Press). In many large cities, local governments tightened regulation on the poorly organized and low quality urban private schools for migrants (Denton, Bryan. "One Target in Beijing's Migrant Crackdown: Schoolchildren." nytimes.com).

I add a policy parameter ρ^{edu} in (16) to depict the effect of the urban education policy on the urban-rural tuition gap ($tuition^U_t - tuition^R_t$). Therefore, the tuition is defined as follows:

$$\overline{tuition}_t = tuition^R_t + \left(1 - \rho^{edu}\right) \times (tuition^U_t - tuition^R_t) \times urban^C_t. \tag{26}$$

When $\rho^{edu} = 0$, the government does not subsidize urban education for rural children, and when $\rho^{edu} = 1$, the tuition is the same across rural and urban areas. I set $\rho^{edu} = 0$ in the baseline model[47]. The counterfactual experiments shifts the policy parameter ρ^{edu} from zero up to one. Thus, urban tuition from kindergarten to high school can be as low as rural tuition at the same education level[48].

Figure 3 shows the effects of the policy on a set of outcomes, including the average education attainment of the children, the fraction of migrating parents, the fraction of grandparents who spend on healthcare, and the average consumption level of the migrants.

Effects on the targeted generation The policy increases the average education attainment of the children by up to 0.1 year (Figure 3a). However, it does not change the fraction of left-behind children (Figure 8b in Appendix B).

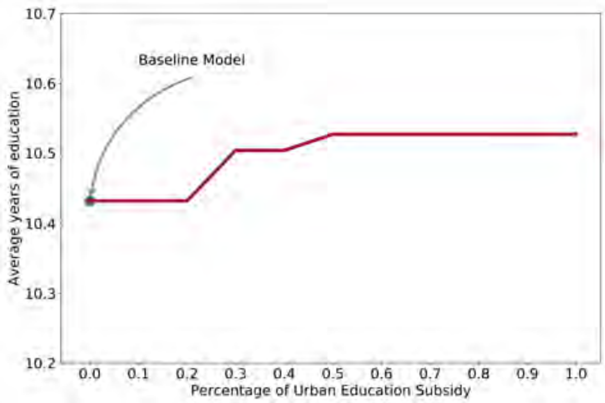
Effects on other generations With my estimated model, I can show the policy effects on other generations in the rural families. The policy changes the social and economic environment that rural households face. Parents' behaviors change because they face different price levels. Grandparents confront a different sequential game because parents' choice set changes. Therefore, in the new steady state, the policy could affect all generations in a household.

One might think that lowering urban tuition gives parents more incentive to migrate with children when the grandparents are sick, and that thus, more parents renege on the informal contract. The lower panel of Figure 4 shows that the parents' choices between fulfilling or reneging on the informal contract do not change. This means that the guilt associated with reneging on the contract is high enough to secure the welfare of the grandparents. In fact, the upper panel of Figure 4 shows that the policy generates a welfare gain to parents, while the grandparents only suffer a very small welfare loss.

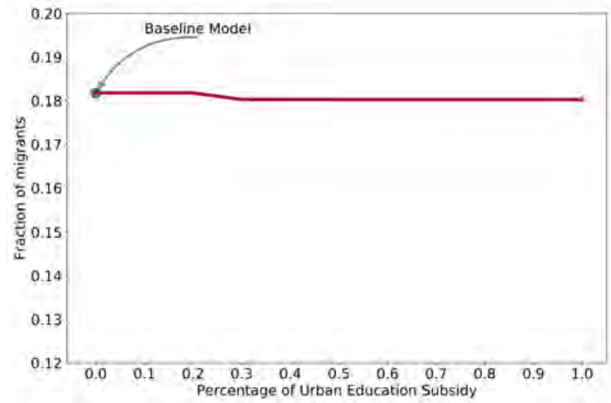
Parents gain utility from the higher education attainment of their children. Lower tuition rates have an income effect on migrating parents, so their consumption level rises by up to 6% as shown in Figure 3d. While the net return of working in the urban area rises for parents who have to migrate with children

[47] 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.

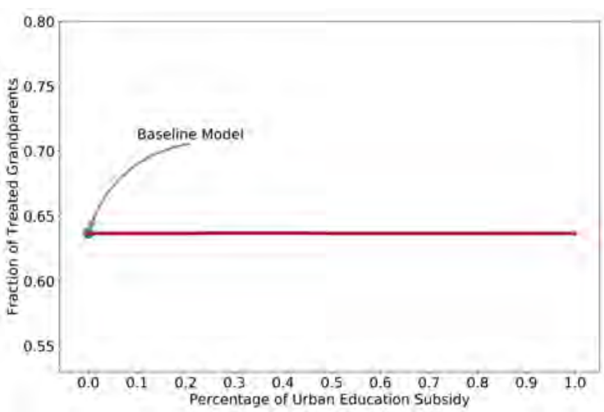
[48] Note that the rural school tuition rates are not affected by this policy. I also examined a similar policy counterfactual that only lowers the tuition for compulsory education periods (i.e. primary school and middle school). However, this policy has no effect on welfare and rural people's behaviors at all.



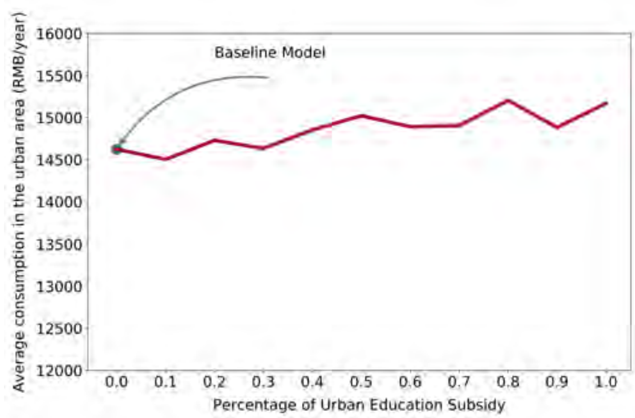
(a) Children have higher education attainment



(b) Parents' migration unaffected



(c) Grandparents' healthcare unaffected



(d) Migrants' consumption rises

Figure 3: Effects of urban education policy on rural households

Note: The above figures show the effects of urban education policy on the three generations. The calculation of the household outcomes accounts for heterogeneity in agricultural endowment of the household, the initial health condition of the grandparents, and the age distribution of the children.

who are enrolled in school, the overall fraction of migrants in the urban area is unaffected by the policy (Figure 3b).

Grandparents are in general unaffected by the policy. As parents' contract decisions are unaffected, grandparents do not suffer or gain from the informal contract. Figure 3c shows that the fraction of grandparents receiving healthcare is constant as the government increases urban education subsidies.

7.2 Expanded public healthcare coverage

The government may increase the public health insurance coverage to improve the grandparents' welfare. Most rural Chinese people are covered by the New Rural Cooperative Medical Care System (NRCMCS) as of 2010. The current coverage of the NRCMCS can be thought of as a plan with 34% reimbursement rate. The counterfactual experiment changes parameter ρ^h in equation (15) from 5% to 95%, to examine

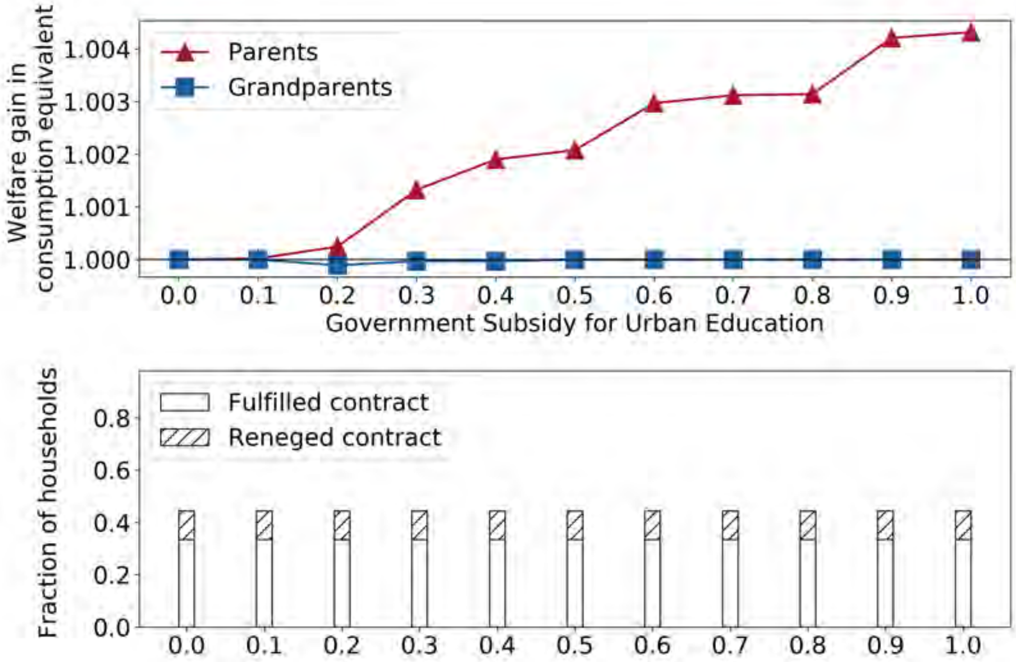


Figure 4: Education policy effects on welfare and the household's contractual space

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.

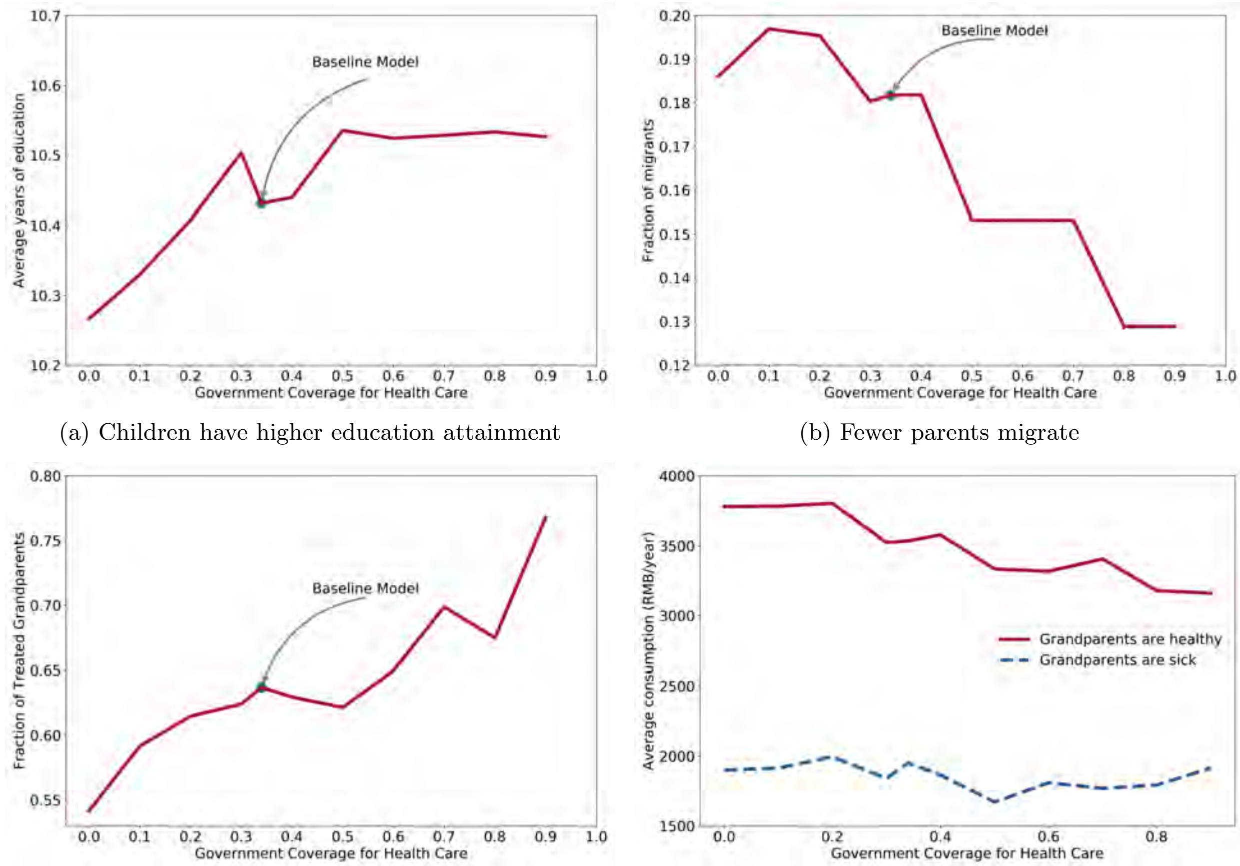
the reaction of the rural household members to this policy.

Figure 5 shows the policy effects on the average education attainment of the children, the fraction of migrating parents, the fraction of grandparents who spend on healthcare, and the average consumption level of the migrants.

Effects on the targeted generation Figure 5c shows that the fraction of grandparents who spend on healthcare increases by up to 13%. They adjust their consumption level when they are healthy in order to save for daily consumption and spending on healthcare when they are sick (Figure 5d).

Effects on other generations In the new steady state, the policy affects the equilibrium of the intra-household behaviors through the informal contract. Then these inter-generational effects change the welfare of all three generations in the family.

Cheaper 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. Affordable healthcare makes it more feasible for grandparents from poor households to stay out of the informal contract and be independent of the parents. As a result, the fraction of households adopting the contract decreases (the lower panel of Figure 6). In addition, the expanded health insurance policy also lowers the amount of transfer that



(a) Children have higher education attainment

(b) Fewer parents migrate

(c) More grandparents spend on healthcare

(d) Healthy grandparents' consumption drops

Figure 5: Effects of public health policy on rural households

a parent must pay to fulfill the contract. Thus, the fraction of reneged informal contract decreases.

The effect of the public health policy on grandparents is positive, and the magnitude of the policy effect is large. This is because grandparents who opt out of the informal contract receive a welfare gain from lowering the risk of not receiving financial support or private care from the parents. Grandparents become more independent of the parents, and are financially more self-supportive.

However, the policy hurts the parents. In households that opted out of the informal contract because of the policy change, parents no longer have the option of leaving the children. These parents' net gain from rural-to-urban migration decreases by the sum of child's daily consumption and tuition in the urban area. Therefore, the fraction of migrants drops by 5% (Figure 5b) and parent's urban consumption falls (Figure 9a in Appendix B).

Public health policy effects on children are mixed. Decreasing parents' migration tightens their budget constraints, and the income effect then causes parents to invest less on children's education. On the other hand, fewer parents leave their children behind because the informal contract is rejected by the

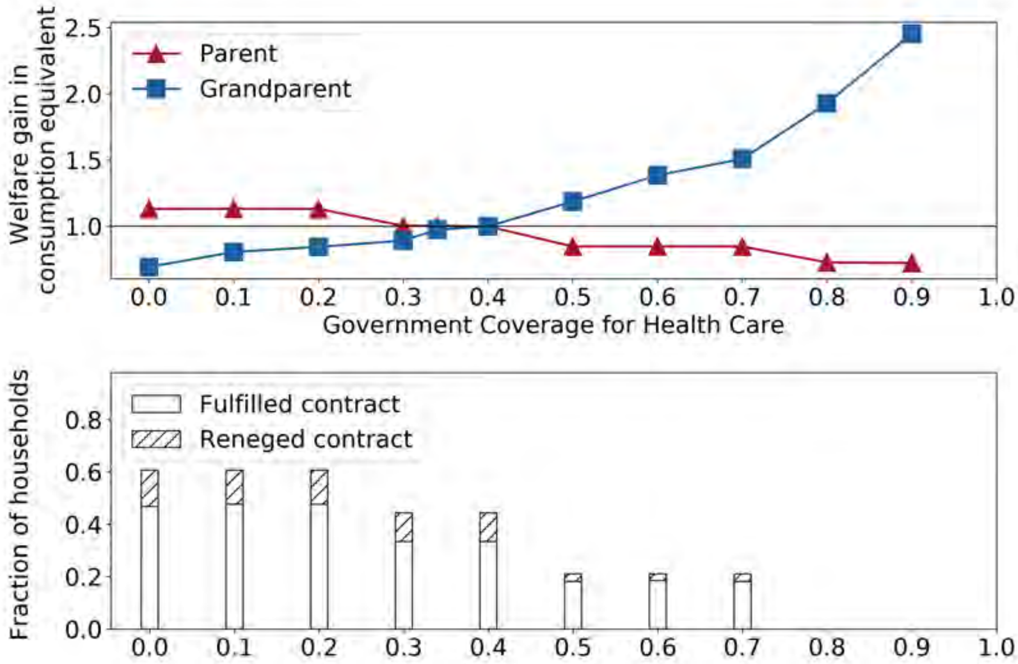


Figure 6: Public health policy effects on welfare and the household's contractual space

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

grandparents (Figure 9b in Appendix B). The children's education attainment increases because they live with their parents and perform better in school (Figure 5a).

7.3 Discussion of policy design

In this subsection, I compare the two policy counterfactuals tested in Section 7.1 and 7.2 to discuss the considerations in the design of policies targeting rural households with three generations.

The government wants to improve the welfare of the rural population while controlling the volume of migrants. The implications of the two policies for these objectives are distinct.

First, the policy effects on welfare distribution within rural households differ. Policy makers must watch out for welfare trade-offs when designing policies targeting one member in a three-generational household. The urban education policy only improves the welfare condition of the parents, and the effect is small (Figure 4). On the other hand, the trade-off appears in the public health policy counterfactual, which causes welfare gains to grandparents but losses to parents (Figure 6).

Second, while both policies target generations in rural households who are never part of the urban labor market, they affect the parent's migration decision through the informal contract. Policy makers should be aware of the policy effects on rural parents' migration decisions. The urban education policy increases the fraction of migrants with high school-aged children, and decreases the fraction of migrants with children over 18 years old (Figure 10 in Appendix B). Additionally, the expanded health policy separates grandparents from the parents' households; thus, parents must migrate with their children. These parents are more likely to migrate when their children are between 6 and 11 years old because urban tuition is cheap (Figure 11 in Appendix B).

8 Conclusion

This paper studies how internal migration in China affects inter-generational behavior in rural households, and how policies targeting migrants and their extended family members affect the welfare of each generation. I find that, in rural households with three generations, migrating parents and left-behind grandparents form an informal contract over child care, financial transfers, and elder care. Policies aiming to improve the welfare of one generation in the household affect the welfare of the other two generations, as well as the migration behavior of parents and children.

I first show empirical facts on migration and extended families in China using five micro datasets. The evidence indicates that most migrating parents leave children and grandparents behind. Grandparents look after children when they are healthy, and receive remittances from the parents. When grandparents are sick, many parents stay in the rural area to provide private care and pay for healthcare. I find a significant correlation among parent's decisions on migration, financial transfer, elder care and grandparent's health condition and decisions on childcare. Their decisions affect rural children's educational outcomes as well.

I develop a model of behavior of rural families with a migration option. The model features an informal contract with limited-commitment that involves these inter-generational behaviors. In my model, parents and grandparents play a sequential game as two independent decision makers. I structurally estimated the model's parameters. The model estimates suggest that poor households adopt the informal contract to improve welfare. Parents from wealthier households want to cooperate while grandparents do not because of potential welfare loss if parents renege on the contract.

I use the model to evaluate two hypothetical policies. An urban education policy that lowers urban tuition rates to rural levels promotes children's education, and encourages migration of both parents and children. The guilt associated with parents reneging on the informal contract secures the grandparent's welfare. Improved public health insurance coverage encourages more grandparents to spend on healthcare, while it causes a welfare loss to the parents by reducing the fraction of grandparents who are willing to accept the informal contract. Grandparents' staying independent of the parents leads more children to live with their parents. Their education attainment increases because parents have a positive influence on the children's school performances. The policy evaluation suggests that, when inter-generational behavior is accounted for, policies intended to improve the welfare condition of one left-behind subpopulation would change the welfare and migration behaviors of other family members.

My work points out that the government may use policies targeted at the left-behind family members to adjust the flow of rural-to-urban migrants. As of today, the government's only means of migration control is to make settling down in the cities more challenging, for instance, demolishing urban villages in which migrants aggregate and limiting the migrants' access to urban public services[49]. These methods caused intense protests advocating for the migrants' rights[50]. To that end, my counterfactual experiment shows that changing the policy environment of the migrants' family members could affect the costs and benefits of a parent's migration, and then lead rural parents to alter their migration decisions voluntarily.

In future work, I plan to study the intra-household bargaining between rural parents and grandparents in China in more depth. A more detailed model with heterogeneous informal contracts will allow me to test the effects of various public services on people with different age and household status. The model prediction of public service policy effects on migration will provide more comprehensive insights than current policy discussions have covered.

[49] Buckley, Chris. "Why Parts of Beijing Look Like a Devastated War Zone". The New York Times (2017).

[50] Kan, Karoline. "China's Migration Control Threatens Lives and Growth". Huffington Post (2017).

A Parameter definitions

Table 9 summarizes the parameters that enter the model and the estimation procedure.

Table 9: Parameter definitions

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

B Additional Figures

B.1 The value of intra-household cooperation

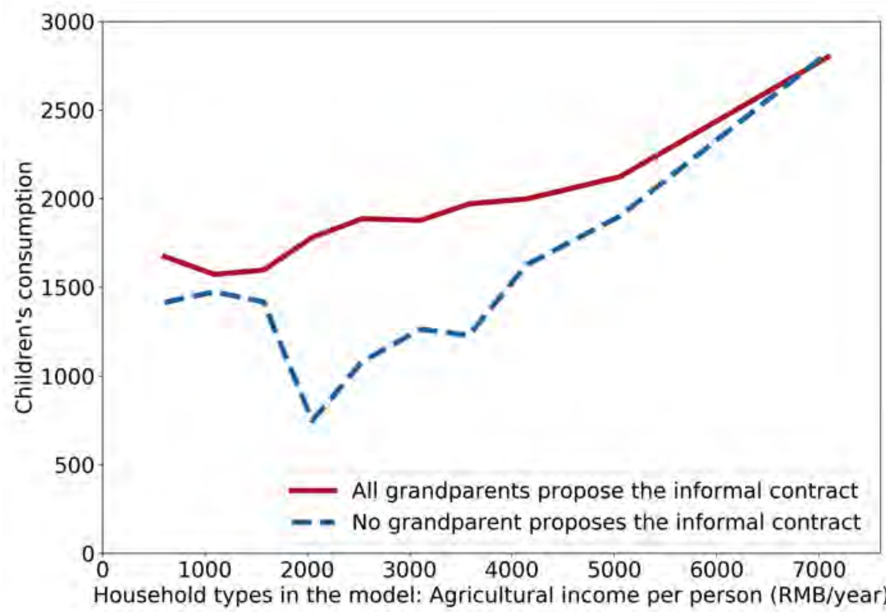
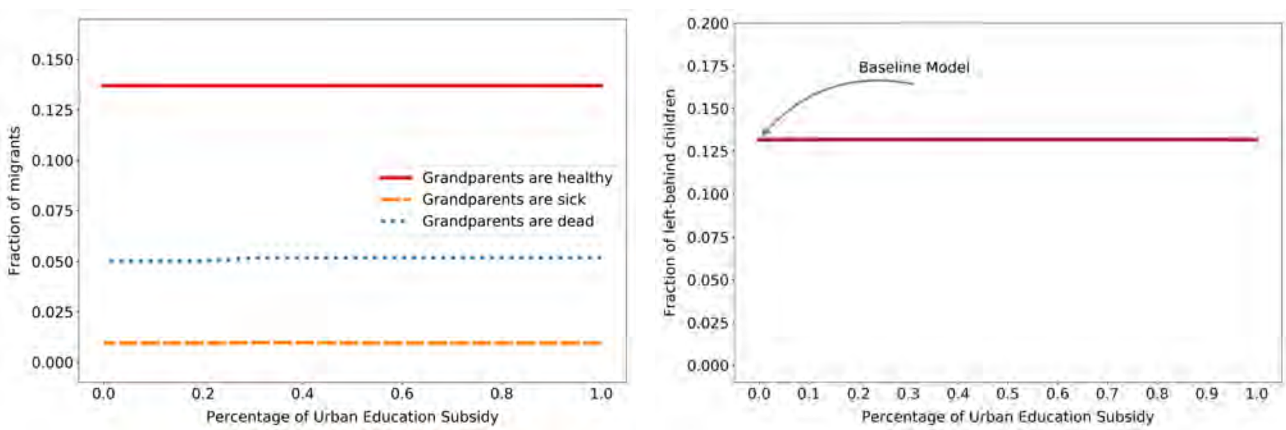


Figure 7: Intra-household cooperation increases children's consumption

Note: The above figure presents the consumption level of the children by their grandparents' informal contract status and their household's agricultural endowment level. 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.

B.2 Urban Education Policy



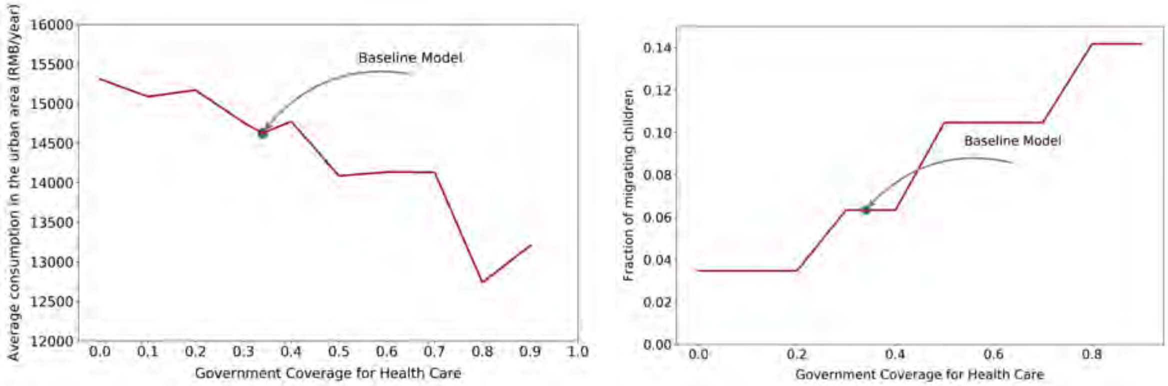
(a) Parents' migration unaffected

(b) Left-behind children problem unresolved

Figure 8: Effects of education policy on parent and child migration

Note: Panel (8a) shows the fraction of migrating parents by the health status of the grandparents. Panel (8b) shows the fraction of left-behind children among all rural children.

B.3 Expanded public healthcare coverage



(a) Parents' urban consumption falls

(b) More children migrate

Figure 9: Effects of public health policy on child migration and parent's urban consumption

Note: Panel (9a) presents the average of parents' consumption levels in the urban area. Panel (9b) shows the fraction of migrating children.

B.4 Discussion of policy design

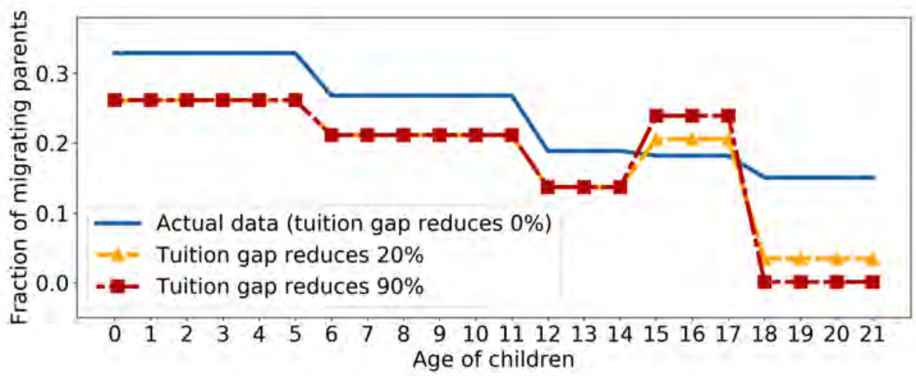


Figure 10: Education policy effect on migration by age of children

Note: The above figure presents the fractions of migrating parents by the age interval of their children with various urban education subsidy rates. The baseline model assumes zero subsidy. The orange line marked with triangles shows the age distribution of migrants when the urban-rural tuition gap is lowered by 20%. The red line marked with squares shows the age distribution of migrants when the urban-rural tuition gap is lowered by 90%. The figure shows that the urban education policy does not have a significant effect on rural parents' migration decisions.

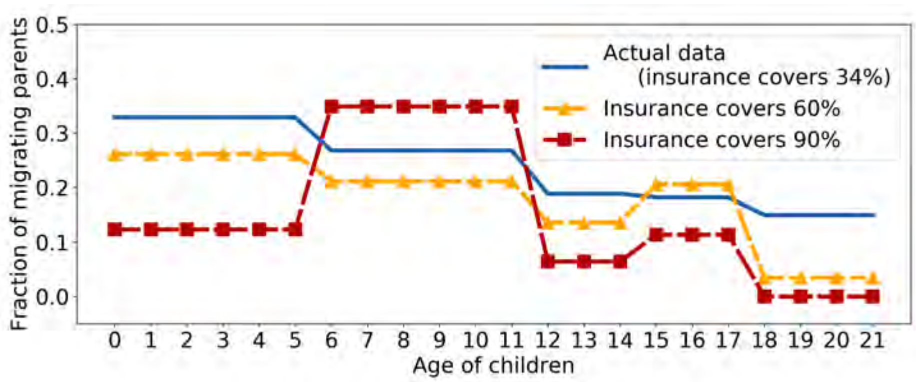


Figure 11: Public health policy effect on migration by age of children

Note: The above figure presents the fractions of migrating parents by the age interval of their children with various public health insurance coverage. The baseline model assumes 34% coverage. The orange line marked with triangles shows the age distribution of migrants when the government pays for 60% of the spending on healthcare. The red line marked with squares shows the age distribution of migrants when the government pays for 90% of the spending on healthcare. The figure shows that as the government pays more for the grandparents' healthcare, more parents migrate when their children are in primary school, and fewer parents migrate when the children are in other education periods.

C Data Appendix

This appendix explains the pre-processing of the data bases 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.

C.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[51]. 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.

C.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.

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

C.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 sick.
- 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 sick grandparents: whether each of the children is providing private care for the sick 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 sick.

D Empirical Evidence Appendix

D.1 Urban labor market for rural migrants

Wage rate does not increase by experience. Table 10 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[52]. 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 10: 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 12 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[53], 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.

D.2 Rural household expenditure

Table 11 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

[52] 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.

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

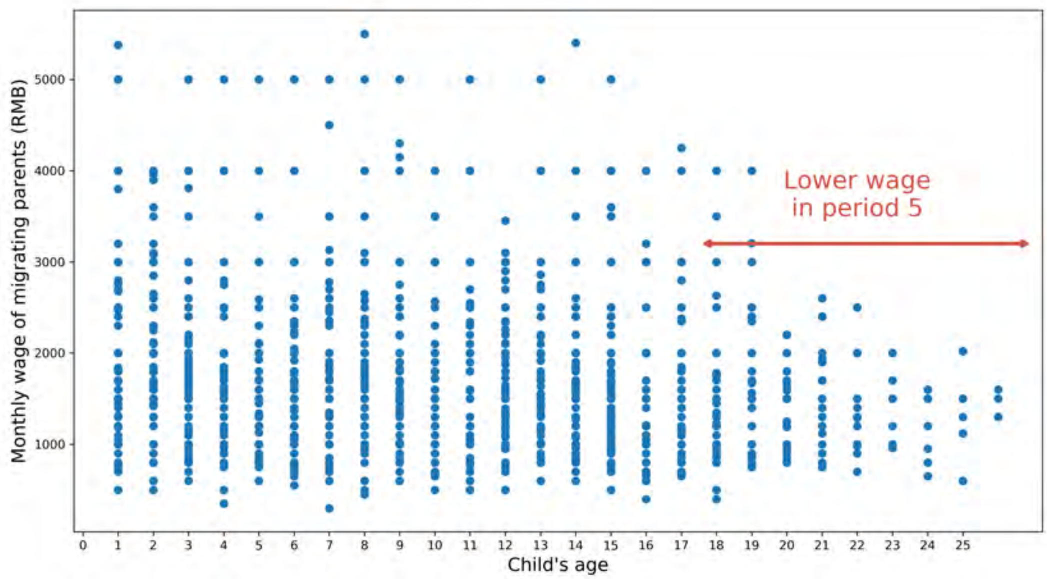


Figure 12: Distribution of migrating parent's wage by child's age

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 of 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 11: 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.

D.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 13. 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.

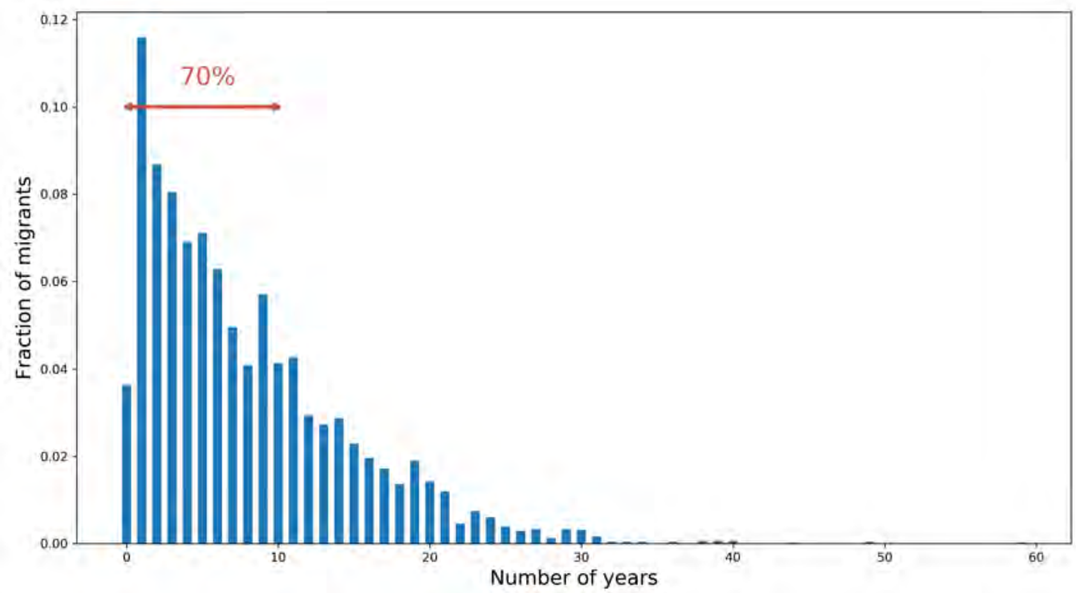


Figure 13: Length of stay in the urban area

D.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 12 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 follow 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 sick, 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 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.

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 14. 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

Table 12: 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

Table 13: OLS regression on migration behavior

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

Note: Standard errors are provided in parenthesis.

need private care are mainly young children and sick 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.

Table 14: 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.

D.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 15, the OLS regression on the amount of transfer, I further restrict the sample to households 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 of 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 sick.

Table 15: 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 sick	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.

D.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 consumptions, 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 16: 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 17 summarizes a selected set of papers on left-behind children in China.

Table 17: 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 18: 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

E Model Appendix

E.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.

E.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 $\omega = 4.14$.

E.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 19 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 19: Change in health status of rural people over 55 years old

	Health condition (two years before)	
Health condition (current)	Healthy	Sick
Healthy	0.3740	0.0276
Sick	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.

E.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 sick 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[54]. 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 sick elderly, and restrict to the sample in which the sick 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 personś 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).

[54] Roberts, Dexter. "Chinaś Rural Poor Bear the Brunt of the Nations Aging Crisis." Bloomberg (2017).

F Estimation Appendix

F.1 Externally estimated parameters

F.1.1 Distribution of agricultural income of rural households

Rural households have heterogeneous agricultural productivity. Appendix E.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 20: 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

F.1.2 Rural-urban price ratio

I use a rural-urban price ratio to standardize the daily consumption levels in the rural and urban area into comparable purchasing powers. 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 scatterplot smoothing (LOWESS) regression. The predicted Engel curves are presented in Figure 14.
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 (22) in footnote 31.

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.

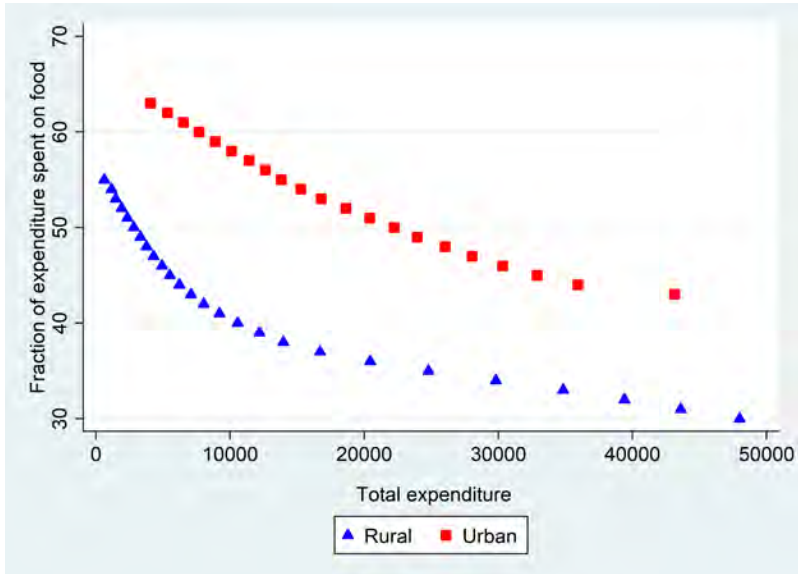


Figure 14: Estimated Engel curves

F.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.

F.1.4 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 21.

Table 21: Distribution of grandparent's health condition by child's age group

	Grandparent's health condition		
Child's age	Healthy	Sick	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.

- $Pr(\text{sick at } t+3 \mid \text{healthy at } t) = 0.15$
- $Pr(\text{dead at } t+3 \mid \text{sick at } t) = 0.65$

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 sick 3 year before. I assume that parent's migration is equivalent to the scenario in which parents do not provide elder care. Table 22 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 sick. 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 sick increases grandparent's mortality rate by 67%.

Table 22: 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

F.1.5 Education dropout probability

As stated in Section5.1.1, direct estimation of

$$Pr(dropout_t \mid tuition_t > 0, guardian, edu^{guardian}) \tag{27}$$

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 23
- $Pr(edu^{guardian} \mid guardian)$: Table 24

Table 23: Children's dropout probabilities by guardian's education attainment

Child's education level upon dropping out	Guardian's education				
	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 24: 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

The two sets of probabilities are presented in Tables 23 and 24. The matrix multiplication of these two marginal probability distribution results in the desired joint distribution.

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 $\frac{1.19}{4.41} = 0.2698$.

G Computation Appendix

G.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 $\left\{contract^G, \{X_t, choice_t^P, choice_t^G\}_{t=0}^4 | A_j, h_0\right\}$, 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 consumptions, 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 (19) - (21), 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

$$\begin{aligned} & Pr\left\{X_t, choice_t^P, choice_t^G, A_j, h_0\right\} \\ =& Pr\{contract^G, \left\{X_t, choice_t^P, choice_t^G\right\}_{t=1}^5 |A_j, h_0\} \\ & \times Pr(A_j) \times Pr(h_0) \times Pr(t) \end{aligned} \tag{28}$$

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

G.2 Computation algorithm

G.2.1 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 follow:

- 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 (29).

$$score(\vec{v_i}) = [(Q_0 - Q(\vec{v_i}))' W^{-1} (Q_0 - Q(\vec{v_i}))]^{-1} \tag{29}$$

It is the reciprocal of the objective function in equation (24). 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

$$pr(\vec{v_i} \in S_{k+1}) = \frac{score(\vec{v_i})}{\sum_{\vec{v_i} \in P_{k,nonelite}} score} \tag{30}$$

   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 $\vec{v_1\prime}$ and $\vec{v_2\prime}$ 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:

$$\begin{cases} \vec{v_1\prime}_j = \vec{v_2}_j \text{ and } \vec{v_2\prime}_j = \vec{v_1}_j & \text{with } \; pr_{crossover} \\ \vec{v_1\prime}_j = \vec{v_1}_j \text{ and } \vec{v_2\prime}_j = \vec{v_2}_j & \text{with } \; 1 - pr_{crossover} \end{cases} \tag{31}$$

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

      iii. Mutation on each of the two individuals. For each component in $\vec{v_i\prime}$, 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:

$$\begin{cases} \vec{v_1\prime\prime}_j \sim N(\vec{v_1\prime}_j, \sigma_j^2) & \text{with } \; pr_{mutation} \\ \vec{v_1\prime\prime}_j = \vec{v_1\prime}_j & \text{with } \; 1 - pr_{mutation} \end{cases} \tag{32}$$

      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 $\vec{v_1\prime\prime}$ and $\vec{v_2\prime\prime}$ 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

$$\max\{score(\vec{v_i})\}_{\vec{v_i} \in P_k} - \min\{score(\vec{v_i})\}_{\vec{v_i} \in P_k} < c \tag{33}$$

   where c is a constant tolerance parameter.

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

G.2.2 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, sick 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 15. It shows that the distributed system with careful allocation design (1) achieves full usage of all of the 720 CPUs most of the time and (2) eliminates most of the context switching overhead problems.

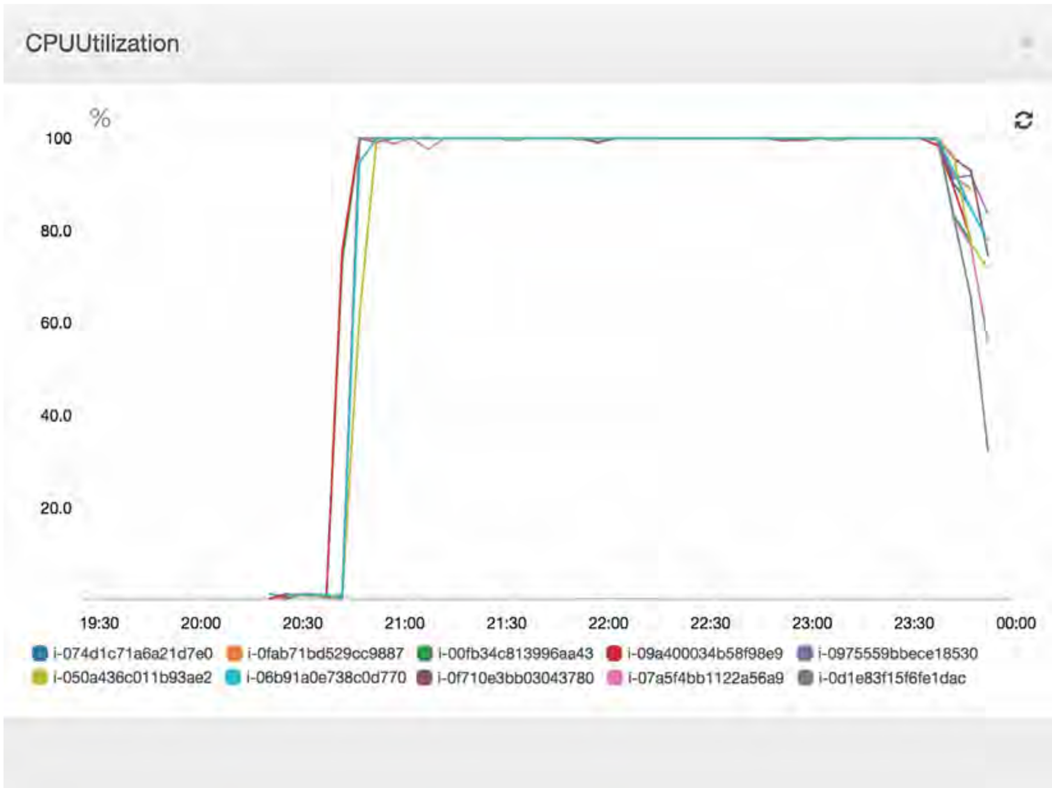


Figure 15: CPU usage of 10 72-core high performance computers controlled by a one-core machine on the Amazon Web Services

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.

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.

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.

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.

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.*

Greenwood, Michael J. 1997. "Internal migration in developed countries." *Handbook of population and family economics*, Volumn 1, 647–720.

Guriev, Sergei, and Elena Vakulenko. 2015. "Breaking out of poverty traps: Internal migration and interregional convergence in Russia." *Journal of Comparative Economics*, 43(3): 633–649.

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.

Hoddinott, John. 1994. "A model of migration and remittances applied to Western Kenya." *Oxford economic papers* 459–476.

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.

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.

Lucas, Robert EB. 1997. "Internal migration in developing countries." *Handbook of population and family economics*, Volumn 1, 721–798.

Luo, Baozhen. 2009. "The impact of rural-urban migration on familial elder care in rural China."

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. 2015. "Long-term care insurance and the family." *unpublished, November.*

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.

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

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.

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."

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.

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.

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.

Zhuo, Yue, and Zai Liang. 2015. "6. Migration and wellbeing of the elderly in rural China." *Handbook of Chinese Migration: Identity and Wellbeing*, p. 126.