ARC Centre of Excellence in Population Ageing Research

Working Paper 2019/10

Why is inequality higher among the old? Evidence from China

Katja Hanewald, Ruo Jia, and Zining Liu

Why is inequality higher among the old? Evidence from China*

Katja Hanewald, Ruo Jia, and Zining Liu†

August 2019

Abstract: This paper studies income inequality in old age and its development over the life cycle. We develop a theoretical framework and a new empirical method to show that income is more unequally distributed in old age than in working age. We combine the regression-based inequality decomposition method and the three-step mediating effect test to analyze the transmission of income inequality from initial socioeconomic differences to income inequality in old age. Our study is based on a panel of over 4,000 old households from the China Health and Nutrition Survey during 1991-2015. We find that the urban-rural gap and educational inequality are the primary causes of old-age income inequality. The effect of the urban-rural gap is partially mediated by educational inequality. Inequality accumulates with age and is reinforced in old age by the Chinese public pension system, which is fragmented by occupational sector.

Keywords: Inequality; Decomposition; Urban-rural gap; Pensions; China

JEL Classifications: H55; J26; O15; P36

* The authors acknowledge the financial support from the Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR). We are grateful for comments received from Martin Eling, Ming Gao, Bo Li, Can Wang, Yaojing Wang, Cheng Yuan, Wei Zheng, and participants at the Urbanization and Social Development Forum (Qingdao, June 2016), the 24th Annual Colloquium of Superannuation Researchers (Sydney, July 2016), the 7th China International Conference on Insurance and Risk Management (Xi'an, July 2016), the Social Welfare and Governance Forum (Hangzhou, May 2018), the seminar of the Modern Risk Society (Online, July 2018), the Labour and Health Economics Workshop, Peking University (Beijing, June 2019) and the 5th Annual Workshop on Population Ageing and the Chinese Economy (Sydney, July 2019). This study uses data from the China Health and Nutrition Survey (CHNS). We thank the National Institute for Nutrition and Health, China Center for Disease Control and Prevention, Carolina Population Center (P2C HD050924, T32 HD007168), the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, R24 HD050924, and R01-HD38700) and the NIH Fogarty International Center (D43 TW009077, D43 TW007709) for financial support for the CHNS data collection and analysis files from 1989 to 2015 and future surveys, and the China-Japan Friendship Hospital, Ministry of Health for support for the CHNS 2009, Chinese National Human Genome Center at Shanghai since 2009, and Beijing Municipal Center for Disease Prevention and Control since 2011.

† Katja Hanewald (k.hanewald@unsw.edu.au) is affiliated with the Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales. Ruo Jia (ruo.jia@pku.edu.cn) and Zining Liu (liuzining94@pku.edu.cn) are affiliated with the School of Economics, Peking University.

1. Introduction

Population aging and increasing inequality are two megatrends experienced in many countries. The worldwide population aged 60 and over (hereafter, the old) increased from 382 million (8.6%) in 1980 to 906 million (12.3%) in 2015 and is expected to increase to 1.4 billion (16.4%) in 2030 (United Nations, 2017a). In addition, income inequality has increased in almost all countries and regions in recent decades and particularly rapidly in North America, China, India, and Russia since 1980 (Alvaredo et al., 2018). Between 1980 and 2015, the top 10% income share rose from 35% to 47% in the U.S. and from 26% to 40% in China (Alvaredo et al., 2017). The countries where most old people live and where population aging occurs rapidly are also those most disadvantaged by inequality (United Nations, 2017b). Therefore, it is of great importance to study the inequality of the old and the transmission process causing this inequality.

On the one hand, inequalities accrue and become reinforced over an individual's life cycle because advantages and disadvantages associated with one's location, gender, socioeconomic status, and other characteristics accumulate by age (United Nations, 2017b; Chen et al., 2018). On the other hand, social security systems can pool risks and thus limit the growth of inequality over the life cycle, moderating inequality among the old compared with the young (Hurd and Shoven, 1985; Deaton et al., 2002). Understanding where the inequality among the old comes from and investigating how inequality changes from young to old is important for the design of public policies aimed at mitigating inequality to improve the quality of life of the old.

This paper studies income inequality in old age and analyzes the factors and transmission process causing this inequality. We develop an overlapping generation (OLG) model to describe the life-cycle evolution of inequality as a transmission process from initial socioeconomic differences to differences in educational attainment, then to wage inequality, and eventually to pension inequality in old age. To test the predictions of the model, we develop a new empirical method, which we refer to as the *inequality mediation decomposition* approach. We apply this method and conduct both repeated cross-sectional and longitudinal analyses to empirically identify the transmission process of income

inequality. Our analysis is based on data from the China Health and Nutrition Survey (CHNS) for the period 1991-2015. The CHNS is a high-quality household-level dataset covering a long period and has been widely used in development economics studies (e.g., Cao and Birchenall, 2013; Chamon et al., 2013; Santaeulalia-Llopis and Zheng, 2018).

China provides a persuasive context to analyze the income inequality of the old because it has experienced increasing inequality and rapid population aging in recent decades (Liao, 2013; Song et al., 2011; Alvaredo et al., 2017). China also attracts general interest because of the large size of its old population. By the end of 2017, 241 million Chinese were 60 or older, accounting for one quarter of the world's old population (United Nations, 2017a; Ministry of Civil Affairs of China, 2018). Moreover, China faces large urban-rural disparities in development: urban residents receive higher incomes, better social welfare, and better education than rural residents (Cao and Birchenall, 2013). Our results for China are informative for other developing countries that face large urban-rural gaps, rapidly aging populations, fragmented pension systems, and increasing income inequality.

There is limited research that theoretically models and/or empirically identifies the transmission process of income inequality. Young (2013) estimates that the urban-rural gap accounts for 40% of inequality using a dataset of 65 countries. Molero-Simarro (2017) documents that the increasing urban-rural gap, especially the rise in the income share of the top 10% of urban households, explains the overall increase in inequality in China. Both studies focus on the direct impact of the urban-rural gap on income inequality. Burzynski et al. (2019) develop a dynamic model to investigate the inequality between urban and rural regions within one country and the inequality among countries. One key dimension of the urban-rural gap is unequal access to education (Liu, 2005; Golley and Kong, 2018; Burzynski et al., 2019). Empirical evidence shows that reducing the inequality of educational attainment is key to reducing income inequality and poverty (Gregorio and Lee, 2002; Knight et al., 2010).

Our study makes three main contributions to the existing literature. First, to the best of our knowledge, we are the first to theoretically model and empirically identify the transmission process from a cohort's socioeconomic origins to income inequality during working age and

to greater income inequality in old age. Our model allows the quality of education to vary by initial rural-urban status and describes how this educational inequality transmits to wage inequality during working age and pension inequality in old age. We empirically confirm this transmission process over the life cycle across the 25-year sample period and across different population cohorts. We confirm that education is a mediator in the transmission process and show that fragmented public pension programs reinforce inequality from working age to old age. The model, together with the empirical results, provides a microeconomic foundation to explain the higher levels of inequality among the old than among the young.

Our second contribution is a novel approach to empirically identify the transmission process of inequality over the life cycle. Our new approach combines the regression-based inequality decomposition method (Shorrocks, 1982; Fields, 1998; Bourguignon et al., 2001; Morduch and Sicular, 2002; Wan, 2004) and the three-step mediating effect test (Baron and Kenny, 1986). We show how the new approach, which we refer to as *inequality mediation decomposition*, can be applied to model and quantify the multiple steps in an inequality transmission process. The new method is a useful tool to study mechanisms and contribution channels in inequality studies.

Third, we provide new evidence quantifying the level of inequality among the old in China and the drivers of this inequality. Most previous studies on China focus on the inequality of the entire population, including both old and young households (e.g., Li and Sicular, 2014; Ding and He, 2018; Santaeulalia-Llopis and Zheng, 2018). Chen et al. (2018) analyze the distributional effects of population aging on inequality and conclude that population aging has the overall effect of aggravating inequality. We focus on the inequality among the rapidly growing old population and analyze the transmission process that makes old-age inequality larger than the inequality among the young. Only a few studies have examined the drivers of income inequality among the old in China. These studies have analyzed the impact of specific socioeconomic disparities on inequality among the old, such as gender differences (Saunders, 2007; Zhao and Zhao, 2018), differences in living arrangements (Park et al., 2012; Connelly and Maurer-Fazio, 2016), and urban-rural disparities (Wang et al., 2014). We are the first to compare the contributions of different socioeconomic characteristics and to identify the

primary contributors to old-age income inequality in China.

We find that the income inequality of old households was higher than that of young households in all survey years during the sample period 1991-2015. The urban-rural gap, capturing one's socioeconomic origin, and educational attainment are the primary socioeconomic factors explaining the income inequality among old households. In recent years (2000-2015), these two factors explain 11%-22% of the overall old-age income inequality. Approximately one-third to half of the contribution of the urban-rural gap to old-age income inequality was mediated by educational inequality between rural and urban areas. The urban-rural gap and educational inequality mainly contribute to old-age income inequality by contributing to pension inequality, which accounts for 45%-60% of the old-age income inequality and explains almost all of the inequality increase from 2000 to 2015 for the household cohort that was young in 2000 and became old in 2015. We empirically show that old-age income inequality is reinforced by the pension system, which is fragmented by occupational sector.

Our results suggest that reforms of the public resource allocation for social security and education can reduce inequality. Future pension reforms need to balance the trade-off between the adequacy of one pension program and the equality across different pension programs. The current practice to increase pension payments by a fixed percentage for all pensioners (e.g., 5% every year) increases income adequacy but also increases income inequality among the old. Long-term strategies to reduce old-age income inequality in China include improving access to and the quality of education in rural areas, as this would improve the labor market opportunities and wages of rural residents, allowing them to increase their pension savings. These reforms would require a shift in the structure of government spending away from investments aimed at economic growth to income redistribution and improving welfare. Our conclusions and policy implications are also relevant for other countries with a fragmented pension system and large inequality among the old, such as Mexico and India (World Bank, 2005).

The remainder of the paper is organized as follows. Section 2 reviews the literature and develops the theoretical framework. Section 3 describes the empirical methodology to

decompose income inequality. Section 4 introduces the CHNS dataset and our samples. Section 5 reports the results from repeated cross-sectional and longitudinal analyses. Section 6 presents eight robustness tests. Section 7 summarizes the conclusions and policy implications.

2. Literature review and theoretical framework

2.1. Income inequality among the old and the young

There are different hypotheses in the literature regarding the evolution of income inequality over a population cohort's life cycle. Some theories suggest that income is more unequally distributed among the old than among the young, some predict the opposite, and others argue that the two inequalities should be similar.[1]

Both the permanent income hypothesis (Friedman, 1957) and the life cycle hypothesis of saving (Modigliani, 1966) imply that the income inequality within a population cohort increases as the cohort ages. Intuitively, an individual's income is determined by his/her history of educational attainment, employment, living arrangements, and other characteristics (Chen et al., 2018). These qualitative differences, advantages, and disadvantages accumulate during working life so that income inequality increases with age. Standard human capital models also predict increasing income inequality over the life cycle considering that the rate of return to human capital differs by individual, and different individuals choose different levels of education and professional training (e.g., Mincer, 1974). Therefore, the income differences among individuals should increase with different education and professional training.

Counter arguments suggest that the public transfers embedded in social security programs are often designed to offset an individual's cumulative (dis)advantage by redistributing resources from the more advantaged to the less advantaged (Radner, 1982; Hurd and Shoven,

[1] A related but separate strand of literature analyzes the impact of population aging (i.e., the change in the population age structure) on the aggregate inequality of an economy (e.g., Chu and Jiang, 1997; Zhong, 2011; Luo et al., 2018; Dong et al., 2018). These studies usually assess the inequality of the entire population, while we focus on the inequality among the old. Another strand of literature studies the development of inequality across generations (i.e., cohort effects) or during different time periods (i.e., time effects) (e.g., Tsakloglou, 1993; Heathcote et al., 2005). Our research focuses on the age effects on income inequality.

1985; Crystal and Shea, 1990; Crystal and Waehrer, 1996; Deaton et al., 2002). Additionally, older individuals rely less on income from work, which is usually a key source of income inequality. Based on these arguments, the income inequality of the old should be smaller than that of the young.

A third hypothesis argues that income inequality remains stable with age because similar factors determine income before and after retirement (Henretta and Campbel, 1976). This inequality maintenance hypothesis assumes that the redistribution effect of the retirement system offsets the accumulated inequality in old age.

The empirical literature comparing the income inequality among the old with that among the young provides mixed results. Several previous studies find that the income of the old is more unequally distributed than that of the young (Deaton and Paxson, 1994, for the U.S., U.K., and Taiwan; Deaton and Paxson, 1997, for Thailand; Bönke et al., 2010, for Germany; Chen et al., 2018, for mainland China). Some studies find the opposite (Prus, 2000, for Canada) or similar levels of inequality among the old and the young (Coder et al., 1989, for some OECD countries). We analyze the factors and transmission process causing income inequality in old age.

2.2. Transmission process from socioeconomic origin to income inequality

To analyze the income inequality among the old and explain how it evolves over the life cycle, we conceptualize a transmission process from initial socioeconomic differences, through the life course, to income inequality in old age.

We argue that income inequality in China originates from urban-rural disparities, which are often considered the primary cause of inequality in developing economies. In China, urban-rural disparities are especially large due to the *hukou* system.[2] Previous research shows that the urban-rural separated *hukou* system is the main source of educational inequality in China (Liu, 2005; Golley and Kong, 2018). Previous studies also show that educational inequality affects the income distribution (Gregorio and Lee, 2002) and is an important cause of income inequality between rural and urban residents (Zhang and Xu, 2016). Therefore, we argue that

[2] The *hukou* system is a household registration system that determines where individuals can receive education and other public services. One's initial hukou status is largely determined by his/her parents' hukou.

the urban-rural gap and educational inequality both have a standalone direct effect on income inequality. In addition, we argue that educational attainment mediates the relationship between the urban-rural gap and income inequality.

Old households have two main income sources: pensions and income from work. We argue that pensions in China are more strongly determined by one's socioeconomic origin and educational attainment than by income from work. This is because an individual's occupational sector is highly related to his/her socioeconomic origin and educational attainment, and an individual's occupational sector determines more of his/her public pension benefits than income from work in China (Zheng et al., 2019). Individuals with urban hukou and more years of formal education are more likely to work in the formal sector and to participate in the high-benefit Employees' Basic Pension Program. Individuals with rural hukou and less education are more likely to work in the informal sector and to participate in the low-benefit Residents' Basic Pension Program.[3] For example, a rural middle school graduate may have a good income from work by running a household business but is likely to have the low-benefit Residents' Basic Pension Program. The adherence to pension programs is not easy to change by migration from rural to urban areas during working life and in old age. As a result, the income of the old, which consists of both pensions and income from work, is likely to be more unequally distributed than the income of the young, who rely primarily on income from work.

2.3. A stylized OLG model

To formalize our conceptual framework and derive the hypotheses to be empirically tested, we develop a theoretical model in the framework of overlapping generations (Samuelson, 1958; Acemoglu, 2009) to analyze the transmission from the urban-rural gap to income inequality in young age and in old age. The model describes how income inequality

[3] China's public pension system favors employees in the formal sector, who participate in the Employees' Basic Pension Program, and disadvantages workers in the informal sector, who participate in the Residents' Basic Pension Program (Wang et al., 2014; Zheng et al., 2019). The fragmented Chinese public pension system is partially a historical legacy. The Chinese government shifted the subsidies for urban formal sector employees from work related-welfare programs (e.g., food coupons, housing) to high-benefit pension programs (Santaeulalia-Llopis and Zheng, 2018). Informal sector workers, who are mostly rural residents, historically received low public subsidies and are still participating in the recently established low-benefit Residents' Basic Pension Program.

accumulates over the life cycle and how differences in education and human capital result in income inequality during working life and in old age. Under assumptions that reflect the environment of a developing economy, the model predicts that the initial educational disadvantages of rural residents lead to lower income over their lifetime than urban residents. The urban-rural income inequality increases as the agents age, and thus, the income inequality is greater in old age than in young age.

The model focuses on the optimization problem of a representative individual over a life cycle with three periods: childhood, working age, and old age. The economy is populated by a constant number of homogeneous individuals who are born in different periods. Time *t* is discrete, and the time horizon is infinite with overlapping generations. We use a double subscript to indicate the birth period followed by the current time period.

During childhood, an individual born at the beginning of period *t* consumes $C_{t,t}$ and invests $K_{t,t}$ in education in period *t*. The child's consumption and educational investment are funded and determined by his/her parent, who was born in period *t*-1 and makes a transfer $T_{t-1,t}$ to the child in period *t* ($C_{t,t} + K_{t,t} \leq T_{t-1,t}$).[4] At the end of period *t*, the individual acquires human capital $H_{t,t+1} = K_{t,t} \times H_{t,t}$ from the educational investment $K_{t,t}$, where $H_{t,t}$ is an endowment, capturing the marginal return of educational investment. In the context of this paper, we consider $H_{t,t}$ to capture the quality of education that the child can access, and we allow this parameter to differ between urban and rural areas (see below).

During working age, the individual works and raises a child. He/she supplies human capital ($H_{t,t+1}$) and earns a wage income ($I_{t,t+1} = W_{t+1}H_{t,t+1}$). The wage rate ($W_{t+1}$) equals the marginal return of human capital and can therefore be considered the equilibrium wage rate. The individual consumes $C_{t,t+1}$, invests $K_{t,t+1}$ in assets, which earn a real gross return rate of R_{t+1} in the next period, and transfers $T_{t,t+1}$ to his/her child. The sum of the individual's consumption, investment, and transfers is subject to the constraint $C_{t,t+1} + K_{t,t+1} + T_{t,t+1} \leq W_{t+1}H_{t,t+1}$. Working-age individuals also make transfers to the retired generation in the form of tax contributions to a pay-as-you-go pension program. We assume

[4] As population size does not change, one pair of parents has two children, and each individual has one child.

that the representative individual owns the production sector and employs himself/herself to produce *Y*. A part of *Y* is distributed as wage income, and the rest contributes to the pension system. Thus, the amount of pension transfer is defined as the difference between total production and wage income $Y_{t+1} - W_{t+1}H_{t,t+1}$.

In old age, the individual consumes $C_{t,t+2}$, subject to the constraint $C_{t,t+2} \leq R_{t+1}K_{t,t+1} + \left(\mathrm{Y}_{t+2} - W_{t+2}H_{t+1,t+2}\right)$, where $R_{t+1}K_{t,t+1}$ is the asset investment income and $\mathrm{Y}_{t+2} - W_{t+2}H_{t+1,t+2}$ is the pension income.

We assume the following production function: $Y_{t+1} = A_{t+1}{H_{t,t+1}}^{\alpha}$ with $0<\alpha<1$. A_t denotes the technology (or total factor productivity). The individual's wage rate is thus $W_{t+1} = \alpha A_{t+1}{H_{t,t+1}}^{\alpha-1}$.

The individual maximizes his/her lifetime utility over the three-period consumption stream, as shown in the optimization problem (1), where β is the subjective discount factor ($0<\beta<1$). In childhood and old age, the individual cares only about his/her own consumption. During working age, he/she also cares about his/her child's consumption, which is scaled by a factor $\theta>0$, capturing how much the individual cares about his/her child's utility (Andreoni, 1989). The individual chooses his/her own consumption $C_{t,t+1}, C_{t,t+2}$, his/her asset investment $K_{t,t+1}$, the transfer to his/her child $T_{t,t+1}$, and the child's consumption and educational investment $C_{t+1,t+1}, K_{t+1,t+1}$, to maximize his/her lifetime utility, subject to the budget constraints in the three periods (inequalities (2)-(4)).

$$\max_{C_{t,t+1},K_{t,t+1},C_{t,t+2},T_{t,t+1},C_{t+1,t+1},K_{t+1,t+1}} \left\{U\left(C_{t,t}\right) + \beta\left[U\left(C_{t,t+1}\right) + \theta U\left(C_{t+1,t+1}\right)\right] + \beta^2 U(C_{t,t+2})\right\} \quad (1)$$

s.t. $C_{t+1,t+1} + K_{t+1,t+1} \leq T_{t,t+1}$ (2)

$C_{t,t+1} + K_{t,t+1} + T_{t,t+1} \leq W_{t+1}H_{t,t+1}$ (3)

$C_{t,t+2} \leq R_{t+1}K_{t,t+1} + Y_{t+2} - W_{t+2}H_{t+1,t+2}$ (4)

We operationalize the maximization problem by assuming a natural logarithm utility function, which belongs to the family of constant relative risk aversion utility functions. We solve the maximization problem using the Lagrangian function to derive the individual's consumption ($C_{t,t+1}$) and income ($I_{t,t+1}$) during working age as shown in Equations (5) and (6), where $\varphi = [\alpha(1-\alpha)]^{\frac{1}{1-\alpha}}$. At any time *t*, the technology (A_t), gross return rate on asset

investment (R_t), and quality of education ($H_{t,t}$) are exogenously determined and given in the economy (Acemoglu, 2009).

$$C_{t,t+1} = \frac{\alpha\varphi^{\alpha}A_{t+1}^{\frac{1}{1-\alpha}}H_{t,t}^{\frac{\alpha}{1-\alpha}}R_t^{\frac{-\alpha}{1-\alpha}} + \frac{1-\alpha}{\alpha}\varphi A_{t+2}^{\frac{1}{1-\alpha}}H_{t+1,t+1}^{\frac{\alpha}{1-\alpha}}R_{t+1}^{\frac{-1}{1-\alpha}}}{1+\beta+\theta} \quad (5)$$

$$I_{t,t+1} = W_{t+1}H_{t,t+1} = \alpha\varphi^{\alpha}A_{t+1}^{\frac{1}{1-\alpha}}H_{t,t}^{\frac{\alpha}{1-\alpha}}R_t^{\frac{-\alpha}{1-\alpha}} \quad (6)$$

Next, we extend the baseline model of homogeneous individuals to allow for two types of representative individuals: urban and rural residents. The quality of education is assumed to be higher in urban areas than in rural areas ($H_{t,t}^{Urban} > H_{t,t}^{Rural}, \forall\, t$) due to differences in public resources. Since $C_{t,t+1}$ and $I_{t,t+1}$ are positively correlated with $H_{t,t}$, we have $C_{t,t+1}^{Urban} > C_{t,t+1}^{Rural}$ and $I_{t,t+1}^{Urban} > I_{t,t+1}^{Rural}$, assuming for now that the technology (A_t) and gross investment return rate (R_t) are the same between urban and rural areas.

In addition, we obtain the retirement replacement ratios for consumption and income, as shown in Euler Equation (7) and Equation (8). We define $\delta_t = \frac{\beta(1-\alpha)^2 a_{t+1}^{\frac{1}{1-\alpha}} h_t^{\frac{\alpha}{1-\alpha}} r_t^{\frac{-\alpha}{1-\alpha}}}{\alpha(1+\beta+\theta)}$, with $a_{t+1} = \frac{A_{t+2}}{A_{t+1}}, h_t = \frac{H_{t+1}}{H_t}, r_t = \frac{R_{t+1}}{R_t}$. By definition, $\delta_t > 0$. For any *t*, the variables describing economic development (i.e., a_t, h_t, and r_t) are assumed to be the same in rural and urban areas. Thus, both replacement ratios are positive, independent of $H_{t,t}$, and exogenously determined according to the economy. Therefore, both ratios are the same in rural and urban areas in any period *t*. We thus have the consumption and income in old age following the same pattern as in working age, that is, $C_{t,t+2}^{Urban} > C_{t,t+2}^{Rural}$ and $I_{t,t+2}^{Urban} > I_{t,t+2}^{Rural}$.

$$\frac{C_{t,t+2}}{C_{t,t+1}} = \beta R_{t+1} \quad (7)$$

$$\frac{I_{t,t+2}}{I_{t,t+1}} = \frac{\beta}{1+\beta+\theta} R_{t+1} + \delta_t \quad (8)$$

Next, we relax the assumption that the real gross investment return (R_t) is the same between urban and rural areas. Developing economies often have a dual urban-rural structure, and residents in rural areas are likely to have less developed financial markets than those in urban areas. Rural residents therefore face stricter credit constraints and less access to high-return investment opportunities (Galor and Moav, 2004; Coibion et al., 2012). Therefore, we

assume in the following that urban residents have a higher gross investment return rate than rural residents ($R_t^{Urban} > R_t^{Rural}, \forall\, t$).

We now compare the income inequality during working age with that in old age. From Equation (8), we have $\frac{I_{t,t+2}^{Urban}}{I_{t,t+1}^{Urban}} = \frac{\beta}{1+\beta+\theta} R_{t+1}^{Urban} + \delta_t$ and $\frac{I_{t,t+2}^{Rural}}{I_{t,t+1}^{Rural}} = \frac{\beta}{1+\beta+\theta} R_{t+1}^{Rural} + \delta_t$. Given the assumption regarding the gross investment return rate $R_t^{Urban} > R_t^{Rural}$, we have $\frac{I_{t,t+2}^{Urban}}{I_{t,t+1}^{Urban}} > \frac{I_{t,t+2}^{Rural}}{I_{t,t+1}^{Rural}}$, and thus $\frac{I_{t,t+2}^{Urban}}{I_{t,t+2}^{Rural}} > \frac{I_{t,t+1}^{Urban}}{I_{t,t+1}^{Rural}}$.[5] We therefore derive our Hypothesis 1 below.

- *Hypothesis 1: The income inequality is larger among the old than among working-age adults in a developing economy environment.*

Based on Equation (6), the ratio of urban and rural income in working age can be expressed as $\frac{I_{t,t+1}^{Urban}}{I_{t,t+1}^{Rural}} = h\left(\frac{H_{t,t}^{Urban}}{H_{t,t}^{Rural}}\right)$ ($h' > 0$). Similarly, from Equations (6) and (8), the ratio of urban and rural income in old age can be expressed as $\frac{I_{t,t+2}^{Urban}}{I_{t,t+2}^{Rural}} = g\left(\frac{H_{t,t}^{Urban}}{H_{t,t}^{Rural}}\right)$ ($g' > 0$). That is, the income inequality between urban and rural individuals during working age and in old age is driven by the difference in the quality of education between urban and rural areas. The larger the educational gap is, the greater the income inequality. We therefore derive Hypothesis 2 below.[6]

- *Hypothesis 2: The income inequality during working age and in old age is driven by the difference in education between urban and rural areas.*

3. Empirical methodology

We empirically test these hypotheses and the inequality transmission process described in Section 2.2 by decomposing income inequality into its socioeconomic determinants and income

[5] To ensure that urban income is larger than rural income in working age and in old age, i.e., $I_{t,t+1}^{Urban}/I_{t,t+1}^{Rural} > 1$ and $I_{t,t+2}^{Urban}/I_{t,t+2}^{Rural} > 1$, a sufficient and necessary condition is required that $H_{t,t}^{Urban}/H_{t,t}^{Rural} > R_t^{Urban}/R_t^{Rural}$, which states that the urban-rural gap in the marginal return of educational investment is larger than the gap in the marginal return of asset investment. This condition is likely to hold in less-developed economies, where human capital is more critical to inequality and economic growth than physical capital (Galor and Moav, 2004).

[6] We can extend our model to allow for migration from rural to urban areas. We assume that a fraction of m ($0<m<1$) rural residents migrate to urban areas and earn an urban income. Therefore, the income inequality between the original urban and rural residents in working age becomes $I_{t,t+1}^{Urban}/\left[(1-m)I_{t,t+1}^{Rural} + mI_{t,t+1}^{Urban}\right]$, which is smaller than the income inequality $I_{t,t+1}^{Urban}/I_{t,t+1}^{Rural}$ in our main model without migration. Similarly, the income inequality in old age becomes $I_{t,t+2}^{Urban}/\left[(1-m)I_{t,t+2}^{Rural} + mI_{t,t+2}^{Urban}\right]$, which is smaller than $I_{t,t+2}^{Urban}/I_{t,t+2}^{Rural}$. We can prove that both hypotheses remain true when we allow for migration. The proof is available from the authors upon request. We also consider the migration between rural and urban areas in our empirical analyses.

components and by comparing the decomposition results for the old and the young.

3.1. Income inequality measures

To measure income inequality,[7] we estimate the Gini coefficient, the Theil index, and the P90/P10 ratio of the real net household income per capita (hereafter, household income per capita) for each survey year in our 1991-2015 sample. Equations (9) and (10) define the Gini coefficient and the Theil index, respectively. I_n is the household income per capita of household *n*. I is a vector consisting of all I_n ranked in nondecreasing order. N is the number of households in a given year. $\bar{I}$ is the mean of household income per capita in a given year. $R_n = \frac{2n-1}{2N}$ is the fractional rank of household *n*'s per capita income, where *n* represents the rank of household *n* according to per capita income. The P90/P10 ratio is the ratio between the 90th and the 10th quantile of the distribution of I_n.

$$\text{Gini coefficient } (I) = \frac{2}{N}\sum_{n=1}^{N}\frac{R_n I_n}{\bar{I}} - 1 \tag{9}$$

$$\text{Theil index } (I) = \frac{1}{N}\sum_{n=1}^{N}\frac{I_n}{\bar{I}}\left(\ln\left(\frac{I_n}{\bar{I}}\right)\right) \tag{10}$$

To analyze the income inequality among the old and to compare it with that among the young, we estimate the Gini coefficient, the Theil index, and the P90/P10 ratio separately for old and young households. We compare the estimated values in each year. We also compare the average values over all years for old households and young households using the mean difference *t*-test.

3.2. Decomposition of the Gini coefficient

To empirically examine the two hypotheses and the transmission process described in Section 2, we conduct three decompositions. First, we decompose the inequality of household income per capita by socioeconomic characteristics that determine income (Morduch and Sicular, 2002; Wan, 2004) (Decomposition (1)). Second, we decompose the inequality of

[7] Some economists analyze consumption inequality instead of income inequality considering that households can smooth their consumption by borrowing (Meyer and Sullivan, 2013). Empirical evidence shows that inequality of consumption and inequality of income in China are strongly correlated and closely track each other both over time and over the life cycle (Cai et al., 2010; Ding and He, 2018), perhaps because of the limited access of households to credit and financial markets (Li and Sicular, 2014). Therefore, we focus on income inequality in the empirical part of the paper. Our theoretical model can also be used to study consumption inequality. The results are available from the authors upon request.

household income per capita into income components (Shorrocks, 1982; Lerman and Yitzhaki, 1985) (Decomposition (2)). Third, we decompose the inequality of the income component that has the largest contribution to overall income inequality (pensions for old households and income from work for young households) into its socioeconomic determinants (Decomposition (3)). We compare the decomposition results for old households and young households to explain the income inequality differences between the old and the young.

We apply the method developed by Shorrocks (1982) and Lerman and Yitzhaki (1985) to decompose the Gini coefficient of total income into contributions of several income components, as shown in Equation (11):

$$\text{Gini coefficient } (I) = \sum_{m=1}^{M} \left(\frac{\bar{\imath}_m}{\bar{I}}\right) C_m = \sum_{m=1}^{M} S_m C_m \tag{11}$$

We use $i_{1n}, i_{2n}, \ldots, i_{Mn}$ to denote the income components of household n and $I_1, I_2, \ldots, I_M$ to denote the vectors of the respective income components ranked by the total income of each household I_n. The total income I is the sum of all income component vectors. $\bar{\imath}_m = \frac{1}{N}\sum_{n=1}^{N} i_{mn}$ is the average of income component m over all households N. S_m is income component m's share of total income. C_m is called the concentration index, capturing the inequalities of income component i_m ranked by total income I, with $C_m = \frac{2}{N}\sum_{n=1}^{N} \frac{R_n i_{mn}}{\bar{\imath}_m} - 1$.[8]

The contribution to the overall income inequality of income component m is the concentration index C_m times the share S_m of income component m of total income. The sum of the contributions of all income components equals the Gini coefficient of total income. The contribution of income component m can also be presented as a percentage of the overall income inequality, i.e., m's contribution divided by the Gini coefficient of total income. The Gini coefficient decomposition method given by Equation (11) is applicable only to decompose income inequality into contributions of the income components because the

[8] C_m can be viewed as the Gini coefficient of income component m multiplied by the "Gini correlation" between income component m and total income I. The Gini correlation represents the correlation between the income component m and total income. For example, when the income component m is a monotonically increasing (decreasing) function of total income, which means income component m has the same (opposite) rank of total income I, then the Gini correlation equals +1 (-1).

method requires that the sum of the decomposed elements equal the object to be decomposed.

To decompose income inequality into its socioeconomic determinants, we use a regression-based approach (Fields, 1998; Morduch and Sicular, 2002; Wan, 2004). The idea to decompose inequality into its socioeconomic determinants (or population groups) dates back to Oaxaca (1973) and Shorrocks (1980). Since then, there has been continuous interest to improve the method (see, e.g., Fields, 1998; Bourguignon et al., 2001; Morduch and Sicular, 2002; Wan, 2004). Fields (1998) developed a strategy to decompose the variance of log income into comparable and additive factor components. Morduch and Sicular (2002) and Wan (2004) improved Fields' (1998) method to use a linear regression to decompose inequality measured by the Gini coefficient into the contribution of its determinants. The regression-based decomposition gauges the contribution of one independent variable to inequality while controlling for the effects of other covariates and enables the comparison among contributors. The regression-based decomposition approach has been widely used to decompose inequality, for example, in Europe (Doorslaer and Koolman, 2004), Africa (Lambert et al., 2014), and Asia (Fields and Yoo, 2000; Wagstaff et al., 2003; Wan and Zhou, 2005; Zhong, 2011).

To apply the regression-based approach, we first specify a regression, as shown in Equation (12), with household income per capita as the dependent variable and its socioeconomic determinants as independent variables, including the household head's *hukou* status (rural or urban), gender, age, and completed years of formal education, as well as the household size and living arrangement (see Section 4 for more details). We also include a set of binary variables indicating the province in which a household resides. We apply the regression-based approach separately for each survey year.

$$I_n = \alpha + \sum_{k=1}^{K} \beta_k x_{kn} + e_n \tag{12}$$

Second, we substitute Equation (12) into the definition of the Gini coefficient in Equation (9). The resulting Equation (13) decomposes the Gini coefficient by its socioeconomic determinants:

$$\text{Gini coefficient } (I) = \frac{2}{N}\sum_{n=1}^{N}\frac{R_n}{\bar{I}}\left(\alpha + \sum_{k=1}^{K}\beta_k x_{kn} + e_n\right) - 1 \quad (13)$$

We define the concentration index $C_k = \frac{2}{N}\sum_{n=1}^{N}\frac{R_n x_{kn}}{\bar{x}_k} - 1$, which represents the inequality of socioeconomic determinant *k*. We also define the concentration index for the error term as $C_e = \frac{2}{N}\sum_{n=1}^{N} R_n e_n$. With these, we obtain Equation (14) below.

$$\text{Gini coefficient } (I) = \sum_{k=1}^{K}\left(\frac{\beta_k \bar{x}_k}{\bar{I}}\right) C_k + \frac{C_e}{\bar{I}} \quad (14)$$

When we define $S_k = \frac{\beta_k \bar{x}_k}{\bar{I}}$ as the income elasticity of socioeconomic determinant *k*, we obtain Equation (15), which shows how much of the estimated total income inequality can be expressed as the sum of the contributions of its socioeconomic determinants. The contribution of each socioeconomic determinant is the product of C_k and S_k, representing a weighted inequality of factor *k*, where the weights are the income elasticity S_k.

$$\text{Gini coefficient } (I) = \sum_{k=1}^{K} S_k\, C_k + \frac{C_e}{\bar{I}} \quad (15)$$

The conventional decomposition approach in Equation (11) can be considered a special case of the regression-based approach described in Equations (14) and (15). Because total income *I* is exactly the sum of all income components, β equals one for each income component and there is no error term in Equation (11). In a robustness test, we decompose the Theil index into its socioeconomic determinants and obtain results that are consistent with the Gini coefficient decomposition (see Section 6).

3.3. Inequality mediation decomposition

Our conceptual framework described in Section 2 hypothesizes that the transmission process from the urban-rural gap to income inequality involves educational attainment as a mediator. The existing decomposition approaches do not allow us to test for this potential mediating effect. We thus develop a new approach, which we call *inequality mediation decomposition*, by combining the regression-based decomposition approach (Fields, 1998; Morduch and Sicular, 2002; Wan, 2004) and the three-step mediating effect test (Baron and

Kenny, 1986).[9] Instead of conducting three regressions as in the original mediating test, the new approach conducts three regression-based decompositions. In Step 1, we test how the urban-rural gap affects educational inequality. Step 2 tests how the urban-rural gap affects income inequality. Step 3 tests whether the urban-rural gap still significantly affects income inequality when education is included in the model and whether educational attainment significantly affects income inequality. The regression equations are as follows:

Step 1: $$Education_n = \alpha_1 + \beta_{11} hukou_n + \beta_{1X} X_n + \varepsilon_{1n} \quad (16)$$

Step 2: $$I_n = \alpha_2 + \beta_{21} hukou_n + \beta_{2X} X_n + \varepsilon_{2n} \quad (17)$$

Step 3: $$I_n = \alpha_3 + \beta_{31} hukou_n + \beta_{32} Education_n + \beta_{3X} X_n + \varepsilon_{3n} \quad (18)$$

Educational attainment is identified as a mediator between the urban-rural gap and income inequality if all of the following conditions are satisfied: (1) the impact of the urban-rural gap on educational inequality in Step 1 is significant; (2) the impact of the urban-rural gap on income inequality in Step 2 is significant; (3) the impact of educational attainment on income inequality in Step 3 is significant; and (4) the impact of the urban-rural gap on income inequality in Step 3 is smaller than that in Step 2. There is perfect mediation if the urban-rural gap has no significant impact on income inequality in Step 3 when education is controlled for. We report the magnitude of the impact by the *contribution* of a socioeconomic factor to inequality $S_k \times C_k$, which is in the same unit as the Gini coefficient. We also report the *contribution in percentage* $S_k \times C_k / Gini$. We test the significance of the impact by checking whether the regression coefficients are significantly different from zero.

4. Data

4.1. CHNS dataset and our sample

We use panel data from the China Health and Nutrition Survey (CHNS) for the period 1991-2015. The CHNS is a collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention. CHNS data are available

[9] Both approaches have been widely used in the economic literature to identify the drivers and sources of inequalities (e.g., Gustafsson and Shi, 2002; Lambert et al., 2014) or to identify a mediating effect of a variable (e.g., Buchan et al., 2008; Ding et al., 2018). We are the first to integrate the two approaches.

from ten waves (1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, 2011, and 2015), in which surveyed households were revisited and new households were added to replace households no longer participating in the study. We exclude the first survey year 1989 from our sample because the questions asked in the 1989 survey are different from those in other waves, and it is therefore difficult to compare the 1989 data with other waves (see also, Zhong, 2011; Baeten et al., 2013; Cao and Birchenall, 2013). The CHNS uses a multistage, random cluster-sampling scheme to select the households from each sampled province. Counties in each province were selected using a weighted method. Villages and townships in each county and urban and suburban neighborhoods in each city were randomly selected. The survey collects detailed information on income, assets, and demographics for each household. The CHNS has been widely used in household finance studies (e.g., Chamon et al., 2013, Yu and Zhu, 2013) and economic development studies (e.g., Cao and Birchenall, 2013; Santaeulàlia-Llopis and Zheng, 2018) in China.

The CHNS covers the following provinces: Guangxi, Guizhou, Heilongjiang (from 1997 onward), Henan, Hubei, Hunan, Jiangsu, Liaoning (except 1997), and Shandong. Since 2011, Beijing, Shanghai, and Chongqing have been included in the CHNS. We construct an unbalanced panel of households using CHNS data from all available waves during the period 1991-2015. We exclude Beijing, Shanghai, and Chongqing in 2011 and 2015 to keep the sample comparable across waves.[10] The nine provinces in our sample vary substantially in terms of geography, economic development, and other socioeconomic factors and provide a good representation of China (Chamon et al., 2013).

We define households with at least one member aged 60 or older as old households (Cai et al., 2006; Connelly and Maurer-Fazio, 2016).[11] We refer to this sample of old households as Sample A. We refer to the remaining sample consisting of households with only young members as young households (Sample B). We use the demographic characteristics of the

[10] In a robustness test, we repeat our analyses including the observations of Beijing, Shanghai, and Chongqing in 2011 and 2015. The results are consistent with our main results (see Section 6).

[11] Studies on the U.S. identify old households using a similar logic, i.e., classify a household as old if one member is old (Deaton and Paxson, 1994; Costa, 1997). In a robustness test, we repeat our analyses in two subsamples of the old households: "pure old" households (all members are 60 or older) and "mixed" households (with both old and young members). The results are consistent with our conclusions (see Section 6).

household head to describe the household.[12] Alternatively, in a robustness test, we use a household's average demographic information to describe the household (see Section 6). We identify the household head based on self-reporting in the CHNS questionnaire. The CHNS defines a household based on sharing physical presence (members eat and live with each other) instead of *hukou* registration. We use the original *hukou* status[13], gender, age, and education of the household head.

We compute the following six income categories for each household: (i) income from work as the sum of wages, net agricultural income, and net business income; (ii) pension benefits;[14] (iii) capital income; (iv) private transfers; (v) subsidies, and (vi) other unspecified income. The definition and variable sources for each income category are summarized in Appendix A. We compute the household income per capita as the sum of all income components divided by the number of household members.[15] We convert all income to real values in 2015 using the consumer price index provided by the CHNS (Santaeulàlia-Llopis and Zheng, 2018). The indices are based on the standard consumer basket supplied by the National Bureau of Statistics of China. This price index differs by province, by urban/rural areas within each province, and by year. It is set to 100 for urban Liaoning in 2015. It adjusts incomes for different costs of living in different regions and years and makes income comparable across years and provinces. To reduce the bias from extreme values, we trim the top and bottom 1% of real net household income per capita in each year (Chamon et al.; 2013; Santaeulàlia-

[12] Household head characteristics are often used to describe the household (Cai et al., 2006; Cao and Birchenall, 2013). In the CHNS, about three-quarters of the household heads of old households are aged 60 or older (Connelly and Maurer-Fazio, 2016), indicating that household head information well captures the characteristics of old Chinese.

[13] To capture the socioeconomic origin of the household head, we would like to include his/her original *hukou* status at birth. As this information is not available in the CHNS, we approximate the original *hukou* status of the household head by the *hukou* status of his/her parents recorded in CHNS. We assume the original *hukou* status as urban if one of the parents has urban and the other has rural *hukou*. If this information is also missing, we assign the *hukou* status based on the birthplace (rural vs. urban). If this information is also missing, we use the reported *hukou* status in the earliest survey year.

[14] Pension benefits in the CHNS include both public and private pensions. However, private pension benefits are less than 1% of total pension benefits in China (Zheng et al., 2019). Thus, the scale and trend of total pension benefits mostly reflects the public pension benefits.

[15] Some studies estimate household income per capita using adult equivalents (see e.g., Santaeulàlia-Llopis and Zheng, 2018). To capture the living standard of a household, we assume that all household members equally share the total household income, which is consistent with the practice of the Chinese government, which uses the household income per capita based on all household members to measure the living standard. The assumption that children or seniors consume less than working adults may not be accurate in China, where education and medical services are relatively expensive. Therefore, we give equal weight to each household member.

Llopis and Zheng, 2018).[16] We assume that a missing value in a given income category represents a zero when a household reports at least one other income category.

There are a total of 7,822 households and 37,740 household-year observations in the nine waves of the CHNS 1991-2015 (excluding Beijing, Shanghai, and Chongqing in 2011 and 2015). We exclude 1,585 observations that have one or more missing values in the household characteristics or for household income per capita. Of the remaining 36,155 observations (7,654 households), 14,802 observations (4,239 households) have at least one household member aged 60 or older – these are the old households in Sample A. The remaining 21,353 observations (5,780 households) are young households in Sample B.[17]

We also construct a balanced sample for our longitudinal analysis. We identify a cohort of households (Sample C) that were young households in 2000,[18] became old households in 2015, and continuously appeared in the waves of 2004, 2006, 2009, and 2011. This sample is smaller than Samples A and B because it consists of the same 762 households in each wave during 2000-2015.

For Sample C, we have information on the occupational sector of most household heads. This information is not available for most old households in Sample A since the CHNS only records the occupational information for individuals who are working in the survey year but not the occupation of the retired and the unemployed (Chamon et al., 2013). We classify a household head as belonging to the *formal sector* if he/she is working or has worked in the government, public institutions, or formally established enterprises. According to the pension regulation in China, formal sector employees participate in the high-benefit Employees' Basic Pension Program, while informal sector residents participate in the low-benefit Residents' Basic Pension Program (Zheng et al., 2019). Thus, the occupational sector in essence captures which public pension program an individual should participate in. A total of 17.4% of household heads belong to the formal sector, while the rest belong to the informal sector.

[16] The trimmed observations are treated as missing values. Alternatively, we trim the top and bottom 0.5% and repeat all analyses. All our conclusions hold. The results are available from the authors upon request.

[17] In a robustness test, we repeat our analyses excluding young households having early retired members who receive a pension before 60. The results support our conclusions (see Section 6).

[18] We use 2000 instead of 1991 as the starting year because the sampled provinces remained the same since 2000, allowing us to follow the same households. Choosing 2000 as the starting year also yields the largest longitudinal sample. Using earlier starting years would significantly reduce the sample size due to panel attrition in the CHNS. Using later years would also reduce the sample size because fewer households changed from young to old during the shorter sampling period.

Table 1: Descriptive statistics: Demographics and household income per capita

Year	N	Urban	Male	Age	Education	Household size	Living arrangement (100%)				Household income per capita				
							With adult children	Single	Couple	Other	Mean	Std. Dev.	P10	Median	P90
Panel A: Cross-sectional sample (Sample A)															
Full sample	14,802	0.41	0.75	63.09	5.81	3.45	70%	11%	16%	2%	9,623	11,220	1,167	5,616	23,551
1991	1,142	0.44	0.77	56.17	4.45	4.12	80%	6%	11%	4%	3,119	2,014	841	2,752	5,792
1993	1,120	0.43	0.75	57.58	4.66	4.06	79%	6%	12%	3%	3,498	2,590	878	2,817	6,872
1997	1,257	0.42	0.75	59.95	4.81	3.68	70%	9%	15%	5%	4,041	2,972	994	3,277	8,017
2000	1,429	0.44	0.75	61.55	5.42	3.48	65%	10%	20%	5%	5,391	4,553	947	4,224	11,384
2004	1,618	0.44	0.75	64.65	5.84	3.10	59%	14%	24%	3%	7,510	6,864	1,281	5,379	16,705
2006	1,806	0.42	0.74	64.73	5.85	3.27	72%	13%	15%	1%	8,170	8,135	1,172	5,429	18,615
2009	1,970	0.38	0.74	64.91	6.06	3.30	70%	13%	15%	1%	11,919	11,198	1,816	8,783	26,098
2011	2,090	0.39	0.76	66.09	6.21	3.27	68%	13%	18%	1%	13,462	11,905	1,637	10,150	29,285
2015	2,370	0.35	0.74	65.12	7.18	3.38	75%	11%	13%	2%	18,422	16,956	1,919	14,184	39,822
Panel B: Longitudinal sample (Sample C)															
Full sample	4,572	0.20	0.85	56.87	6.72	3.88	85%	3%	7%	5%	10,226	10,900	1,635	6,899	23,000
2000	762	0.20	0.88	49.69	6.73	4.10	80%	1%	5%	14%	5,575	4,103	1,327	4,691	11,130
2004	762	0.20	0.87	53.40	7.03	3.79	80%	2%	10%	7%	7,083	6,237	1,473	5,264	14,571
2006	762	0.20	0.86	55.87	6.57	3.91	90%	2%	5%	3%	7,777	7,440	1,482	5,719	16,268
2009	762	0.20	0.85	58.61	6.57	3.86	88%	4%	7%	1%	10,922	9,139	2,118	8,507	22,222
2011	762	0.21	0.83	60.58	6.51	3.76	84%	5%	9%	1%	13,279	11,572	2,222	9,789	28,062
2015	762	0.21	0.80	63.09	6.92	3.86	90%	6%	3%	1%	16,723	17,189	1,884	11,530	36,930

Notes: Mean values are reported. All variables are defined in Appendix A. Urban, Male, Age, and Education are the characteristics of the household head. Education refers to the completed years of formal education. Household income per capita is inflation-adjusted to 2015. The descriptive statistics of young households (Sample B) are available from the authors upon request.

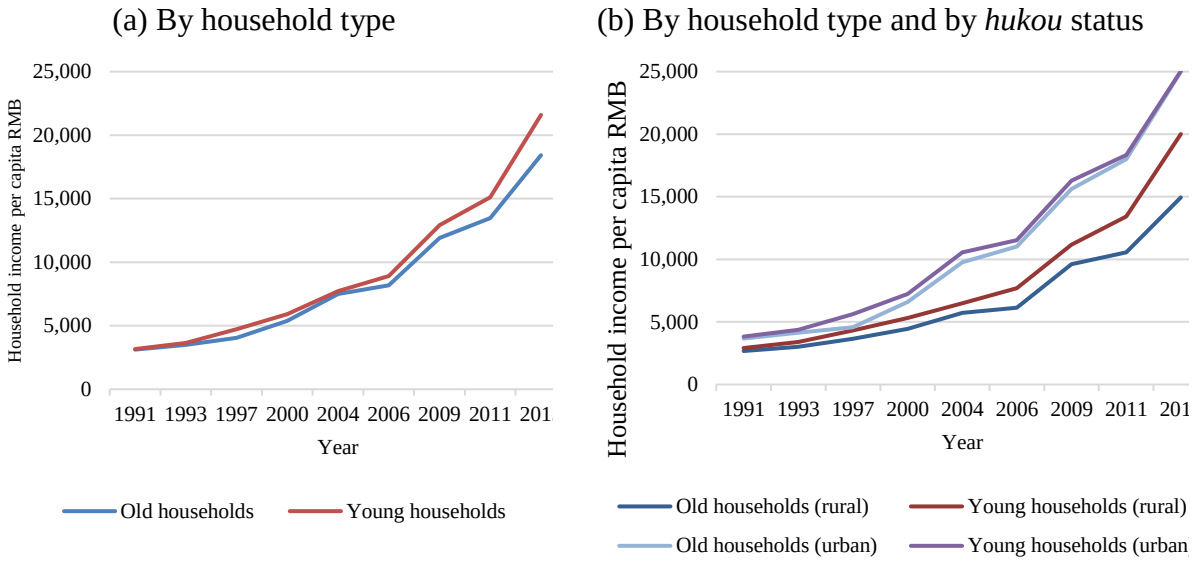
4.2. Descriptive statistics

Panel A of Table 1 reports the descriptive statistics of the demographic variables and (real) household income per capita of the old households (Sample A) for each survey year and the average values for all years. The sample size increases over time for two reasons: The survey participants age and more households have members aged 60 and older in later waves, and new households were added to the survey. In total, 75% of household heads are male. The average years of completed formal education of the household head increased from 4.5 years in 1991 to 7.2 years in 2015. The average household size decreased from 4.1 in 1991 to 3.4 in 2015. The share of single households increased from 6% in 1991 to 11% in 2015. The summary statistics for age, gender, and education show similar trends as those reported in Cao and Birchenall (2013) and Lee and Malin (2013) based on the CHNS.

Real household income per capita steadily increased from 1991 to 2015. Its standard deviation also increased during the entire sample period, indicating that the dispersion of household income per capita increased. This greater dispersion is also evident in the 10^{th} and 90^{th} percentiles of household income per capita. The 90^{th} percentile increased more than six times, from RMB 5,792 (USD 839) in 1991 to RMB 39,822 (USD 5,771) in 2015, while the 10^{th} percentile increased less than threefold, from RMB 841 (USD 122) in 1991 to RMB 1,919 (USD 278) in 2015. These numbers indicate that the income of high-income old households grew much faster than that of low-income old households, resulting in a widened income gap over the sample period.

Figure 1 compares the income of the old households in Sample A with that of the young households in Sample B. Figure 1(a) shows the average household income per capita for the two groups over the sample period 1991-2015. Income levels have rapidly increased for both household types and show similar trends. Figure 1(b) compares the income trends by *hukou* status of the household head. The average income is always lower for rural households in both samples. The urban-rural income gap is larger for old households than for young households, indicating greater income inequality among the old.

Figure 1: Trends in household income per capita

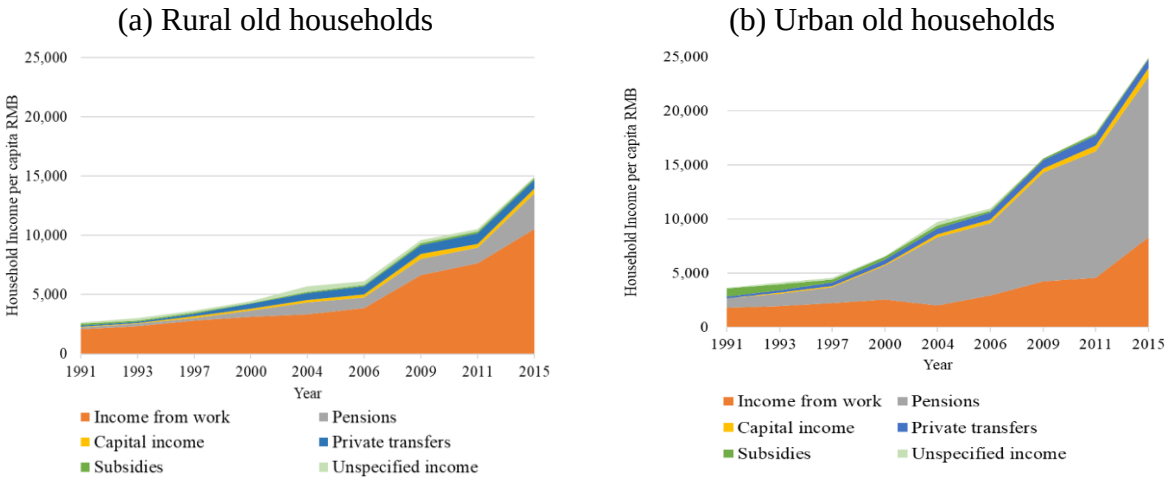


Notes: Average real net household income per capita.

Panel B of Table 1 reports the summary statistics of Sample C, which is the longitudinal sample following the same households from 2000 to 2015. Sample C includes fewer urban households than Sample A because the rate of successful revisits is lower for urban households than for rural households. For this reason, Sample C reports a higher fraction of male household heads, a larger household size, and a higher rate of living with adult children than Sample A. The annual average age in Sample C increased over time as the cohort ages.

Figure 2 shows the trends in household income per capita by income components. Figure 2(a) shows that rural old households mainly rely on income from work (of both old and young household members), while Figure 2(b) shows that urban old households mainly rely on pension income. For both urban and rural households, the fraction of pension income has increased over time, reflecting the expansion of public pension programs in China (Zheng et al., 2019). However, the pension benefits and their increase are much smaller for rural households than for urban households, implying that the contribution of the urban-rural gap to income inequality is likely to be driven by pension income. Private transfers from outside the household are a larger income component for rural old households than for urban old households. Capital income and subsidies contribute little to household income for both rural and urban households.

Figure 2: Trends in household income per capita by income components



Notes: Average real net household income per capita.

5. Results

In Sections 5.1.-5.3., we present the results of repeated cross-sectional analyses based on the full (unbalanced) samples of old households (Sample A) and young households (Sample B). In Section 5.4., we provide the results of the longitudinal analyses based on the balanced sample of households that were young in 2000 and old in 2015 (Sample C). In Section 5.5., we analyze the role of the fragmented pension system based on the old households in 2015 in Sample C.

5.1. Income inequality among old households compared with young households

We first conduct cross-sectional analyses to compare income inequality for different age groups in each survey year. The differences in income inequality for old households and for young households in each wave can be driven by age effects or cohort effects.[19] Therefore, we repeat the cross-sectional analysis for each of the nine survey waves over 1991-2015 to minimize the impact of cohort effects.

Table 2 presents the Gini coefficients, Theil indices, and P90/P10 ratios for old households and young households, respectively. All three measures of inequality show that the income inequality of old households was higher than that of young households in all

[19] Changes in inequality can be driven by age effects when the age structure of the population changes and income differs by age, or it can be due to cohort effects when income varies by cohort and over generations. We are interested in comparing and explaining the income inequality by age groups; that is, we are interested in the age effects.

survey years during the sample period 1991-2015. We also compute the mean values of the Gini coefficients, Theil indices, and P90/P10 ratios by averaging each measure over all survey years. The differences between these averages for old and young households are all statistically significant based on mean difference *t*-tests. The finding that income inequality is higher among old households confirms *Hypothesis 1* and is consistent with the theoretical view that inequality accumulates during working life (Deaton and Paxson, 1994).

Table 2: Income inequality by household type

	Old households (Sample A)			**Young households** (Sample B)			**Difference between old and young** (Sample A minus Sample B)		
Year	Gini	Theil	P90/P10	Gini	Theil	P90/P10	Gini	Theil	P90/P10
1991	0.351	0.201	6.9	0.350	0.198	6.5	0.002	0.003	0.4
1993	0.393	0.253	7.8	0.388	0.245	7.7	0.005	0.008	0.1
1997	0.389	0.249	8.1	0.370	0.222	7.1	0.019	0.026	0.9
2000	0.438	0.319	12.0	0.393	0.255	8.5	0.046	0.064	3.5
2004	0.464	0.352	13.0	0.451	0.330	11.9	0.014	0.022	1.1
2006	0.489	0.395	15.9	0.463	0.352	12.5	0.026	0.043	3.4
2009	0.467	0.360	14.4	0.441	0.314	11.2	0.026	0.046	3.2
2011	0.463	0.349	17.9	0.425	0.291	11.7	0.038	0.058	6.1
2015	0.471	0.362	20.8	0.463	0.344	18.5	0.008	0.018	2.2
Average	0.436	0.316	13.0	0.416	0.283	10.6	0.020***	0.033***	2.4***

Notes: The Gini coefficients and Theil indices are estimated based on household income per capita using Equations (9) and (10), respectively. We use *t*-tests to assess the differences in the average Gini coefficient, average Theil index, and average P90/P10 ratio. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

We also note that the income inequalities of old households and young households show similar trends over time. For both samples, the income inequality measured by the Gini coefficient and the Theil index increased between 1997 and 2006 and decreased between 2006 and 2011. Previous studies report similar trends for the total population, including old and young (see Appendix B), confirming the validity of our sample. The P90/P10 ratios increased continuously from 1991 to 2015 in both samples. The P90/10 ratios also varied more over time than the Gini coefficients, indicating that the income distributions changed more at the margins than at the median levels over the sample period.

5.2. Decomposition of income inequality among old households

Decomposition (1) by socioeconomic characteristics

Table 3 reports the results of the three-step inequality mediation decomposition using data for 2015, which is the most recent available year in the CHNS. We report the results

for the other survey years in Appendix C and discuss them in a separate subsection below. We report the estimated coefficients (β_k), their standard deviations, the contribution to income inequality (the product of C_k and S_k) for each socioeconomic determinant, and the contribution in percentage (the contribution divided by the Gini coefficient). We discuss the results for old households in Panel A of Table 3 in this section and compare them with the results for young households in Panel B in Section 5.3.

Step 1 analyzes the impact of the urban-rural gap on educational inequality. The regression coefficient of urban *hukou* is 2.742, indicating that the urban old, on average, completed three more years of education than the rural old. The decomposition results in Columns 3 and 4 in Table 3 indicate that the urban-rural gap makes the largest contribution (9.3%) to educational inequality among all socioeconomic determinants.

Step 2 tests how the urban-rural gap affects income inequality without explicitly considering the mediating effect of education. The per capita income of urban old households exceeds that of rural old households by RMB 8,651 (USD 1,272), on average, in 2015 (Column 1). The urban-rural gap contributes 8.4% (Column 4) to the Gini coefficient of the overall income inequality among old households, which is 0.471.

Step 3 repeats the analyses in Step 2 but adds education as an explanatory variable. Both the urban-rural gap and education have a significantly positive impact on old households' income and contribute to old-age income inequality. Compared to Step 2, the contribution of the urban-rural gap to income inequality among the old households is reduced from 8.4% to 5.7% (Column 4). The reduced 2.7 percentage points are mediated by educational inequality. Education in total explains 8.5% (Column 4) of the income inequality among the old households. Of the other inequality determinants, household size (5.6%) and regional differences (5.2%) are also important.

Table 3: Decomposition (1) by socioeconomic characteristics in 2015

Decomposition (1)	**Panel A: Old households** (Sample A, N=2,370)				**Panel B: Young households** (Sample B, N=1,843)			
	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)
Step 1: Education								
Urban hukou	2.742***	0.155	0.030	9.3%	2.078***	0.157	0.018	9.9%
Male	2.225***	0.18	0.020	6.3%	0.130	0.209	0.000	-0.1%
Age	-0.103***	0.007	0.027	8.2%	-0.053***	0.013	0.002	1.4%
Household size	-0.001	0.048	0.000	0.0%	-0.127**	0.063	0.001	0.8%
Single	-0.424	0.285	0.002	0.5%	0.021	0.409	0.000	0.0%
Couple	0.503**	0.239	0.001	0.4%	0.506*	0.267	0.001	0.3%
Other living arrangements	0.467	0.588	0.000	0.0%	0.518**	0.227	0.002	0.8%
Province dummies	Yes	Yes	0.012	3.7%	Yes	Yes	0.003	1.7%
	R^2=0.287		Gini=0.322	ê: 71.5%	R^2=0.142		Gini=0.179	ê: 85.1%
Step 2: Household income per capita								
Urban hukou	8.651***	0.693	0.039	8.4%	4.487***	1.007	0.009	1.9%
Male	1.698**	0.804	0.001	0.1%	-1.768	1.345	0.001	0.2%
Age	-0.032	0.031	0.000	-0.1%	-0.156*	0.082	-0.001	-0.3%
Household size	-1.626***	0.213	0.026	5.6%	-2.030***	0.406	0.019	4.0%
Single	-2.567**	1.274	0.000	-0.1%	-2.270	2.628	-0.001	-0.1%
Couple	1.291	1.067	0.002	0.5%	0.453	1.716	0.000	0.1%
Other living arrangements	-3.726	2.627	0.000	0.1%	-4.280***	1.46	0.004	0.9%
Province dummies	Yes	Yes	0.028	6.2%	Yes	Yes	0.034	7.4%
	R^2=0.163		Gini=0.471	ê: 79.2%	R^2=0.098		Gini=0.463	ê: 85.9%
Step 3: Household income per capita								
Urban hukou	5.901***	0.719	0.027	5.7%	2.695**	1.045	0.005	1.2%
Education	1.003***	0.09	0.040	8.5%	0.862***	0.149	0.019	4.0%
Male	-0.534	0.809	0.000	0.0%	-1.880	1.333	0.001	0.2%
Age	0.0710**	0.031	0.001	0.1%	-0.110	0.082	-0.001	-0.2%
Household size	-1.625***	0.207	0.026	5.6%	-1.920***	0.403	0.018	3.8%
Single	-2.142*	1.242	0.000	-0.1%	-2.288	2.605	-0.001	-0.2%
Couple	0.787	1.041	0.001	0.3%	0.016	1.702	0.000	0.0%
Other living arrangements	-4.194	2.56	0.000	0.1%	-4.727***	1.449	0.005	1.0%
Province dummies	Yes	Yes	0.024	5.2%	Yes	Yes	0.032	7.0%
	R^2=0.205		Gini=0.471	ê: 74.6%	R^2=0.114		Gini=0.463	ê: 83.2%

Notes: The results are based on the three-step inequality mediation decomposition method described in Section 3.3. The OLS regression coefficients and standard errors of Steps 1, 2, and 3 are from estimations of Equations (16), (17), and (18), respectively. All monetary variables are scaled by 1/1000. The regression omits the binary variable 'households with adult children', as the baseline living arrangement. The decomposition includes eight province dummies, and we report their total contribution to inequality. The last row in each decomposition shows the regression R^2, the Gini coefficient of household income per capita, and the contribution in percentage of the error term to income inequality. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

All four conditions to establish the mediating effect of education are satisfied. We thus conclude that the impact of the urban-rural gap on income inequality is partially mediated by education. Among old households, the urban-rural gap is the largest contributor to educational inequality (9.3% in Step 1). Together, the urban-rural gap and education are also the two largest contributors to income inequality (5.7% and 8.5%, respectively, in Step 3). Therefore, we conclude that the urban-rural gap and education attainment are the two primary causes of old-age income inequality, together contributing 14.2% of old-age income inequality. We are also confident that the transmission process from the urban-rural gap via educational inequality to income inequality is one of the most important channels for income inequality among old households in China. The results confirm *Hypothesis 2*.

We note that the regression error terms and their contribution to income inequality remain large in our model. To address this concern, we add more control variables in a robustness test to identify other potential inequality determinants, which improves the model fit while our conclusions hold (see Section 6). Endogeneity is less a concern in our regressions. We use the household head's original *hukou* status, which is predetermined when we assess the income and education of working-age and old individuals. The educational attainment is also largely predetermined because most Chinese finish formal education in early age. In a robustness test, we exclude households whose head changed his/her *hukou* status and/or educational attainment. The results support our conclusions (see Section 6). We also control for the migration of household heads (mostly from rural to urban areas) in Section 6, and the results support our conclusions.

Decompositions (2) by income components

To analyze how the urban-rural gap and educational attainment affect income inequality through different income components, we conduct two complementary decompositions. First, we decompose the inequality by income components (Decomposition (2), see Table 4). Second, we decompose the inequality of the income component that contributes most to the overall income inequality by its socioeconomic determinants (Decomposition (3), see Table C1 in Appendix C). These two decompositions reveal how the transmission process works through different income components and are particularly relevant to the design of public policies aimed at income redistribution. In this section, we focus on the results for old households shown in Panel A of Table 4. We compare the results with those for young households in Section 5.4.

Table 4: Decomposition (2) by income components in 2015

Decomposition (2)	Concentration index (C_m)	Income share (S_m)	Contribution ($C_m \times S_m$)	Contribution in % ($C_m \times S_m$ / Gini)
Panel A: Old households (Sample A, N = 2,370)				
Pensions	0.592	0.384	0.227	48.2%
Income from work	0.398	0.532	0.212	44.9%
Capital income	0.545	0.030	0.016	3.4%
Private transfers	0.260	0.043	0.011	2.3%
Subsidies	-0.072	0.008	-0.001	-0.1%
Unspecified income	1.371	0.004	0.006	1.2%
Household income per capita		1.000	Gini=0.471	100.0%
Panel B: Young households (Sample B, N=1,843)				
Pensions	0.457	0.060	0.027	5.9%
Income from work	0.462	0.884	0.408	88.1%
Capital income	0.530	0.026	0.014	2.9%
Private transfers	0.311	0.015	0.005	1.0%
Subsidies	0.077	0.004	0.000	0.1%
Unspecified income	0.808	0.011	0.009	1.9%
Household income per capita		1.000	Gini=0.463	100.0%

Notes: The results are estimated based on Equation (11).

The results in Panel A of Table 4 show that pensions are more unequally distributed than the total household income per capita and are the most unequally distributed income component for old households. The concentration index for pension income (C_m = 0.592) is larger than the Gini coefficient of household income per capita (0.471), and it is the largest concentration index among all income components (Column 1). In other words, a general increase, for example, of 5%, in pensions for each old person would increase income inequality. As pension income accounts for a relatively large share of total household income for old households (S_m = 38.4%), it contributes 48.2% to the income inequality among old households. This result is likely driven by the limited redistribution between the fragmented public pension programs in China.

The income from work constitutes more than half of the total income of old households (S_m = 53.2%).[20] It is more equally distributed than total income, with C_m = 0.398, which is smaller than the Gini coefficient of 0.471. This result is consistent with the finding of Paul et al. (2017) that the contribution of wages and salaries to inequality in China has declined in recent years. In our sample, income from work contributes 44.9% of the overall income inequality among old households. We therefore conclude that pensions and income from work are the two main sources of income inequality among old households in China.

The total contribution of other income components to income inequality among old

[20] In a robustness test, we distinguish the income from work earned by the old and by the young using individual income information in the CHNS. The results show that pension income remains the largest contributor to the income inequality among old households (see Section 6).

households is less than 7%, as these components account for very small shares in household income. In line with previous literature, we find that capital income is more unequally distributed than income from work (Zhuang and Li, 2016) and total household income per capita. Subsidies and private transfers from outside the household mitigate income inequality. However, the redistribution effects of subsidies and private transfers are very limited due to their small shares in income (Santaeulalia-Llopis and Zheng, 2018).

Table C1 in Appendix C reports the results of Decomposition (3) for pension income, which has the largest contribution to income inequality among old households. The results suggest that the urban-rural gap and educational inequality contribute to the income inequality of old households mainly through the channel of pension inequality. The total contribution of the urban-rural gap to pension inequality is 21.3%, of which 5.8 percentage points are mediated by educational attainment. The urban-rural gap (15.5%) and education (13.1%) are the two largest contributors to pension inequality, much larger than their contributions to the overall income inequality of old households.

Inequality decompositions in other survey years

We repeat the decomposition analyses for the survey years before 2015. The results in Table C2 in Appendix C confirm our *Hypothesis 2* and the transmission process from the urban-rural gap to educational inequality and then to income inequality in old age in all survey years except in the first year, 1991. The contribution of the urban-rural gap to old-age income inequality was 8.9% in 2011, 7.2% in 2009, 9.7% in 2006, and 9.1% in 2004, of which approximately one-third to half was mediated by education. The urban-rural gap and educational attainment remain two primary causes of old-age income inequality, contributing 11%-22% of old-age income inequality during 2000-2015.[21]

Figure C1 in Appendix C summarizes the decomposition results by income components during 1991-2015. Pensions have been the largest contributor to income inequality among old households since 2004. Before that, income from work was the largest contributor. The reason for this shift is that China has experienced a gradual expansion of public pension coverage, first to urban employees during 1997-2005, then to rural residents during 2009-2014 and to other urban residents during 2011-2014 (Zheng et al., 2019). Before 1997, most old Chinese received no pension, and therefore, pensions could not contribute much to income inequality among the old. During 2000-2011, the contribution of pension inequality

[21] We note the increasing impact of the urban-rural gap on education in Step 1. This is because before the 1950s, there was no *hukou* system, and most individuals received very little education.

to overall old-age income inequality was 45% - 60%.

5.3. Comparing the inequality decomposition between old households and young households

Here, we identify reasons for the greater income inequality among old households than among young households. As shown in Table 3, the contribution of the urban-rural gap to income inequality is larger among old households (0.027 and 5.7% in Step 3, Panel A) than among young households (0.005 and 1.2% in Step 3, Panel B). The same is true for educational attainment, whose contribution to the inequality of old households is 0.040 or 8.5%, which is larger than the contribution of educational attainment to the inequality of young households (0.019 or 4.0%). The contributions of the remaining socioeconomic characteristics among old households are very similar in level to those among young households.[22] The results in other survey years also show that the contributions of the urban-rural gap and educational attainment to income inequality are always larger in the old household sample (Table C2 in Appendix C) than in the young household sample (Table C3 in Appendix C) since 2000. The results indicate that the transmission process is stronger among old households than among young households. That is, the advantages and disadvantages caused by the urban-rural gap and educational attainment accumulate over the life cycle, resulting in greater income inequality among old households than among young households.

Panel B of Table 4 shows that income from work has the largest contribution percentage to income inequality among young households (88.1%). Moreover, income from work is the single major source of income of young households, with a share of 88.4%. As shown in Figure C1 in Appendix C, the contributions of income from work to income inequality among young households increased from 70% to 85% during 2004-2011. From the income component perspective, the income inequality among old households is higher than that among young households because of the additional contribution of pensions to inequality among old households in addition to the contribution of income from work. This suggests that the public pension system in China reinforces the inequality of income from work, leading to greater income inequality among old households.

From the socioeconomic determinant perspective, the urban-rural gap contributes very

[22] The results in Table 3 show a larger contribution of provincial differences to income inequality among the young than among the old. This suggests that the pension system in China mitigates regional inequality.

little to the inequality of income from work among young households (0.0%),[23] and the contribution of education is small (3.5%), as shown in Table C1 in Appendix C. However, both the urban-rural gap and education contribute substantially to the inequality of pension income among old households (15.5% and 13.1%, Table C1). The results suggest that the contributions of the urban-rural gap and education to income inequality are mainly through the channel of pension income (pension inequality) instead of income from work.

5.4. Longitudinal analyses

This section reports the results for the longitudinal sample (Sample C), which follows one cohort of households that were young households in 2000 and old households in 2015. We examine the trend in income inequality and compare the results of the inequality decompositions in 2000 and 2015 for this cohort. Analyzing this longitudinal data allows us to exclude potential cohort effects and minimize the impact of uncontrollable household heterogeneity on income inequality. However, we note that compared to the repeated cross-sectional analyses, the sample size reduces significantly (one-third of old households in Sample A in 2015).

Table 5 shows that the income inequality within Sample C increased from 2000 to 2015: the Gini coefficient was 0.388 in 2000 and 0.503 in 2015. This increase might mix age effects with time effects. We conduct a time-series regression to disentangle the age effects from the time effects, where age effects are captured by the fraction of old households (Luo et al., 2018) and time effects are captured by the income inequality of the entire population in each wave, a more informative time trend than wave dummies. The results in Table C4 in Appendix C show that income inequality, measured by the Gini coefficient, Theil index, and P90/P10 ratio, increased as the fraction of old households increased in Sample C over the period 2000-2015. The results support *Hypothesis 1*.

[23] As a side result, we note that in 2015, the *hukou* system no longer limited the migration of young workers to seek higher income employment or business opportunities in urban areas. Instead, the primary obstacle for the rural young to obtain a higher income is the inequality of education. The results are consistent with Imai and You's (2014) finding that "out-migration" for better jobs significantly reduces the poverty in rural areas, mitigating the income inequality caused by the urban-rural gap.

Table 5: Decomposition (1) by socioeconomic characteristics in 2000 and 2015

Decomposition (1)	**Panel A: Old households in 2015** (Sample C, N=762)				**Panel B: Young households in 2000** (Sample C, N=762)				
	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)	Contribution to change in Gini 2000-2015
Step 1: Education									
Urban hukou	2.736***	0.317	0.021	7.0%	2.593***	0.292	0.024	8.4%	-10.7%
Male	2.175***	0.327	0.016	5.1%	2.437***	0.357	0.011	4.0%	17.9%
Age	-0.109***	0.016	0.017	5.5%	-0.135***	0.022	0.016	5.6%	3.6%
Household size	-0.126*	0.068	0.001	0.4%	-0.176**	0.090	0.002	0.8%	-3.6%
Single	-1.179**	0.580	0.003	0.8%	-2.455**	1.222	0.002	0.7%	3.6%
Couple	0.121	0.716	0.000	0.0%	-0.048	0.558	0.000	0.0%	0.0%
Other living arrangements	0.584	1.300	0.000	0.0%	-0.483	0.380	-0.002	-0.6%	7.1%
Province dummies	Yes	Yes	0.012	3.7%	Yes	Yes	0.009	3.3%	10.7%
	R^2=0.228		Gini=0.309	ê: 77.5%	R^2=0.218		Gini=0.281	ê: 77.9%	100.0%
Step 2: Household income per capita									
Urban hukou	11.85***	1.470	0.045	8.9%	1.381***	0.365	0.011	2.8%	29.6%
Male	1.298	1.520	0.000	-0.1%	-0.839*	0.446	0.002	0.6%	-1.7%
Age	-0.117	0.072	0.001	0.2%	0.008	0.028	0.000	0.1%	0.9%
Household size	-1.492***	0.316	0.028	5.6%	-0.500***	0.112	0.019	4.9%	7.8%
Single	-2.367	2.691	-0.001	-0.2%	-2.112	1.528	0.000	0.1%	-0.9%
Couple	-3.047	3.323	0.000	0.0%	0.632	0.698	0.001	0.4%	-0.9%
Other living arrangements	-5.935	6.036	0.001	0.2%	-0.732	0.475	0.001	0.3%	0.0%
Province dummies	Yes	Yes	0.037	7.4%	Yes	Yes	0.030	7.8%	29.6%
	R^2=0.174		Gini=0.503	ê: 78.0%	R^2=0.150		Gini=0.388	ê: 83.2%	100.0%
Step 3: Household income per capita									
Urban hukou	10.15***	1.529	0.038	7.6%	0.920**	0.380	0.007	1.8%	27.0%
Education	0.621***	0.169	0.017	3.4%	0.178***	0.045	0.011	2.7%	5.2%
Male	-0.052	1.551	0.000	0.0%	-1.272***	0.456	0.004	0.9%	-3.5%
Age	-0.049	0.074	0.000	0.1%	0.033	0.029	0.001	0.3%	-0.9%
Household size	-1.414***	0.315	0.027	5.4%	-0.469***	0.111	0.018	4.6%	7.8%
Single	-1.635	2.676	-0.001	-0.1%	-1.676	1.517	0.000	0.1%	-0.9%
Couple	-3.122	3.295	0.000	0.0%	0.641	0.692	0.001	0.4%	-0.9%
Other living arrangements	-6.298	5.986	0.001	0.2%	-0.646	0.471	0.001	0.3%	0.0%
Province dummies	Yes	Yes	0.035	7.0%	Yes	Yes	0.030	7.6%	4.3%
	R^2=0.189		Gini=0.503	ê: 76.5%	R^2=0.167		Gini=0.388	ê: 81.4%	100.0%

Notes: The results are based on the three-step inequality mediation decomposition method described in Sections 3.3. The OLS regression coefficients and standard errors of Steps 1, 2, and 3 are from estimations of Equations (16), (17), and (18), respectively. All monetary variables are scaled by 1/1000. The regression omits the binary variable 'households with adult children' as the baseline living arrangement. The decomposition includes eight province dummies, and we report their total contribution to inequality. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

The results in Panel A of Table 5 show that the urban-rural gap is the largest contributor to income inequality in old age (8.9% in total, Step 2). Education remains a mediator between the urban-rural gap and income inequality in old age, which is evidenced by the significant coefficient of education and the smaller contribution of the urban-rural gap in Step 3 than in Step 2. The contribution of the urban-rural gap and education to income inequality is larger when the cohort is old in 2015 (0.038 or 7.6%, and 0.017 or 3.4%) than when it is young in 2000 (0.007 or 1.8%, and 0.011 or 2.7%). The results support *Hypothesis 2*.

Because Sample C contains the same households in 2000 and 2015, we can directly compare the contributions of the socioeconomic determinants and income components by estimating the contribution to the change in the Gini coefficient, as shown in the last columns of Tables 5 and 6 following the comparison method from Zhong (2011). For example, we calculate the difference of the contribution of urban *hukou* to the Gini coefficient between 2000 and 2015 and the difference of the Gini coefficient between the two years. The ratio of the two differences is the contribution of the urban-rural gap to the change in the Gini coefficient from 2000 to 2015. This analysis is applicable only to a balanced panel (Zhong, 2011). *Hukou* status is the largest contributor (29.6% in total and 27.0% after controlling for education) to the change in inequality from young to old. Education is also an important contributor (5.2%). Regional inequality and household size also contribute to the increase in total income inequality (4.3% and 7.8%). Therefore, we conclude that the increase in inequality from young to old (i.e., the age effect) is primarily driven by the urban-rural gap.

The results in Table 6 show that pension income and income from work remain the two largest sources of income inequality in old age, contributing 30% and 64%, respectively. Consistent with the repeated cross-sectional analyses, income from work remains the main source of income inequality among the young, contributing over 80%. Pension income contributes 106% to the increase in inequality from young to old. The second largest contributor, income from work, contributes only 10.4% to the increase in inequality. Private transfers and subsidies reduce old-age income inequality by 3.5% and 19.1%, respectively. The results confirm the transmission process hypothesized in Section 2.2 in a longitudinal setup and over a cohort's life cycle.

Table 6: Decomposition (2) by income components in 2000 and 2015

Decomposition (2)	Concentration index (C_m)	Income share (S_m)	Contribution ($C_m \times S_m$)	Contribution in % ($C_m \times S_m$ / Gini)	Contribution to change in Gini 2000-2015
Panel A: Old households in 2015 (Sample C, N = 762)					
Pensions	0.638	0.236	0.151	29.9%	106.1%
Income from work	0.471	0.685	0.323	64.1%	10.4%
Capital income	0.527	0.034	0.018	3.6%	3.5%
Private transfers	0.258	0.037	0.010	1.9%	-3.5%
Subsidies	0.210	0.006	0.001	0.3%	-19.1%
Unspecified income	0.580	0.002	0.001	0.2%	4.3%
Household income per capita		1.000	Gini=0.503	100.0%	100%
Panel B: Young households in 2000 (Sample C, N=762)					
Pensions	0.636	0.046	0.029	7.6%	
Income from work	0.371	0.838	0.311	80.1%	
Capital income	0.454	0.032	0.014	3.7%	
Private transfers	0.457	0.030	0.014	3.5%	
Subsidies	0.694	0.033	0.023	6.0%	
Unspecified income	-0.178	0.021	-0.004	-0.9%	
Household income per capita		1.000	Gini=0.388	100.0%	

Notes: The results are estimated based on Equation (11).

5.5. Role of fragmented pension programs

The results presented in the previous sections indicate that the fragmented public pension system in China reinforces old-age income inequality by linking an individual's pension benefits to his/her socioeconomic origin and education. This section presents direct evidence on how the different pension programs reinforce income inequality in old age. We again apply the three-step *inequality mediation decomposition*, focusing on the occupational sector as a potential mediator. We use a binary variable indicating whether the individual works or worked in the formal sector to capture which public pension program the individual is eligible for. We focus on the old households in the longitudinal sample (Sample C), for which this information is available. As the starting point, we adopt the model used in Step 3 of Table 5 (Panel A), which includes both *hukou* status and education.[24]

The results of Step 1 in Table 7 show that household heads with original urban *hukou* and more formal education are more likely to work in the formal sector, which makes them more likely to participate in the high-benefit Employees' Basic Pension Program. The results of Step 2 and Step 3 show that the inequality between the fragmented pension programs, captured by the formal sector status, mediates and reinforces the contributions of the urban-rural gap and educational inequality to income inequality. In Step 2, the rural-urban gap and education contribute 12.5% to income inequality. In Step 3, the contribution of urban hukou and education is reduced to 10% when the formal sector is included in the model. The formal

[24] Due to missing values in the binary variable *formal sector*, the sample size is reduced to 678 in Table 7 compared to N = 762 households in Table 5.

sector itself contributes 4.8% to old-age income inequality in Step 3. Together, the urban-rural gap, educational attainment, and the fragmented pension system contribute 15% of income inequality in total, as shown in Step 3, which is larger than the 12.5% total contribution of the urban-rural gap and educational attainment in Step 2. The results provide direct evidence on how old-age income inequality is reinforced by the Chinese public pension system, which is fragmented by occupational sector.

Table 7: Decomposition (1) by socioeconomic characteristics in 2015

Decomposition (1)	**Old households** (Reduced Sample C, N=678)			
	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)
Step 1: Formal sector				
Urban hukou	0.328***	0.035	0.116	14.0%
Education	0.0278***	0.004	0.079	9.6%
Male	-0.126***	0.039	0.006	0.8%
Age	0.00381**	0.002	0.002	0.2%
Household size	-0.010	0.007	0.006	0.7%
Single	-0.085	0.067	0.000	0.0%
Couple	-0.030	0.071	0.000	0.0%
Other living arrangements	0.467***	0.147	0.008	1.0%
Province dummies	Yes	Yes	0.018	2.3%
	R^2=0.286		Gini=0.825	ê: 71.4%
Step 2: Household income per capita				
Urban hukou	12.52***	1.667	0.049	9.8%
Education	0.544***	0.182	0.014	2.7%
Male	-0.557	1.866	0.000	0.1%
Age	-0.107	0.081	0.002	0.4%
Household size	-1.317***	0.324	0.026	5.2%
Single	1.224	3.240	0.001	0.1%
Couple	-2.420	3.435	0.000	0.0%
Other living arrangements	-7.319	7.069	0.001	0.3%
Province dummies	Yes	Yes	0.036	7.3%
	R^2=0.213		Gini=0.500	ê: 74.2%
Step 3: Household income per capita				
Urban hukou	10.46***	1.763	0.041	8.2%
Education	0.369**	0.188	0.009	1.8%
Formal sector	6.286***	1.860	0.024	4.8%
Male	0.236	1.866	0.000	0.0%
Age	-0.131	0.081	0.002	0.5%
Household size	-1.256***	0.322	0.025	4.8%
Single	1.757	3.218	0.001	0.2%
Couple	-2.234	3.408	0.000	0.0%
Other living arrangements	-10.260	7.068	0.002	0.4%
Province dummies	Yes	Yes	0.035	6.9%
	R^2=0.226		Gini=0.500	ê: 72.3%

Notes: The results are based on the three-step inequality mediation decomposition method described in Section 3.3. The OLS regression coefficients and standard errors of Steps 1, 2, and 3 are from the estimations of Equations (16), (17), and (18), respectively. Household income per capita is scaled by 1/1000. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

6. Robustness tests

To verify the robustness of our main results, we conduct eight tests. We provide the results of these tests in Appendix D and discuss them below. Test 1 includes additional control variables in our inequality mediation decomposition (Table D1). Tests 2 to 5 involve sample variations (Tables D2, D3, and D4). Tests 6 and 7 consider additional or alternative inequality determinants or inequality measures in the decompositions (Table D3). Test 8 distinguishes the income from work of old and young individuals (Table D4).

In Test 1, we include additional control variables to analyze other potential determinants of old-age income inequality. The sample size is reduced due to missing values in these control variables. The additional control variables are *migrate to urban* (=1 if the household head's *hukou* status changed from originally rural to currently urban), *migrate to rural* (=1 if the household head's *hukou* status changed from originally urban to currently rural), *health problems and disability* (=1 if the household head has any health problem recorded in the CHNS, including having been sick or injured in the past four weeks, suffered from chronic or acute disease in the past four weeks, diagnosed of myocardial infarction or stroke in the past year, and/or had difficulties with performing at least one of the activities of daily living (ADLs) including bathing, dressing, eating, using the toilet, and continence),[25] *han ethnic group* (=1 if the household head belongs to *Hanzu*, the main Chinese ethnic group), *internet access* (=1 if the household head has access to internet), *business owner* (=1 if the household owns a home business), *medical insurance* (=1 if the household head has medical insurance), and *medical expense* (medical expenses of the household during the past four weeks). We also add the quadratic terms of *age* and *education* wherever applicable.

The results in Table D1 are consistent with the results in Table 3, supporting *Hypothesis 2* and the transmission process from socioeconomic origin to educational attainment and to income inequality in old age. The significant and positive coefficients of $education^2$ in Table D1 suggest that the impact of education on old-age income might be slightly nonlinear, with larger returns to higher education. The mediating effects of education between the urban-rural gap and income inequality remain unchanged.[26]

[25] In a separate test, we alternatively control for the self-reported health status as a determinant of inequality (Baeten et al., 2013). The CHNS records the self-reported health status until 2006, but not in later waves. We thus repeat the three-step inequality mediation decomposition analyses in 2006. The results support our conclusions and are available from the authors upon request.

[26] The access to internet also significantly contributes to the old-age income inequality (2.0% in Table D1). Internet access approximates the access to information and knowledge and can be potentially associated with financial literacy, which is an important determinant of old-age inequality in the U.S. (Lusardi et al., 2017). The CHNS does not contain any direct measure of financial literacy. We capture its effect partially by education and partially by internet access in our model.

In Test 2, we consider two subsamples of old households (Sample A): "pure old" households having only old members (Sample A1) and "mixed old" households having both old and young members (Sample A2). Three-quarters of old Chinese live with their adult children (see Table 1), whose income from work constitutes an important part of the household income. Therefore, the pattern of income inequality and its drivers might differ between "pure old" and "mixed old" households. We present the income inequality estimations for these two subsamples in Row 1 of Table D2. The average of the annual Gini coefficients over the sample period is 0.436 for "pure old" households and 0.433 for "mixed old" households. Both of these average Gini coefficients are larger than the average Gini coefficient of the young households (0.416). The same is true if we measure the inequality by Theil index and P90/P10 ratio. This confirms that income inequality is larger among the old than among the young in China.

We also repeat the regression-based inequality decompositions by socioeconomic characteristics and by income components with these two subsamples in 2015. The results in Table D3 show that the transmission process from the urban-rural gap to educational inequality and then to income inequality is significant and economically large in both subsamples. However, the size of the impact differs: the impact of the urban-rural gap and educational attainment is much stronger among the "pure old" households than among the "mixed old" households. This finding is consistent with our conceptual framework and main empirical results that the urban-rural gap and education contribute more to the income inequality of old households than young households. The results in Table D4 show that pension inequality is the main source of income inequality for "pure old" households (with a contribution percentage of 84.6%), while income from work is the main source of inequality for "mixed old" households (with a contribution percentage of 58.3%), and pensions are the second most important source of inequality (34.5%). This is because pensions comprise 74.2% of household income for "pure old" households ($S_m = 0.742$), while income from work has the largest income share for "mixed" households ($S_m = 0.634$). In both subsamples, income from work is less unequally distributed than the overall income ($C_m = 0.229$ vs. Gini = 0.405 for pure old, $C_m = 0.440$ vs. Gini = 0.479 for mixed old), while pensions are more unequally distributed than the overall income ($C_m = 0.461$ for pure old, $C_m = 0.586$ for mixed old). Therefore, we conclude that pensions are the primary cause of inequality among the old.

In Test 3, we repeat our analyses using all provinces included in the CHNS, that is, including the observations from Beijing, Shanghai, and Chongqing in 2011 and 2015. The results are consistent with our main results that the income of old households is more

unequally distributed than that of young households (Table D2) and that the transmission process from the urban-rural gap to education and to income inequality remains significant and important (Table D3).

In Test 4, we exclude young households with early retired members from Sample B. We define early retirement as receiving a pension before age 60. This alternative definition of young households excludes the impact of the pension system on the income inequality among the young, resulting in a cleaner comparison of the income inequality between the old (who mainly rely on pensions) and the young (who typically do not receive pensions). The results in Tables D2 and D3 confirm that the income of old households is more unequally distributed. The results also confirm the existence of the hypothesized transmission process.

In Test 5, we exclude old households in 2015 whose head's *hukou* status changed from urban to rural or rural to urban, and/or whose head's education changed during our sample period 1991-2015. This test addresses concerns about the endogeneity of original *hukou* status and education. The results in Tables D2 and D3 are consistent with our main results presented in Section 5.

In Test 6, we use as demographic characteristics the household average values for original *hukou* status, education, gender, and age instead of the household head's characteristics as in the main analyses. The results in Table D3 are consistent with those in Section 5.

In Test 7, we follow Shorrocks (1980) and Lin et al. (2010) to decompose the Theil index into its socioeconomic determinants. Different from the Gini coefficient decomposition, the decomposition of the Theil index is not regression-based and considers one determinant at a time without controlling for the impact of other determinants.[27] Therefore, it cannot be used to test mediating effects. Moreover, it is applicable only to discrete determinants. We thus convert education, age, and household size into categorical variables. We consider four age groups (<60, 60-69, 70-79, and >80), five education groups (no formal education, primary school, middle school, high school, and college or above), and two household size groups (one or two people and more than two people). The results in Table D3 confirm that the urban-rural gap and educational attainment remain the two largest contributors to income inequality among old households.

In Test 8, we distinguish the income from work of old and young household members in an old household using the individual income information in the CHNS. This separation

[27] Morduch and Sicular (2002) propose a unified regression-based method to decompose both the Theil index and the Gini coefficient. However, their method to decompose the Theil index cannot isolate the contribution of each determinant because the formula of the Theil index includes non-additive logarithmic terms, leading to "qualitatively different results" (pp. 104) from the Gini coefficient decomposition in their paper.

provides additional insights on the second largest contributor – income from work – to income inequality among old households. The results in Table D4 show that the income from work of the young is more unequally distributed and contributes more to the overall inequality than the income from work of the old. However, pension income remains the most unequally distributed income component and contributes most to the inequality among old households.

7. Conclusion

This paper studies income inequality in old age and its development over the life cycle. We theoretically model and empirically verify a transmission process from socioeconomic origins to income inequality in old age, explaining the increase in income inequality over the life cycle. We develop a three-period overlapping generation model that allows the quality of education to vary by initial rural-urban status and describes how this educational inequality transmits to wage inequalities in working age and pension inequalities in old age. We develop and apply a new three-step *inequality mediation decomposition* method to empirically identify this transmission process. Our empirical analyses are based on household panel data from the China Health and Nutrition Survey (CHNS) over the period 1991-2015.

First, we find that income is more unequally distributed among old households than among young households. This result holds for different measures of inequality, including the Gini coefficient, the Theil index, and the P90/P10 ratio, for all nine data waves between 1991 and 2015, and for a cohort of households that were young in 2000 and became old in 2015.

Second, we empirically identify the primary transmission process causing income inequality among old households in China, which starts from the urban-rural origins, via educational inequality, to wage inequality, and then to pension inequality, which is the main source of old-age income inequality. This finding is robust across repeated cross-sectional analyses and longitudinal analyses. Our finding is consistent with previous studies that find that educational inequality in China mainly comes from the urban-rural gap (Liu, 2005; Golley and Kong; 2018) and that improving educational equality would reduce (regional) income inequality (Fleisher et al., 2010). We add to the literature evidence of the complete mediating chain of how the urban-rural gap affects income inequality through education. In addition, we show that old-age income inequality is reinforced (compared to income inequality in young age) by the Chinese public pension system because it is fragmented by occupational sector. This explains the results that the urban-rural gap and education contribute more to old-age income inequality than to young-age inequality.

Third, we provide new empirical evidence quantifying the level of inequality among the old in China and the drivers of this inequality. We show that the urban-rural gap and educational inequality are the primary two causes of old-age income inequality, together explaining 11%-22% of the old-age income inequality during 2000-2015. Approximately one-third to half of the contribution of the urban-rural gap is mediated by educational inequality. Both the urban-rural gap and educational inequality also have a standalone impact on old-age income inequality. Their contributions are mainly to pension inequality, which accounts for 45%-60% of the old-age income inequality during 2000-2015 and explains almost all of the inequality increase within a cohort that aged from young to old during 2000-2015.

Our results have important policy implications. First, future pension reforms need to balance the trade-off between the adequacy of one pension program and the equality across different pension programs. A general increase in pensions for all pensioners will increase the income inequality among old households. Additionally, the public pension system in China is fragmented by occupational sector without cross-subsidies between programs. Thus, reforms should focus on increasing the funding and raising the pension benefits of the low-benefit Residents' Basic Pension Program. Additional private funding (e.g., contribution to an individual account) and government funding (e.g., dividends and shares from state-owned enterprises) would be required to finance higher pension payments for the Residents' Basic Pension Program. Currently, the majority of private and public funding is directed to the high-benefit Employees' Basic Pension Program (Fang and Feng, 2018), which increases income inequality among the old. The existing government subsidies to old households outside of the pension programs contribute very little to income equality and could be integrated into the premium contribution subsidies of the Residents' Basic Pension Program. The National Social Security Fund could also provide financial resources to co-fund the Residents' Basic Pension Program.

Second, our results show that the educational inequality between rural and urban areas contributes to income inequality over the life cycle and to pension inequality in old age. Thus, a long-term strategy to reduce old-age income inequality should involve higher public investments to improve educational quality in rural areas, for example, by increasing the salary of rural teachers and by formalizing the rotation program between urban and rural teachers. These measures can increase the marginal return of educational investment of rural residents. Public funding should be increased to ensure that rural residents receive high-quality vocational training and/or high-quality tertiary education. Such public investment can

moderate the educational inequality between rural and urban areas, increase the productivity of rural households, and eventually mitigate the old-age income inequality between urban and rural households.

Third, further *hukou* reforms supporting urbanization should allow migrants from rural to urban areas to obtain better education in their urban residence. This would allow migrants from rural areas to join the formal sector and participate in the high-benefit Employees' Basic Pension Program. A higher participation rate in the high-benefit program driven by urbanization and more equal pension payments between the two programs provide a basis to integrate the two public pension programs in the long run.

Finally, the private market for retirement income products should be developed to provide old households with alternative sources of funding. Wealth inequality is usually assumed to be larger than income inequality (Saez and Zucman, 2016). However, this might not be true given the remarkably high rate of home ownership in both urban and rural areas in China, which moderates wealth inequality (Zheng et al., 2019). Given that housing wealth accounts for 70%-80% of retired households' assets (Park and Shen, 2015), it is possible that reverse mortgages and other home equity release products (see, e.g., Hanewald et al., 2016) can play an important role in financing the income of China's growing old population and reducing income inequality. A recent study by Hanewald et al. (2019) confirms that there is a large interest in reverse mortgages in urban China despite an unsuccessful ongoing reverse mortgage pilot project. Zheng et al. (2019) show that annuitizing home equity to increase retirement income would reduce the income inequality among the old in China by 17.5%.

Our results regarding the role of initial socioeconomic differences and a fragmented pension system in the transmission process leading to higher old-age income inequality provide an evidence base for other developing countries. In several developing countries, the social security system is also fragmented (e.g., Mexico and India) or covers only a part of the population (e.g., Bangladesh and Sri Lanka) (World Bank, 2005). These countries face similar problems of very unequal pension income. If these pension programs with unequal coverage are subsidized by the government, then the subsidies are very likely to benefit the rich instead of the poor and cause large inequality in old age.

References

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. NJ: Princeton University Press.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2017). Global inequality dynamics: New findings from WID. world. *American Economic Review*, 107(5), 404-409.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (Eds.) (2018). *World Inequality Report 2018*, Cambridge: Belknap Press.

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy*, 97(6), 1447-1458.

Baeten, S., Van Ourti, T., and Van Doorslaer, E. (2013). Rising inequalities in income and health in China: Who is left behind? *Journal of Health Economics*, 32(6), 1214-1229.

Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.

Bönke, T., Schröder, C., and Schulte, K. (2010). Incomes and inequality in the long run: The case of German elderly. *German Economic Review*, 11(4), 487-510.

Bourguignon, F., Fournier, M., and Gurgand, M. (2001). Fast development with a stable income distribution: Taiwan, 1979–94. *Review of Income and Wealth*, 47(2), 139-163.

Burzynski, M., Deuster C., and Docquier F. (2019). Geography of skills and global inequality. *Journal of Development Economics*, in press.

Cai, H., Chen, Y., and Zhou, L. (2010). Income and consumption inequality in urban China, 1992–2003. *Economic Development and Cultural Change*, 58(3), 385–413.

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

Cao, K. H., and Birchenall, J. A. (2013). Agricultural productivity, structural change, and economic growth in post-reform China. *Journal of Development Economics*, 104(3), 165-180.

Chamon, M., Liu, K., and Prasad, E. (2013). Income uncertainty and household savings in China. *Journal of Development Economics*, 105, 164-177.

Chen, X., Huang, B., and Li, S. (2018). Population ageing and inequality: Evidence from China. *The World Economy*, 41(8), 1976-2000.

Chu, C., and Jiang, L. (1997). Demographic transition, family structure, and income inequality. *Review of Economics and Statistics*, 79(4), 665–669.

Coder, J., Rainwater, L., and Smeeding, T. (1989). Inequality among children and elderly in ten modern nations: The United States in an international context. *American Economic Review*, 79(2), 320-324.

Coibion, O., Gorodnichenko, Y., Kueng, L., and Silvia, J. (2012). Innocent bystanders? Monetary policy and inequality in the US. NBER Working Paper No. 18170. National Bureau of Economic Research.

Connelly, R., and Maurer-Fazio, M. (2016). Left behind, at-risk, and vulnerable elders in rural China. *China Economic Review*, 37, 140-153.

Costa, D. L. (1997). Displacing the family: Union army pensions and elderly living arrangements. *Journal of Political Economy*, 105(6), 1269-1292.

Crystal, S., and Shea, D. (1990). Cumulative advantage, cumulative disadvantage, and inequality among elderly people. *Gerontologist*, 30(4), 437-443.

Crystal, S., and Waehrer, K. (1996). Later-life economic inequality in longitudinal perspective. *Journals of Gerontology,* 51B(6), S307-S318.

Deaton, A. Gourinchas, P., and Paxson, C. (2002). Social security and inequality over the life cycle, in *The Distributional Effects of Social Security Reform*, Feldstein and Liebman (eds.), Chicago: University of Chicago Press.

Deaton, A., and Paxson, C. (1994). Intertemporal choice and inequality. *Journal of Political Economy*, 102(3), 437-467.

Deaton, A., and Paxson, C. (1997). The effects of economic and population growth on national saving and inequality. *Demography*, 34(1), 97-114.

Ding, H., and He, H. (2018). A tale of transition: An empirical analysis of economic inequality in urban China, 1986–2009. *Review of Economic Dynamics*, 29, 106-137.

Dong, Z., Tang, C., and Wei, X. (2018). Does population aging intensify income inequality? Evidence from China. *Journal of the Asia Pacific Economy*, 23(1), 66-77.

Doorslaer, E. V., and Koolman, X. (2004). Explaining the differences in income-related health inequalities across European countries. *Health Economics*, 13(7), 609–628.

Fang, H., and Feng, J. (2018). The Chinese Pension System. NBER Working Paper No. 25088. National Bureau of Economic Research.

Fields, G. S. (1998). Accounting for income inequality and its change. Department of Economics, working paper, Cornell University.

Fields, G. S., and Yoo, G. (2000). Falling income from work inequality in Korea's economic growth: Patterns and underlying causes. *Review of Income and Wealth*, 46(2), 139-159.

Fleisher, B. M., Li, H., and Zhao, M. Q. (2010). Human capital, economic growth, and regional inequality in China. *Journal of Development Economics*, 92, 215-231.

Friedman, M. (1957). *A theory of the consumption function. Princeton*, NJ: Princeton University Press.

Galor, O., and Moav, O. (2004). From physical to human capital accumulation: Inequality and the process of development. *Review of Economic Studies*, 71(4), 1001-1026.

Golley, J., and Kong, S. T. (2018). Inequality of opportunity in China's educational outcomes. *China Economic Review*, 51, 116-128.

Gregorio, J. D., and Lee, J. (2002). Education and income inequality: New evidence from cross-country data. *Review of Income and Wealth*, 48(3), 395-416.

Gustafsson, B., and Shi, L. (2002). Income inequality within and across counties in rural China 1988 and 1995. *Journal of Development Economics*, 69(1), 179-204.

Hanewald, K., Bateman, H., Fang, H., and Wu, S. (2019). Reverse mortgage demand in China. NBER Working Paper No. 25491. National Bureau of Economic Research.

Hanewald, K., Post, T., and Sherris, M. (2016), Portfolio Choice in Retirement - What is the Optimal Home Equity Release Product? *Journal of Risk and Insurance*, 83(2): 421–446.

Heathcote, J., Storesletten, K., and Violante, G. L. (2005). Two views of inequality over the life cycle. *Journal of the European Economic Association*, 3(2-3), 765-775.

Henretta, J. C., and Campbell, R. T. (1976). Status attainment and status maintenance: A study of stratification in old age. *American Sociological Review*, 41(6), 981-992.

Hurd, M. D., and Shoven, J. B. (1985). The distributional impact of social security, pp 193-215, in D. Wise (ed.) *Pensions, Labor and Individual Choice*, Chicago: University of Chicago Press.

Imai, K. S., and You, J. (2014). Poverty dynamics of households in rural China. *Oxford Bulletin of Economics and Statistics*, 76(6), 898-923.

Knight, J., Li, S., and Deng, Q. (2010). Education and the poverty trap in rural China: Closing the trap, *Oxford Development Studies*, 38(1), 1-24,

Lambert, S., Ravallion, M., and Van de Walle, D. (2014). Intergenerational mobility and interpersonal inequality in an African economy. *Journal of Development Economics*, 110, 327-344.

Lee, S., and Malin, B. A. (2013). Education's role in China's structural transformation. *Journal of Development Economics*, 101, 148-166.

Lerman, R. I., and Yitzhaki, S. (1985). Income inequality effects by income component: A new approach and applications to the United States. *Review of Economics and Statistics*, 67(1), 151-156.

Li, S., and Sicular, T. (2014). The distribution of household income in China: Inequality, poverty and policies. *China Quarterly*, 217, 1-41.

Li, J., Wang, X., Yuan, C., and Xu, J. (2018). The role of public pensions in income inequality among old households in China 1988–2013, Working paper, Hitotsubashi University.

Liao, P. J. (2013). The one-child policy: A macroeconomic analysis. *Journal of Development Economics*, 101, 49-62.

Lin, T., Zhuang, J., Yarcia, D., and Lin, F. (2010). Decomposing income inequality: People's Republic of China, 1990−2005. *Asian Development Review*, 25, 119-136.

Liu, Z. (2005). Institution and inequality: The hukou system in China. *Journal of Comparative Economics*, 33(1), 133-157.

Luo, Z., Wan, G., Wang, C., and Zhang, X. (2018). Ageing and inequality: The link and transmission mechanisms. *Review of Development Economics*, 22(3), 885-903.

Lusardi, A., Michaud, P., and Mitchell, O. S. (2017). Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, 125(2), 431-477.

Meyer, B. D., and Sullivan, J. X. (2013). Consumption and income inequality and the great recession. *American Economic Review*, 103(3), 178-83.

Mincer, J. A. (1974). Schooling and earnings. In NBER Chapters in Schooling, Experience, and Earnings, pages 41-63, National Bureau of Economic Research.

Ministry of Civil Affairs of China (2018). Statistical bulletin on 2017, Social Service Development, available at http://www.mca.gov.cn/article/sj/tjgb/.

Modigliani, F. (1966). The life cycle hypothesis of saving, the demand for wealth and the supply of capital. *Social Research*, 33(2), 160-217.

Molero-Simarro, R. (2017). Inequality in China revisited. The effect of functional distribution of income on urban top incomes, the urban-rural gap and the Gini index, 1978–2015. *China Economic Review*, 42, 101-117.

Morduch, J., and Sicular, T. (2002). Rethinking inequality decomposition, with evidence from rural China. *The Economic Journal*, 112(476), 93-106.

Oaxaca, R. L. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709.

Park, A., Shen, Y., Strauss, J., and Zhao, Y. (2012). Relying on whom? Poverty and consumption financing of China's old. In: *Ageing in Asia: Findings from New and Emerging Data Initiatives*, ed. J. P. Smith and M. Majmundar, Washington DC: National Academies Press.

Paul, S., Chen, Z., and Lu, M. (2017). Contribution of household income components to the level and rise of inequality in urban China. *Journal of the Asia Pacific Economy*, 22(2), 212-226.

Prus, S. G. (2000). Income inequality as a Canadian cohort ages: An analysis of the later life course. *Research on Aging*, 22(3), 211-237.

Radner, D. (1982). Distribution of family income: Improved estimates. *Social Security Bulletin*, 45(7), 13-21.

Ravallion, M., and Chen, S. (2004). China's (uneven) progress against poverty. *Journal of Development Economics*, 82(1), 1-42.

Saez, E., and Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131 (2), 519–578.

Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66(6), 467-482.

Santaeulàlia-Llopis, R., and Zheng, Y. (2018). The price of growth: Consumption insurance in China 1989-2009. *American Economic Journal: Macroeconomics,* 10(4): 1-35.

Saunders, P. (2007). Comparing poverty among older people in urban China internationally. *China Quarterly*, 190, 451-465.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, 48(3), 613-625.

Shorrocks, A., F. (1982). Inequality decomposition by factor components, *Econometrica*, 50(1), 193-211.

Song, Z., Storesletten, K., and Zilibotti, F. (2011). Growing like China. *American Economic Review*, 101(1), 196-233.

Tsakloglou, P. (1993). Aspects of inequality in Greece: Measurement, decomposition and intertemporal change: 1974, 1982. *Journal of Development Economics*, 40(1), 53-74.

United Nations (2017a). World population prospects: The 2017 revision. Department of Economic and Social Affairs and Population Division, available at https://population.un.org/wpp/Publications/.

United Nations (2017b). Economic inequalities in old age. Department of Economic and Social Affairs, available at https://www.un.org/development/desa/ageing/news/2017/06/briefing-paper-economic-inequalities-in-old-age/.

Wagstaff, A., Van, D. E., and Watanabe, N. (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics*, 112(1), 207-223.

Wan, G. (2004). Accounting for income inequality in rural China: A regression-based approach. *Journal of Comparative Economics*, 32(2), 348-363.

Wan, G., and Zhou, Z. (2005). Income inequality in rural China: Regression-based decomposition using household data. *Review of Development Economics*, 9(1), 107-120.

Wang, L., Béland, D., and Zhang, S. (2014). Pension fairness in China. *China Economic Review*, 28(1), 25-36.

World Bank (2005). Old-age income support in the twenty-first century: An international perspective on pension systems and reform, available at https://www.un.org/chinese/esa/ageing/pdf/HolzmannHinz_OldAgeSupport.pdf

Xie, Y., Zhang, X., Xu, Q., and Zhang, C. (2015). Short-term trends in China's income inequality and poverty: Evidence from a longitudinal household survey. China economic journal, 8(3), 235-251.

Young, A. (2013). Inequality, the urban-rural gap, and migration. *Quarterly Journal of Economics*, 128(4), 1727-1785.

Yu, J., and Zhu, G. (2013). How uncertain is household income in China. *Economics Letters*, 120(1), 74-78.

Zhang, J., and Xu, L. C. (2016). The long-run effects of treated water on education: The rural drinking water program in China. *Journal of Development Economics*, 122, 1-15.

Zhao, R., and Zhao, Y. (2018). The gender pension gap in China. *Feminist Economics*, 24(2), 218-239.

Zheng, W., Liu, Z., and Jia, R. (2019). How private sector participation improves retirement preparation: A case from China. *Geneva Papers on Risk and Insurance - Issues and Practice*, 44(1): 123-147.

Zhong, H. (2011). The impact of population ageing on income inequality in developing countries: Evidence from rural China. *China Economic Review*, 22(1), 98-107.

Zhuang, J., and Li, S. (2016). Understanding recent trends in income inequality in the People's Republic of China. Asian Development Bank Economics Working Paper Series, (489).

Appendices

Appendix A: Variable definitions

Table A1: Variable definitions

Variable	Definition	Source in CHNS
Household head	Self-reported household head	Variable a5
Urban *hukou*	Household registration status of the individual's parents (urban=1, rural=0)	Variables a8b1 and a20b (if missing, supplemented sequentially by birthplace, *hukou* in the earliest survey wave, and *hukou* of spouse and children)
Male	Gender of the household head (male=1, female=0)	Variable gender
Age	Age of the household head	Variable age
Education	Years of formal education completed in a regular school by the household head	Variable a11
Formal sector	Binary variable indicating occupational sector (=1 if working in government, public institutions, state-owned enterprise, large collective enterprise, or foreign invested enterprises)	Variable b6
Household size	Number of household members; includes household members surveyed and not surveyed	Variable hhsize (supplemented by number of surveyed household members if missing)
Living arrangement	Binary variables indicating different living arrangements: single, couple, living with adult children, and other living arrangements	Variable a5
Income from work	The sum of wages, net agricultural income, and net business income	Sum of wages, net agricultural income, and net business income
Wages	Income from being employed in the labor market	Variable hhnrwage
Net agricultural income	Total net income from agricultural activities including farming, fishing, gardening, and livestock	Sum of hhfarm, hhfish, hhgard, and hhlvst
Net business income	Total net income from household businesses	Variable hhbus
Pensions	Income from retirement pensions or retirement salaries	Variable hhretire
Capital income	Household income from leased land, asset rental, and rental income received from lodgers or boarders[a]	Sum of j2, j3, and j4
Subsidies	Subsidies for poverty and disability, firm-level nonwage compensation to workers, subsidies for gas, food, education, and housing, and allowances for children.	Sum of hhsub and j6
Private transfers	Transfers from children, parents, relatives, and friends who are not household members	Sum of j7a, j7b, j7c, j10, j9b, j9d, j9f, and j10b
Unspecified income	Other unspecified income components	j8

Notes: a. The CHNS provides no information about investment income from financial assets (e.g., interest on deposits or stock returns). However, the investment return from financial assets accounts for a very small fraction of household income in China (less than 2%, Zheng et al., 2019). Therefore, we do not expect our main results to change if investment income were included.

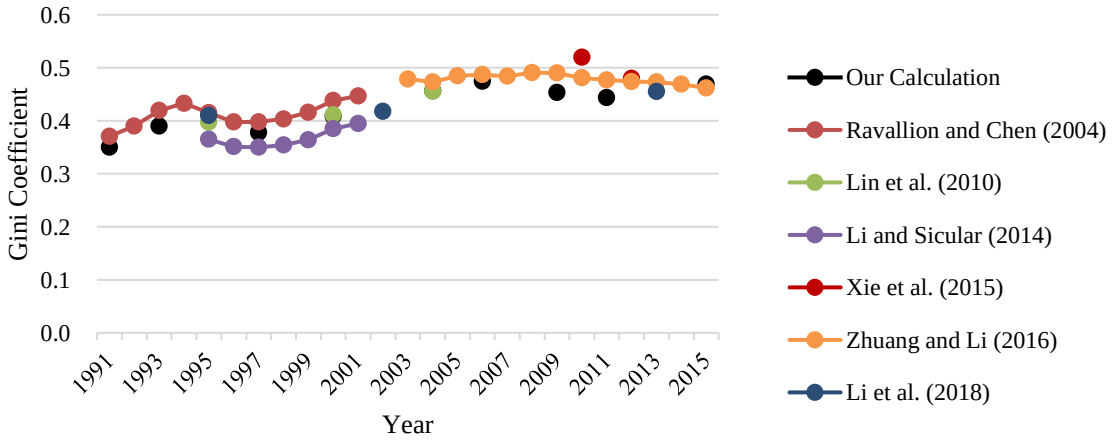
Appendix B: Sample validation

Table B1: Comparison of the Gini coefficients for the entire population in China

	Our paper	Ravallion and Chen (2004)	Lin et al. (2010)	Li and Sicular (2014)	Xie et al. (2015)	Zhuang and Li (2016)	Li et al. (2018)
Dataset	CHNS	CNBS	PSY	CHIP	CFPS	CNBS	CHIP
1991	0.350	0.371					
1992		0.390					
1993	0.390	0.420					
1994		0.433					
1995		0.415	0.397	0.365			0.410
1996		0.398		0.351			
1997	0.378	0.398		0.350			
1998		0.403		0.354			
1999		0.416		0.364			
2000	0.409	0.438	0.411	0.385			
2001		0.447		0.395			
2002							0.418
2003						0.479	
2004	0.456		0.456			0.473	
2005						0.485	
2006	0.475					0.487	
2007						0.484	
2008						0.491	
2009	0.453					0.490	
2010					0.520	0.481	
2011	0.444					0.477	
2012					0.480	0.474	
2013						0.473	0.455
2014						0.469	
2015	0.469					0.462	

Notes: CNBS refers to the National Bureau of Statistics of China. PSY refers to provincial statistical yearbooks published by provincial bureaus of statistics. CHIP refers to the China Household Income Project. CFPS refers to the China Family Panel Studies.

Figure B1: Comparison of the Gini coefficients for the entire population in China



Appendix C: Supplementary results

Table C1: Decomposition (3) by socioeconomic characteristics in 2015

Decomposition (3)	**Panel A: Old households** (Sample A, N=2,370)				**Panel B: Young households** (Sample B, N=1,843)			
	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)
Step 2			**Pensions**				**Income from work**	
Urban hukou	9.885***	0.453	0.166	21.3%	1.587*	0.954	0.001	0.2%
Male	1.176**	0.526	-0.001	-0.2%	-0.511	1.275	0.000	0.0%
Age	0.137***	0.020	0.023	2.9%	-0.313***	0.078	0.004	0.8%
Household size	-1.032***	0.139	0.041	5.3%	-1.655***	0.385	0.010	2.1%
Single	1.036	0.833	0.001	0.1%	-5.894**	2.491	0.001	0.3%
Couple	7.517***	0.698	0.057	7.3%	-1.608	1.626	0.000	-0.1%
Other living arrangements	2.452	1.717	0.000	0.1%	-4.254***	1.384	0.002	0.4%
Province dummies	Yes	Yes	0.042	5.4%	Yes	Yes	0.036	7.4%
	R^2=0.363		Gini=0.781	ê: 57.8%	R^2=0.084		Gini=0.493	ê: 89.0%
Step 3			**Pensions**				**Income from work**	
Urban hukou	7.187***	0.454	0.121	15.5%	-0.121	0.991	0.000	0.0%
Education	0.984***	0.057	0.102	13.1%	0.822***	0.141	0.017	3.5%
Male	-1.013**	0.511	0.001	0.2%	-0.618	1.263	0.000	0.0%
Age	0.238***	0.020	0.040	5.1%	-0.269***	0.078	0.003	0.7%
Household size	-1.031***	0.131	0.041	5.3%	-1.550***	0.382	0.010	2.0%
Single	1.453*	0.785	0.001	0.1%	-5.911**	2.469	0.001	0.3%
Couple	7.022***	0.658	0.053	6.8%	-2.024	1.613	0.000	-0.1%
Other living arrangements	1.993	1.617	0.000	0.1%	-4.680***	1.373	0.002	0.4%
Province dummies	Yes	Yes	0.029	3.7%	Yes	Yes	0.034	7.0%
	R^2=0.436		Gini=0.781	ê: 50.2%	R^2=0.101		Gini=0.493	ê: 86.4%

Notes: The results are based on the three-step inequality mediation decomposition method described in Section 3.3. The OLS regression coefficients and standard errors of Steps 2 and 3 are from estimating Equations (17) and (18), respectively. The results for Step 1 are the same as in Table 3. All monetary variables are scaled by 1/1000. The regression omits the binary variable *households with adult children* as the baseline living arrangement. The decomposition includes eight province dummies, and we report their total contribution to inequality. The last row in each decomposition shows the regression R^2, the Gini coefficient of household income per capita, and the contribution in percentage of the error term to income inequality. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C2: Decomposition (1) by socioeconomic characteristics for old households

Decomposition (1)		β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)
2011: Old households (N=2,090)					
Step 1	Urban hukou	3.032***	0.168	0.039	10.0%
Step 2	Urban hukou	5.928***	0.505	0.041	8.9%
Step 3	Urban hukou	3.851***	0.529	0.027	5.8%
	Education	0.685***	0.064	0.044	9.6%
2009: Old households (N=1,970)					
Step 1	Urban hukou	3.101***	0.173	0.038	9.3%
Step 2	Urban hukou	4.910***	0.501	0.034	7.2%
Step 3	Urban hukou	2.289***	0.517	0.016	3.4%
	Education	0.845***	0.062	0.066	14.1%
2006: Old households (N=1,806)					
Step 1	Urban hukou	3.202***	0.189	0.037	8.1%
Step 2	Urban hukou	4.036***	0.380	0.048	9.7%
Step 3	Urban hukou	1.888***	0.386	0.022	4.5%
	Education	0.671***	0.045	0.087	17.8%
2004: Old households (N=1,618)					
Step 1	Urban hukou	2.696***	0.189	0.026	6.1%
Step 2	Urban hukou	3.391***	0.326	0.042	9.1%
Step 3	Urban hukou	1.941***	0.329	0.024	5.2%
	Education	0.538***	0.041	0.059	12.8%
2000: Old households (N=1,429)					
Step 1	Urban hukou	2.734***	0.205	0.026	5.6%
Step 2	Urban hukou	2.019***	0.239	0.028	6.5%
Step 3	Urban hukou	1.284***	0.247	0.018	4.1%
	Education	0.269***	0.030	0.035	8.1%
1997: Old households (N=1,257)					
Step 1	Urban hukou	2.541***	0.212	0.019	3.7%
Step 2	Urban hukou	0.918***	0.171	0.012	3.0%
Step 3	Urban hukou	0.621***	0.179	0.008	2.0%
	Education	0.117***	0.023	0.014	3.7%
1993: Old households (N=1,120)					
Step 1	Urban hukou	1.897***	0.222	0.005	1.0%
Step 2	Urban hukou	1.112***	0.158	0.023	6.0%
Step 3	Urban hukou	0.923***	0.161	0.019	4.9%
	Education	0.100***	0.021	0.012	3.0%
1991: Old households (N=1,142)					
Step 1	Urban hukou	1.431***	0.220	-0.003	-0.7%
Step 2	Urban hukou	1.006***	0.121	0.025	7.1%
Step 3	Urban hukou	0.866***	0.122	0.022	6.1%
	Education	0.098***	0.016	0.013	3.8%

Notes: All monetary variables are scaled by 1/1000. All control variables and the constant in Equation (12) are included but not reported. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C3: Decomposition (1) by socioeconomic characteristics for young households

Decomposition (1)		β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)
2011: Young households (N= 2,217)					
Step 1	Urban hukou	2.834***	0.136	0.035	17.6%
Step 2	Urban hukou	3.285***	0.530	0.013	3.0%
Step 3	Urban hukou	0.886***	0.566	0.003	0.8%
	Education	0.846***	0.081	0.030	7.1%
2009: Young households (N= 2,377)					
Step 1	Urban hukou	2.659***	0.136	0.030	14.5%
Step 2	Urban hukou	3.792***	0.474	0.020	4.6%
Step 3	Urban hukou	2.018***	0.502	0.011	2.5%
	Education	0.667***	0.071	0.029	6.6%
2006: Young households (N= 2,470)					
Step 1	Urban hukou	2.894***	0.140	0.035	15.2%
Step 2	Urban hukou	2.804***	0.353	0.021	4.6%
Step 3	Urban hukou	1.118***	0.372	0.008	1.8%
	Education	0.583***	0.049	0.044	9.4%
2004: Young households (N= 2,612)					
Step 1	Urban hukou	2.552***	0.128	0.029	14.0%
Step 2	Urban hukou	3.293***	0.280	0.029	6.5%
Step 3	Urban hukou	2.056***	0.293	0.018	4.1%
	Education	0.485***	0.042	0.033	7.4%
2000: Young households (N= 2,771)					
Step 1	Urban hukou	2.572***	0.127	0.030	12.8%
Step 2	Urban hukou	1.489***	0.179	0.015	3.9%
Step 3	Urban hukou	0.711***	0.187	0.007	1.9%
	Education	0.303***	0.026	0.026	6.5%
1997: Young households (N= 2,447)					
Step 1	Urban hukou	2.705***	0.138	0.033	12.6%
Step 2	Urban hukou	1.012***	0.141	0.012	3.4%
Step 3	Urban hukou	0.460***	0.149	0.006	1.5%
	Education	0.204***	0.020	0.021	5.7%
1993: Young households (N= 2,229)					
Step 1	Urban hukou	2.436***	0.164	0.020	6.9%
Step 2	Urban hukou	0.746***	0.129	0.009	2.4%
Step 3	Urban hukou	0.439***	0.134	0.005	1.4%
	Education	0.126***	0.017	0.015	3.8%
1991: Young households (N= 2,387)					
Step 1	Urban hukou	2.317***	0.154	0.019	6.0%
Step 2	Urban hukou	0.663***	0.091	0.014	3.9%
Step 3	Urban hukou	0.400***	0.093	0.008	2.4%
	Education	0.113***	0.012	0.018	5.1%

Notes: All monetary variables are scaled by 1/1000. All control variables and constant in Equation (12) are included but not reported. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure C1: Decomposition (2) by income components

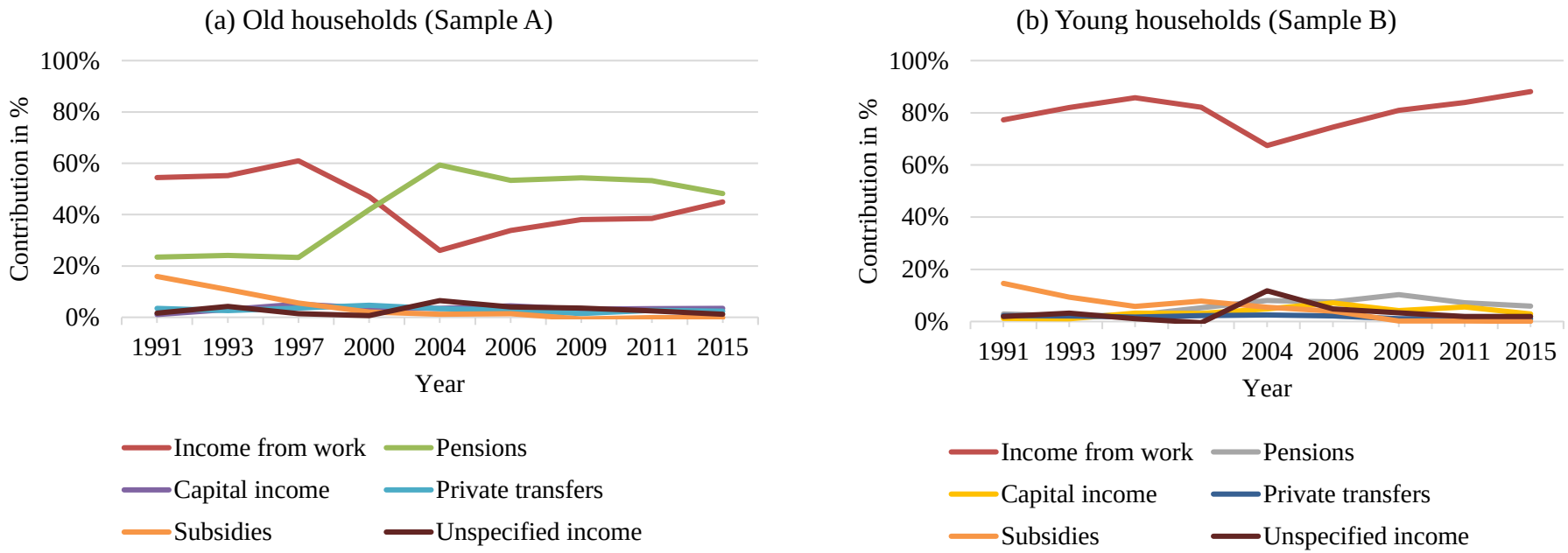


Table C4: Regressions of inequality on the fraction of old households (2000-2015)

	Gini coefficient	Theil index	P90/P10
Fraction of old households	0.060**	0.079*	0.041*
	(0.018)	(0.033)	(3.589)
Inequality of entire population	0.769*	1.075*	1.105*
	(0.259)	(0.367)	(0.366)
Constant	0.070	-0.065	-3.297
	(0.113)	(0.115)	(3.595)
No. of observations (waves)	6	6	6
R^2	0.927	0.908	0.955

Notes: Each column reports the results of a separate OLS regression given by $Inequality_t = \alpha + \beta_1 Fraction\ of\ old\ households_t + \beta_2 Inequality\ of\ entire\ population_t + \varepsilon_t$, where $Inequality_t$ denotes the income inequality among old households in year *t* in Sample C. Inequality is measured by the Gini coefficient in Column 1, the Theil index in Column 2, and the P90/P10 ratio in Column 3. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix D: Robustness tests

Table D1: Additional control variables: Decomposition (1) by socioeconomic characteristics in 2015

Test 1	**Old households** (Reduced Sample A, N=2,198**)**			
	β_k	Std. Dev.	Contribution ($C_k \times S_k$)	Contribution in % ($C_k \times S_k$ / Gini)
Step 1: Education				
Urban hukou	3.072***	0.176	0.034	10.2%
Male	2.245***	0.182	0.021	6.4%
Age	-0.0928***	0.008	0.025	7.7%
Age^2	-0.001	0.000	-0.001	-0.3%
Household size	0.024	0.049	0.000	0.0%
Single	-0.373	0.286	0.002	0.5%
Couple	0.137	0.242	0.000	0.1%
Other living arrangements	0.220	0.600	0.000	0.0%
Migrate to urban	-1.474***	0.391	0.000	-0.1%
Migrate to rural	2.422***	0.272	0.005	1.5%
Health problems and disability	-0.218	0.197	0.000	0.1%
Han ethnic group	0.170	0.258	0.000	0.1%
Medical expense	0.001	0.012	0.000	0.0%
Medical insurance	0.648	0.560	0.000	0.0%
Business ownership	-0.123	0.279	0.000	0.0%
Internet access	1.846***	0.263	0.015	4.5%
Province dummies	Yes	Yes	0.009	2.7%
	R^2=0.334		Gini=0.328	ê: 66.7%
Step 2: Household income per capita				
Urban hukou	9.933***	0.787	0.044	9.5%
Male	1.824**	0.811	0.001	0.2%
Age	0.004	0.035	0.000	0.0%
Age^2	-0.00543***	0.002	0.001	0.3%
Household size	-1.499***	0.218	0.023	4.9%
Single	-2.158*	1.276	0.000	0.0%
Couple	0.113	1.083	0.000	0.0%
Other living arrangements	-5.712**	2.677	0.001	0.2%
Migrate to urban	-6.080***	1.747	0.001	0.2%
Migrate to rural	6.701***	1.213	0.006	1.4%
Health problems and disability	-0.725	0.881	0.000	0.0%
Han ethnic group	1.690	1.154	0.001	0.3%
Medical expense	-0.051	0.054	0.000	0.0%
Medical insurance	3.091	2.499	0.000	0.0%
Business ownership	4.838***	1.246	0.002	0.4%
Internet access	7.179***	1.175	0.014	3.0%
Province dummies	Yes	Yes	0.024	5.0%
	R^2=0.201		Gini=0.466	ê: 74.8%
Step 3: Household income per capita				
Urban hukou	7.207***	0.823	0.032	6.9%
Education	0.850***	0.094	0.035	7.4%
Education^2	0.0683***	0.016	0.004	0.9%
Male	0.417	0.830	0.000	0.1%
Age	0.058	0.036	0.000	0.0%
Age^2	-0.00552***	0.002	0.001	0.3%
Household size	-1.502***	0.213	0.023	4.9%
Single	-1.848	1.249	0.000	0.0%
Couple	0.135	1.060	0.000	0.1%
Other living arrangements	-5.880**	2.619	0.001	0.2%
Migrate to urban	-4.594***	1.716	0.001	0.1%
Migrate to rural	4.433***	1.209	0.004	0.9%
Health problems and disability	-0.644	0.863	0.000	0.0%
Han ethnic group	1.384	1.130	0.001	0.3%
Medical expense	-0.050	0.053	0.000	0.0%
Medical insurance	2.631	2.445	0.000	0.0%
Business ownership	5.020***	1.219	0.002	0.4%
Internet access	4.865***	1.176	0.009	2.0%
Province dummies	Yes	Yes	0.020	4.4%
	R^2=0.236		Gini=0.466	ê: 71.4%

Table D2: Income inequality by household type

		Old households			Young households			Difference between old and young		
		Gini	Theil	P90/P10	Gini	Theil	P90/P10	Gini	Theil	P90/P10
Test 2	Pure old	0.436	0.315	15.7	0.416	0.283	10.6	0.020	0.032***	5.1***
	Mixed old	0.433	0.313	12.1				0.017**	0.030***	1.5*
Test 3	Including Beijing, Shanghai, and Chongqing	0.432	0.309	12.5	0.414	0.281	10.1	0.018***	0.029***	2.5***
Test 4	Excluding young households with early retired members	0.436	0.316	13.0	0.420	0.289	11.0	0.016**	0.026***	2.0**
Test 5	Excluding households with *hukou* and/or education changes	0.440	0.321	13.4	0.406	0.276	8.8	0.034***	0.045***	4.7***

Notes: *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively, based on the mean difference *t*-tests.

Table D3: Decomposition (1) by socioeconomic characteristics in 2015

Decomposition (1)		β_k	Std. Dev.	Contribution ($C_m \times S_m$)	Contribution in % ($C_m \times S_m$ / Gini)
Test 2: "Pure old" households (N=395)					
Step 1	Urban hukou	2.692***	0.386	0.036	11.4%
Step 2	Urban hukou	13.481***	1.769	0.074	18.2%
Step 3	Urban hukou	9.363***	1.773	0.051	12.6%
	Education	1.530***	0.222	0.076	18.7%
"Mixed old" households (N=1,975)					
Step 1	Urban hukou	2.711***	0.171	0.028	8.6%
Step 2	Urban hukou	7.202***	0.751	0.028	5.8%
Step 3	Urban hukou	4.895***	0.783	0.019	3.9%
	Education	0.851***	0.098	0.032	6.6%
Test 3: Old households (N=3,096), including Beijing, Shanghai, and Chongqing					
Step 1	Urban hukou	3.254***	0.135	0.045	14.5%
Step 2	Urban hukou	9.028***	0.590	0.045	10.1%
Step 3	Urban hukou	5.557***	0.624	0.028	6.2%
	Education	1.067***	0.077	0.047	10.4%
Test 4: Young households (N=1,674), excluding young households with early retired members					
Step 1	Urban hukou	2.085***	0.172	0.017	9.0%
Step 2	Urban hukou	3.952***	1.086	0.006	1.3%
Step 3	Urban hukou	2.317**	1.125	0.004	0.8%
	Education	0.784***	0.154	0.017	3.5%
Test 5: **Old households** (N=1,477), excluding households with *hukou* and/or education changes					
Step 1	Urban hukou	3.318***	0.188	0.045	13.5%
Step 2	Urban hukou	9.783***	0.838	0.057	12.4%
Step 3	Urban hukou	7.194***	0.908	0.042	9.1%
	Education	0.780***	0.115	0.031	6.7%
Test 6: Old households (N=2,370**)**, household average characteristics					
Step 1	% with urban hukou	2.973***	0.114	0.044	21.8%
Step 2	% with urban hukou	9.326***	0.756	0.042	9.0%
Step 3	% with urban hukou	3.232***	0.816	0.015	3.1%
	Average education	2.050***	0.130	0.070	14.8%
Test 7: **Old households** (N=2,370), Theil index decomposition					
	Urban hukou	N.A.	N.A.	0.031	8.7%
	Education group	N.A.	N.A.	0.032	8.8%
	Male	N.A.	N.A.	0.000	0.0%
	Age group	N.A.	N.A.	0.000	0.0%
	Household size group	N.A.	N.A.	0.019	5.2%
	Living arrangements	N.A.	N.A.	0.008	2.2%
	Province dummies	N.A.	N.A.	0.027	7.4%

Notes: All monetary variables are scaled by 1/1000. All control variables and the constant in Equation (12) are included but not reported. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table D4: Decomposition (2) by income components in 2015

Decomposition (2)	Concentration index (C_m)	Income share (S_m)	Contribution ($C_m \times S_m$)	Contribution in % ($C_m \times S_m$ / Gini)
Test 2: "Pure old" Households (N = 395)				
Pensions	0.461	0.742	0.342	84.6%
Income from work	0.229	0.177	0.040	10.0%
Household income per capita		1.000	Gini=0.405	100.0%
"Mixed" Households (N = 1,975)				
Pensions	0.586	0.282	0.165	34.5%
Income from work	0.440	0.634	0.279	58.3%
Household income per capita		1.000	Gini=0.479	100.0%
Test 8: Old households (N = 2,288), separating income from work of the young and the old				
Pensions	0.586	0.388	0.227	48.8%
Income from work (the old)	0.355	0.216	0.076	16.4%
Income from work (the young)	0.417	0.313	0.131	28.0%
Household income per capita		1.000	Gini=0.466	100.0%

Notes: Other income components are included in the decomposition but not reported.