

8-14-2017

Micro-Data Evidence on Family Size and Chinese Household Saving Rates

Steven Lugauer

University of Notre Dame, slugauer@nd.edu

Jinlan Ni

University of Nebraska at Omaha, jni@unomaha.edu

Zichao Yin

Southwestern University of Finance and Economics, yzc@swufe.edu.cn

Follow this and additional works at: <https://digitalcommons.unomaha.edu/econrealestatefacpub>

Part of the [Economics Commons](#)

Recommended Citation

Lugauer, Steven; Ni, Jinlan; and Yin, Zichao, "Micro-Data Evidence on Family Size and Chinese Household Saving Rates" (2017). *Economics Faculty Publications*. 58.

<https://digitalcommons.unomaha.edu/econrealestatefacpub/58>

This Article is brought to you for free and open access by the Department of Economics at DigitalCommons@UNO. It has been accepted for inclusion in Economics Faculty Publications by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.

Footer Logo

Chinese Household Saving and Dependent Children: Theory and Evidence⁺

Steven Lugauer^a

Jinlan Ni^{b*}

Zhichao Yin^c

^aSteven Lugauer. Department of Economics, Gatton College of Business and Economics, University of Kentucky, 550 S. Limestone, Lexington, KY, 40506, USA. Email: steven.lugauer@uky.edu.

^bJinlan Ni (*corresponding author). Department of Economics, College of Business Administration, University of Nebraska at Omaha, Omaha NE 68182; Institute of Urban Development at Nanjing Audit University, Nanjing, China. Email: jni@unomaha.edu.

^cZhichao Yin. The School of Finance, Capital University of Economics and Business, Beijing, 100070, China. Email: yzc@cueb.edu.cn.

Abstract

This paper examines the impact of family size on household saving. We first study a theoretical life-cycle model that includes finite lifetimes and saving for retirement and in which parents care about the consumption by their dependent children. The model implies a negative relationship between the number of dependent children in the family and the household saving rate. Then, we test the model's implications using new survey data on household finances in China. We use the differential enforcement of the one-child policy across counties to address the possible endogeneity between household saving and fertility decisions in a two-stage least squares Tobit regression. We find that Chinese families with fewer dependent children have significantly higher saving rates. The data yields several additional insights on household saving patterns. Households with college-age children have lower saving rates, and households residing in urban areas have higher saving rates and a lower ratio of education expenditures to income. However, having an additional child reduces saving rates more for households in urban areas than in rural areas. Our regressions also indicate that saving rates vary with age and tend to be higher for households with more workers, higher education, better health, and more assets.

Keywords: China, Household Saving, Demographics, Overlapping Generations

JEL Codes: D12, D91, E21, F63, J11, O12

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1. Introduction

This paper presents evidence supporting the hypothesis that the decrease in dependent children within Chinese families has led to an increase in household saving rates. We estimate the relationship between saving and family size by applying standard regression techniques to data from a new household-level survey (the China Household Finance Survey). The strong response of fertility rates to family-planning policies (e.g. the one-child policy) in China allows us to address the potential endogeneity between saving and birth decisions. Specifically, we instrument for the number of dependent children in the household with the county level number of births because enforcement of the family-planning policies has varied across geographic regions. Thus, we regress household saving on the instrumented number of children (and additional control variables) at the household level via a two-stage Tobit regression. Our main finding is that families with fewer dependent children save significantly more.

Uncovering the determinants of household saving is, of course, an important topic in general, and Chinese saving, in particular, has been receiving considerable attention. China's household saving rate has exploded in recent years, and this saving has helped create investment led growth in China. Excess Chinese saving has flowed towards safe assets in developed countries. Hence, policymakers both in China and abroad would like to understand the factors behind the high saving rate. Looking forward, accumulated household assets may help China cope with its rapidly aging population.

The life-cycle hypothesis is a leading candidate for explaining China's high household saving rate.¹ Modigliani and Cao (2004) was the first paper empirically showing the correlation between China's age structure and saving rate in the aggregated time series data. Curtis, Lugauer, and Mark (2015) develop a structural life-cycle model of household saving decisions to illustrate the theoretical connections between demographics and savings, and, through a series of model simulations, they show that the demographic effect on China's aggregate saving rate is quantitatively large. We build off Curtis, Lugauer, and Mark (2015) by examining the implications of a life-cycle based model for saving behavior in the cross section (i.e. micro data). We focus on the main implication from the model: household saving decreases with family size.

In Section 2, we present a simplified version of the model from Curtis, Lugauer, and Mark (2015). The key model ingredients include finite lifetimes, saving for retirement, and that parents care about the consumption of their dependent children. The explicit valuation of children's consumption enters the parental utility function with the functional form from Barro and Becker (1989). The structural model has stark implications for the relationship between the number of dependent children in the family and household saving behavior. Therefore, our empirical regressions act as a test of the life-cycle hypothesis of household saving behavior.

The life-cycle model motivates our basic research question as to whether the number of children in a household affects saving decisions, but the data allow us to examine additional control variables, including the characteristics of the household head (age, education, risk aversion, and health), the number of elderly people in the home, and household assets. Our empirical strategy, then, is to leverage the policy driven (and therefore plausibly exogenous) differences in the number of children to estimate the effect of family size on household saving, while controlling for these other factors. The key identification assumption is that the enforcement of family-planning policies (as measured by county-level birth rates) affects household saving decisions only through the fertility channel.

The empirical results support the implications for saving from the structural model. Household saving is decreasing in family size, as measured by the number of dependent children. The estimated coefficients are large and statistically significant at the 1 percent level. Thus, our main finding is in line with the model's prediction that fewer children increases saving. Note, though, a few related theories (in particular, Choukhmane, Coeurdacier, and Jin (2017) and Imrohorglu and Zhao (2017)) exist that also predict the negative impact of children on household saving rates. We discuss these alternative explanations further below.

In our regressions, the coefficient estimates for the control variables have the expected signs. We find that households with higher education levels save more on average. Poor health is associated with lower saving rates. Families with more workers have higher savings, and household saving varies by age.

The data allow us to examine saving behavior along several additional dimensions. First, we examine the relationship with nonlinear specifications in terms of the first born, second, and third or more. We find the largest effect from the first-born child. We also examine the education expenditure data. Households with college age children spend about twice as much on education compared to households with children below college age. Households with college age children have significantly lower saving rates. In addition, for the group of households who spend more on their children, having one less child increases their saving by more than for households spending less on their children. These results are noteworthy, since the structural model also implies different saving behavior due to the parent's preferences over providing for their children.

The households in urban areas spend an average of 1/3 more on education; however, urban households still have higher saving rates than rural households due to a lower ratio of education expenditure to income. Interestingly, having an additional child reduces the saving rate more for households in urban areas than in rural areas. We find a similar pattern when comparing households with at least one parent employed in a state owned enterprise or government agency versus households with no family members employed in the public sector.

The rest of the paper is as follows. Section 2 presents a 2-period structural life-cycle model. We use the model to motivate our empirical, reduced form, regressions. Section 3 summarizes the data. Section 4 details the regressions and our identification strategy. Section 5 presents the main empirical results, which are the regression estimates, along with additional experiments. Section 6 concludes.

2. A Life-Cycle Model of Household Saving

This section presents a structural life-cycle model of household saving decisions. The model represents a simplification of the framework employed in Curtis, Lugauer, and Mark (2015) and Curtis, Lugauer, and Mark (2017) to study the effect of demographic changes on aggregate household saving rates over time across several countries.² The model motivates our reduced form empirical regressions based on Chinese household-level data. The main take-away from the model is that household saving decreases with family size.

Households make consumption and saving decisions taking interest rates and wages as given.³ Labor supply is inelastic, and family size (demographics) is exogenous. We think this assumption is reasonable given the evolution of the Chinese economy in regards to family-planning policies. Plus, this assumption maps into our empirical identification strategy.

Generations overlap, but each agent lives for only 3 periods. In the initial period of life, however, agents are dependent children and make no decisions. The main departure from a standard 2-period utility maximization problem is the inclusion of children's consumption in the parental utility function (via Barro-Becker preferences) in the middle period of life. Agents retire in the final period of life and no longer support children.

Budget Constraints

Let C_1 be parental consumption in the first period of the agent's decision making life (i.e. when the agent is no longer a dependent child). The household has n dependent children, each of whom consume in the amount C^c . As a dependent child (prior to period 1), the agent makes no choices and simply consumes what is provided by his or her parents. Thus, during period 1, parents choose their dependent children's consumption C^c , their own consumption C_1 , and saving S to take into the next period. They receive an endowment of income I , which can be interpreted as the real value of total household income net of taxes and transfers.

The budget constraint in the first period of life is

$$nC^c + C_1 + S = I, \quad (1)$$

where households begin their economic lives with no assets.

In period 2, all agents retire. They no longer support their (now grown) children. The budget constraint faced by the retired is

$$C_2 = S, \quad (2)$$

where C_2 is period 2 consumption, the real return on saving equals zero, and asset holdings are required to be non-negative.⁴

Preferences

During the first period, in which parents make decisions for children, household utility takes a Barro and Becker (1989) functional form

$$u_1(C^c, C_1) = (1-\sigma)^{-1} [\mu n^\eta (C^c)^{1-\sigma} + (C_1)^{1-\sigma}],$$

where $\mu < 1$ and $\eta < 1$ determine the degree to which parents care for their children and $\sigma > 0$ is the inverse of the elasticity of intertemporal substitution.

In period 2, utility depends only on the agent's consumption, C_2 . The period utility function for agents in the retirement period is

$$u_2(C_2) = (1-\sigma)^{-1} (C_2)^{1-\sigma}.$$

Let $0 < \delta < 1$ be the discount factor. Then, the lifetime utility problem is to choose C^c , C_1 , C_2 , and S in order to maximize Equation (3)

$$U = (1-\sigma)^{-1} [\mu n^\eta (C^c)^{1-\sigma} + (C_1)^{1-\sigma}] + \delta (1-\sigma)^{-1} (C_2)^{1-\sigma}, \quad (3)$$

subject to the budget constraints given in Equations (1) and (2).

The Household Saving Decision

The household problem admits an analytical solution.⁵ The agent's optimal choice for saving as a function of the underlying parameters and the exogenously given household income and family size is

$$S = \frac{l\delta^{\frac{1}{\sigma}}}{1 + \delta^{\frac{1}{\sigma}} + \mu^{\frac{1}{\sigma}} \frac{\eta + \sigma - 1}{\sigma}}.$$

Dividing by l and rearranging gives a relatively simple expression for the household saving rate (S/l). The Appendix contains the derivation of Equation (4).

$$\text{SavingRate} = \delta^{\frac{1}{\sigma}} \left[1 + \delta^{\frac{1}{\sigma}} + (\mu n^{\eta + \sigma - 1})^{\frac{1}{\sigma}} \right]^{-1}. \quad (4)$$

The model has stark implications for the qualitative relationship between saving and the number of dependent children. As long as $\eta + \sigma > 1$, the household's saving rate is decreasing in n . Quantitatively, the effect can be big, given a large change in family size. For example, plugging in a δ close to unity, σ equal to 1.5, μ equal to 0.65, and η equal to 0.76 (all values used in the literature) and decreasing n from 3 to 1 increases the saving rate by about 10 percentage points. In the regressions below, we test this relationship and find strong empirical support for the life-cycle model of household saving.

There exist a few alternatives to the theory encapsulated in Equation (4) as to why household saving depends on the number of children in China. For example, Choukhmane, Coeurdacier, and Jin (2017) point out that children are an investment because future transfers from children to elderly parents can be substantial, and these transfers take the place of other old-age support such as public pensions (also see Coeurdacier, Guibaud, and Jin (2014) and Imrohoroglu and Zhao (in press)). Imrohoroglu and Zhao (2017) argue that children may also act as insurance against future health shocks. Clearly, all these mechanisms have merit and are inter-related; however, their implications could differ, especially for policies aimed at reducing saving. The model presented above does not take a strong stand on why parents care for their children; the reasons take many forms. What Equation (4) does indicate is that household saving rates depend on the number of children.

Equation (4) also implies that the saving rate increases with δ and decreases with μ and η , as one might expect.⁶ Recall that $\mu < 1$ and $\eta < 1$ determine the degree to which parents care for their children. These parameters also can be interpreted as the household's ability to provide for their children. In the empirical part below, we examine data on household education expenditures that may shed light on this relationship. We find that households with college aged children (age 16 to 25 and not working) save less than households with younger children. The parents likely spend more on their older children, especially for education. We also compare households residing in urban and rural areas and people working at public sector and nonpublic sector jobs, where parents have different abilities to give care to their children. For instance, families in urban areas may have greater options to spend more on children.

3. Data and Descriptive Statistics

We base the empirical analysis on data from the China Household Finance Survey (CHFS) conducted by the Southwestern University of Finance and Economics in China. The CHFS collects detailed information biennially on households' demographic characteristics, assets and liabilities, insurance and social welfare, and income and expenditures. The survey is new, and we primarily use information from 2013. The survey was also conducted in 2011, but the 2011 sample is considerably smaller.⁷ We use the 2011 sample to construct a panel data set (by matching households across the two samples) in a robustness check reported below.

The CHFS data set is particularly suited to our purposes because it contains information on a large sample of families with young children and covers most of China, both urban and rural areas. In addition, the survey asks detailed questions about income, expenditures, and family demographics. There are two other widely used surveys for micro-level research: the China Health and Retirement Longitudinal Study (CHARLS) and the Urban Household Survey (UHS). However, neither of these data sets have the same advantages as the CHFS.

CHARLS consists of a high quality nationally representative sample, but it focuses on individuals aged 45 or older. Therefore, the CHARLS data has been used to analyze questions focusing on aging in China. It does not fit our research question, as we want a large sample of individuals under age 45 who have younger children. The UHS, a survey conducted by the National Bureau of Statistics, contains information on income, consumption, and demographic characteristics of urban households. The UHS focuses on nine provinces and is not as nationally representative as our CHFS data. Furthermore, the UHS data comes from urban families only. In contrast, the CHFS uses a three-stage stratified sampling method and covers 29 provinces and autonomous regions (except Tibet, Xinjiang,

Hong Kong, Macao and Taiwan). The CHFS also has a low non-response rate (10.9% in 2013). As summarized in Gan et al. (2014), the overall representativeness of the CHFS is excellent, and it fits our research purpose well.

After removing outliers and households with missing data, the 2013 CHFS survey provides a sample of 21,861 households from 1,048 different communities in 262 counties. Survey participation was randomized; so, again, the data are highly representative in terms of geographic location and economic development. Matching households from 2011 and 2013 reduces the sample size to 13,120.

Table 1: Summary Statistics

	2013		2011-2013 Panel	
	Mean	Std Dev	Mean	Std Dev
<i>SavingRate</i>	0.283	0.443	0.253	0.448
<i>Children</i>	0.744	0.871	0.791	0.871
<i>Elders</i>	0.020	0.138	0.016	0.125
<i>Workers</i>	1.822	1.267	1.952	1.276
<i>Age</i>	52.050	14.470	51.380	13.940
<i>Education</i>	9.616	4.246	9.530	4.172
<i>Health</i>	0.493	0.500	0.238	0.426
<i>Risk Averse</i>	0.676	0.468	0.642	0.479
<i>Risk Prefer</i>	0.106	0.308	0.116	0.320
<i>Asset (10K RMB)</i>	59.482	129.021	52.955	112.006
<i>Debt (10K RMB)</i>	2.496	20.103	2.619	14.346
<i>Public</i>	0.185	0.389	0.186	0.389
<i>Rural</i>	0.312	0.464	0.361	0.480
<i>Observations</i>	21,861		13,120	

Table 1 reports descriptive statistics. The main variable we wish to explain, *SavingRate*, is defined as one minus the ratio of total household consumption to total income. The average saving rate in our sample is 28 percent, which is consistent with both the available macro data and micro data used in other studies (see Zhou 2014 and Banerjee *et al.* 2014, for example). The main independent variable (*Children*) is the number of dependent children in the household aged 16 and below (reporting no labor income) plus college students with ages between 16 and 25. We assume parents continue to support college students, which is typical in China. On average, households contain less than one dependent child. Both *SavingRate* and *Children* exhibit large variation. The variation in the number of children might seem surprising given that the one-child policy has been in effect for over 30 years. However, enforcement of the policy varies from place to place. We will leverage this policy driven difference in birth rates in our two-stage regression approach.

The remainder of Table 1 lists statistics for the variables employed as controls in the regressions. *Elders* is the number of elder persons (age 45+) in the home without a job. *Workers* is the number of family members currently employed. *Age* is the age of the household head. *Education* is the number of years the household head attended school. Variable *Health* is a dummy concerning the self-reported health condition of the household head. If the head has bad health, then *Health* equals 1, and it equals 0 otherwise. *Risk Averse* is a dummy equal to 1 if the respondent is unlikely to invest or only invests in projects with little risk and small expected returns. Similarly, *Risk Prefer* is a dummy equal to 1 if the respondent is likely to invest in high-risk, high-return projects. *Asset* is the total housing assets (house value). *Debt* is the total housing related debt.

We also use two additional dummy variables, *Public* and *Rural*. The variable *Public* is 1 if at least one of the parents in the household is an employee at a state-owned enterprise or government agency in an urban area. The one child policy applies more strictly to this group. Less than 20 percent of the sample works in the public sector. The variable *Rural* equals 1 if the household resides in a rural area. The *Rural* group differs from the urban group because of the institutional Hukou system, the social security systems, employment opportunities, and for many other reasons. About 30 percent of the sample comes from rural areas.

Each variable has a similar mean in the 2013 and matched sample. On average, a little less than two people work per family. Assets per household exceed 500,000 RMB, with little debt. The average household head is 52 years old with a middle school education, and households tend to be risk averse.

4. Regression Equation and Identification Strategy

In our simple life-cycle model, the household saving rate depends on the number of dependent children in the household.⁸ Thus, we estimate the relationship between family size and saving in the data by running regressions based on Equation (5).

$$SavingRate_i = \alpha + \beta * Children_i + P + X'_i \pi + \varepsilon_i. \quad (5)$$

The *SavingRate* and *Children* variables are as previously defined for each household *i*. The vector X_i includes all the control variables listed in Table 1.⁹ The household head's age enters as a quadratic, and *P* stands for a complete set of province fixed effects.¹⁰

Our primary interest revolves around estimating the relationship between saving and the number of children, captured by β in Equation (5). Figure 1 and Figure 2 display the raw data underlying the analysis. Figure 1 shows the percentage of families with 0, 1, 2, 3, or 4+ dependent children. About 47 percent of households have no dependent children, 37 percent have one dependent child, 13 percent have two dependent children, and a little over 3 percent of families have three or more dependent children. Figure 2 displays the simple relationship between the number of dependent children and the household saving rate. Household saving rates monotonically decrease with the number of dependent children. Families with no children to support save over 30% of their income, on average; those with more than three children save less than 20 percent.

Families with fewer children could differ from those with more children along some dimension. Our regressions attempt to account for the relevant differences by including the set of control variables, X_i . However, even with all the controls, fertility decisions could be endogenously determined with regards to savings. Thus, we use an instrumental variable, a linear combination of the county level birth rates at 2000 and 2010, to address the potential endogeneity in a two-stage regression approach.¹¹

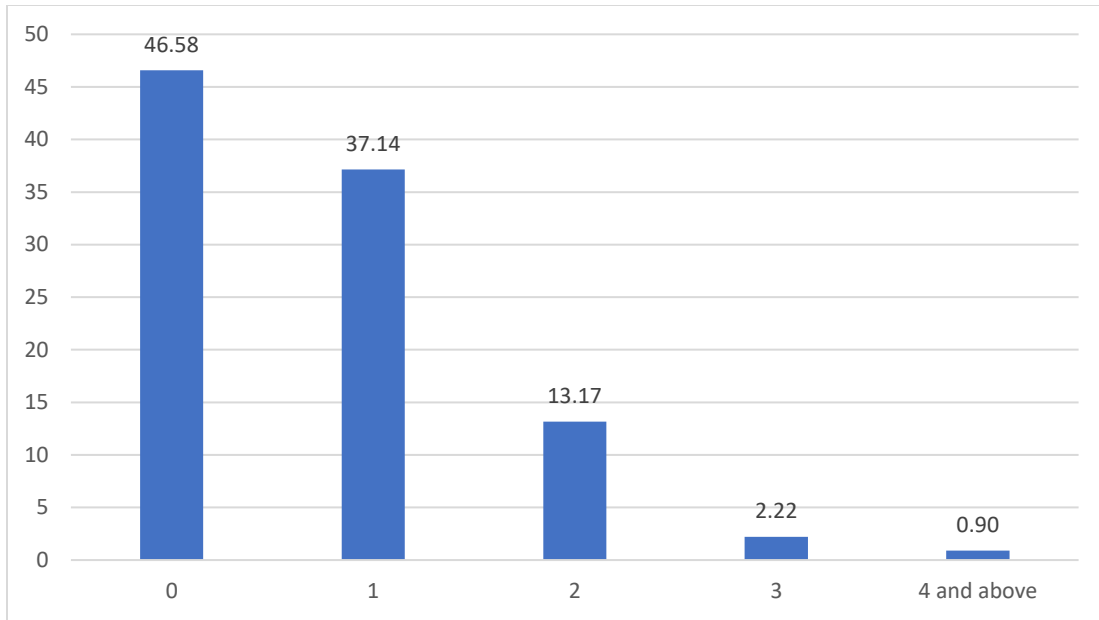


Figure 1: Number of Dependent Children in the Household (% of Households)

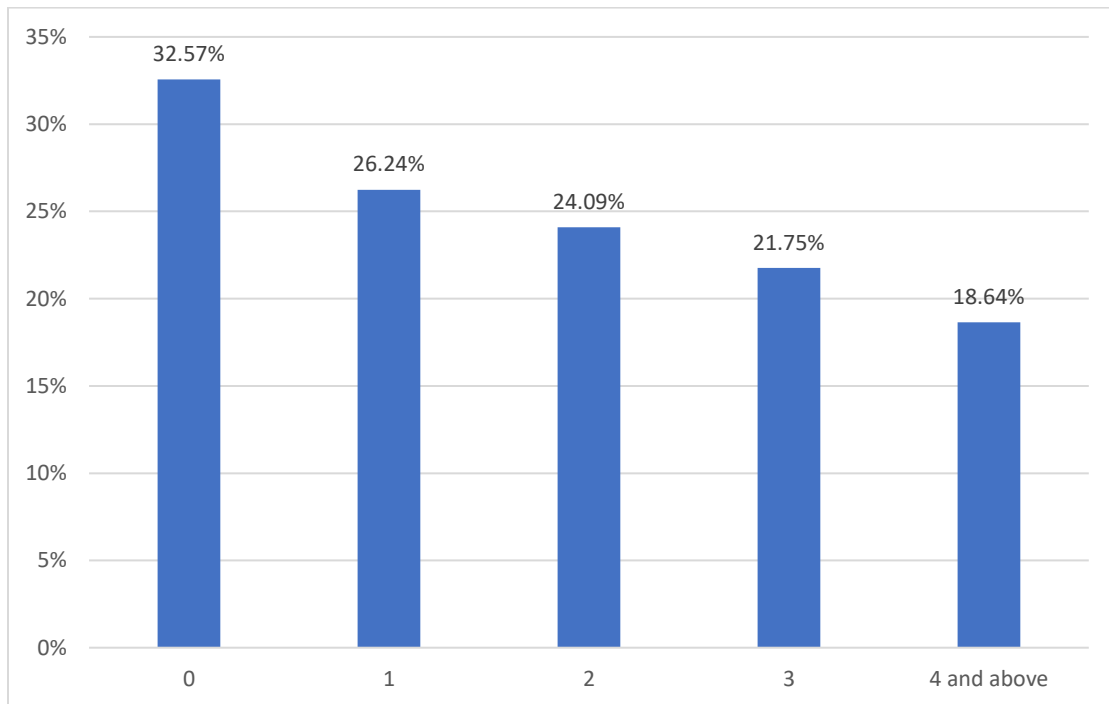


Figure 2: Household Saving Rate by Number of Dependent Children

Equation (6) is the first stage regression equation,

$$Children_i = \alpha + \gamma * County_i + P + X'_i \pi + \varepsilon_i, \quad (6)$$

where the instrument *County* equals the number of births per 1000 people (calculated separately for each county as the average from the 2000 and 2010 census data) in the current county of residence for household *i*. The data come from the Chinese National Bureau of Statistics, which conducts the national census every ten years. Econometrically, a linear combination of multiple instruments often serves as the best one (Wooldridge, 2013). In addition, the timing between 2000 and 2010 is close to the average birth year for the dependent children (average age of 11) in our sample. All the other variables are as defined above.

The county birth rate is a good instrument for a household's number of dependent children because it likely satisfies the two validity conditions. The number of dependent children within a family surely depends on the birthrate within their county. In other words, the first stage is strong. When implementing the policy in 1979, the Chinese government provided economic incentives, such as a monthly subsidy, to encourage compliance. The government also imposed severe punishments, such as dismissal from work (especially from state-owned enterprises) and substantial fines to restrict female fertility. The policy has greatly reduced family size for the whole country, but the effect has not been uniform.

Importantly for our approach, the enforcement of the population control policies has varied from place to place. Figure 3 summarizes the large variation in birth rates across counties. Some counties have birth rates four or five times higher than others do. The strength of our identification of the parameter β depends in part on the extent to which the variation in enforcement of the one-child policy is exogenous to household-level saving decisions.

Several reasons exist for why the one-child policy has had differential effects across counties; none of which seem to be directly related to household saving. Local "fertility czars" and other officials have had a fair amount of autonomy in how to enforce the policy. Methods have ranged from brutal (allowing sex-selective abortions, coerced abortions, and infanticide) to lenient. Fines have been ignored in rural areas, since few families could afford them. Also, farmers with a girl or sickly child have been allowed additional children. Enforcement has tended to be stricter in urban areas; however, even across cities there has been variation. For instance, some areas have allowed additional children if both parents work in high-risk occupations, or are minorities, or if both parents (and sometimes only one) are single children themselves, while other areas have not.

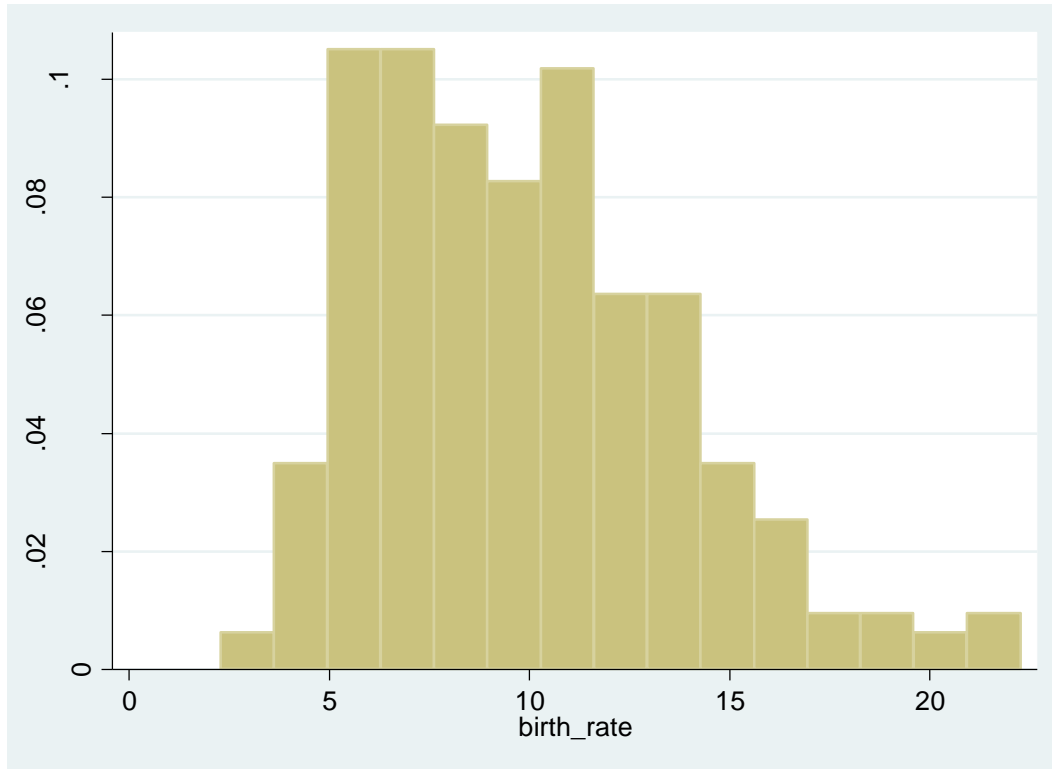


Figure 3: Distribution of County Birth Rates

To summarize, the number of children within a family depends on the birth rate in the county, and we think that the county birth rate has been determined exogenously relative to household saving rates. Thus, our empirical strategy is to first estimate the dependence of family size on county of residence in the first stage regression, Equation (6). The results indicate that the number of dependent children correlates strongly with the county birth rate. Then, second, we estimate Equation (5) with a Tobit regression.

We use a Tobit regression because our dependent variable of interest is a limited dependent variable. The calculated saving rate has a natural upper bound of unity. Furthermore, any observed saving rates below zero have been set to zero. Thus, the saving variable is truncated with a lower and upper bound. In this setting, the Tobit model (a type of censored regression) is preferred to ordinary least squares (OLS). Our main findings are qualitatively robust to using OLS (results are available upon request); however, OLS coefficient estimates are known to be biased, and it is more appropriate to assume that the error term is drawn from a truncated normal distribution. Therefore, we report only the Tobit regressions in the next section, which contains our empirical estimates.

5. Empirical Findings

We begin with the main empirical results, which come from regressing household saving rates on the number of dependent children in a household and the control variables. Then, we present additional analysis on how saving behavior depends on other household characteristics such as residing in urban versus rural areas, being employed in the public sector versus nonpublic sector, and education expenditures, and we also looked at the matched 2011 and 2013 panel. In all cases, the main finding remains the same: families with more children save less.

Household Saving and the Number of Dependent Children

Table 2 contains the main regression results. Column 1 reports the Tobit regression estimates based on Equation (5) and the 2013 CHFS data. The estimate for the coefficient (β) on *Children* equals -0.025, indicating that household saving rates decrease by 2.5 percentage points with each additional child. This estimate is practically large and statistically different from zero at better than the 1 percent level.

Column 2 reports the regression results from the two-stage instrumental variable regression. The coefficient estimate ($\gamma = 0.056$) for the *County* instrumental variable in the first-stage regression based on Equation (6) is highly statistically significant at the 1 percent level.¹² We suppress the rest of the first-stage estimates to conserve space, but the full results are available upon request. The estimate of β in the second stage equals -0.056, and it is significant at the 1 percent level. Taking the estimate literally implies that each additional child decreases a household's saving rate by 5.6 percentage points, on average. We interpret this result as a very large dependent child effect.

To get a sense of the magnitude, consider the cross-sectional estimate in regards to the observed decline in family size over time in China. Prior to the enactment of the one-child policy, families contained around three dependent children on average. Now families have less than one. Thus, our IV estimate of β (-0.056), in a rough back of the envelope calculation, implies that the decline in dependent children (from 3 to 1) increased the average household's saving rate by about 11 percentage points. In the structural model presented in Section 2, reducing the number of dependent children from 3 to 1 increases the saving rate by about 10 percentage points (using standard parameter values). Thus, our empirical findings support the structural model's qualitative and quantitative predications. Chinese families with fewer children have higher saving rates in the data, just as the life-cycle theory of household saving predicts.¹³

Table 2: The Effect of Dependent Children on the Household Saving Rate

	(1) Tobit	(2) IV Tobit Full sample	(3) IV Tobit non-public sector
<i>Children</i>	-0.0249*** (0.00286)	-0.0557*** (0.0213)	-0.0687*** (0.0224)
<i>Elders</i>	-0.0406** (0.0174)	-0.0280 (0.0195)	-0.0207 (0.0218)
<i>Workers</i>	0.0241*** (0.00240)	0.0276*** (0.00336)	0.0274*** (0.00363)
<i>Age</i>	-0.537*** (0.120)	-0.585*** (0.126)	-0.669*** (0.136)
<i>Age Square</i>	0.692*** (0.114)	0.688*** (0.114)	0.734*** (0.121)
<i>Education</i>	0.00576*** (0.000709)	0.00493*** (0.000877)	0.00409*** (0.000967)
<i>Health</i>	-0.00962* (0.00512)	-0.00996* (0.00518)	0.0335*** (0.0107)
<i>Married</i>	0.0174** (0.00726)	0.0275*** (0.0101)	-0.0122** (0.00563)
<i>Risk Averse</i>	8.89e-05 (0.00617)	0.000179 (0.00622)	0.00315 (0.00681)
<i>Risk Prefer</i>	-0.00888 (0.00886)	-0.0106 (0.00897)	-0.00678 (0.00991)
<i>Log Asset</i>	0.00157*** (0.000549)	0.00173*** (0.000574)	0.00209*** (0.000614)
<i>Log Debt</i>	3.85e-05 (0.000618)	0.000262 (0.000628)	0.000298 (0.000710)
<i>Public</i>	0.0316*** (0.00709)	0.0298*** (0.00718)	
<i>Rural</i>	-0.0152** (0.00613)	-0.0138** (0.00637)	-0.0152** (0.00672)
<i>Constant</i>	1.244*** (0.0329)	1.289*** (0.0431)	1.329*** (0.0471)
<i>Observations</i>	21,861	21,861	17,649

Note: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2 Column (3) uses employment to check on the endogenous nature of the instrument. The one-child policy was enforced differentially for nonpublic sector workers, for which the one child policy was applied more loosely. For instance, rich families working in the nonpublic sector could pay a fine to have more children. Although not statistically different from the estimate in column 2, the regression indicates an even larger negative effect (-0.0687) for families employed in the nonpublic sector. We take this as evidence that our instrument is robust to this type of endogeneity.

Table 2 also contains the coefficient estimates for each of the control variables. The estimates are not surprising. The number of older dependents (*Elders*) significantly reduces the saving rates. This finding is consistent with the Chinese tradition that adult children transfer money to and otherwise materially supported their elderly parents, which likely decreases household savings. Household saving rates increase with house assets (*Asset*) and the number of workers (*Workers*). Poor health (*Health*) significantly decreases saving. Household heads with more years of education (*Education*) have higher saving rates, as has been found in other contexts (see Kane and Rouse, 1995).

Saving rates are U-shaped with respect to age. To see this more clearly, Figure 4 shows the saving rate by age in the raw data. Saving is high for younger workers, relatively low for those aged 30-50, and high again for households about to enter retirement. This U-shaped pattern (rather than the hump shape often observed in other countries) has been well documented, but not fully explained. Theories for its emergence include relatively high recent wage growth for younger workers (Song and Yang, 2010) and delayed fertility (Curtis, Lugauer, and Mark, 2015). Neither the risk preference nor the house debt variables affect savings significantly.

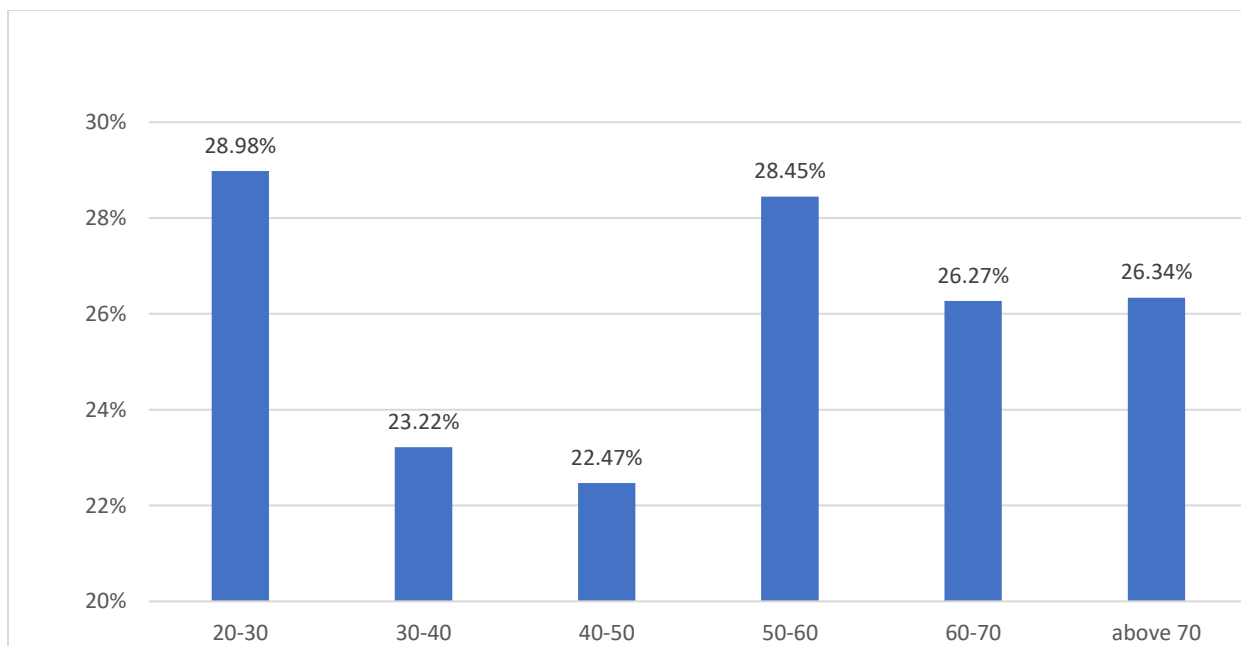


Figure 4: Household Saving Rates by Age Group

Additional Analysis

This section presents additional experiments examining various sub-groups of the population. First, we analyze whether the first-born child has a differential effect on saving relative to the second or third born. After that, we examine household education expenditures, with a particular emphasis on households with college age children. We then present how saving behavior differs in urban versus rural areas, and revisit public versus nonpublic sector employment. Lastly, we include the 2011 CHFS data to study a panel of households.¹⁴ The regressions include the full set of control variables, but we do not report the coefficient estimates in the tables to conserve space.

There may be economies of scale in raising more than one child, and thus additional children may have a smaller effect on the saving rate. Table 3 column (1) examines whether the saving rate has a non-linear relationship with regard to the number of the dependent children.¹⁵ As the table shows, a family with one child decreases its saving rate by 4.5 percent on average, while a family with two children decreases its saving rate by 6 percent. The marginal reduction of the saving rate of the second child is 1.5 percent. The coefficient for three or more children is -0.067, which indicates an even smaller change in the saving rate for each additional child beyond two. Therefore, compared to the average family with no dependent children, the first child has the largest marginal effect on saving.

Table 3: The Effect of Dependent Children on Saving Rate: Nonlinear and by Age Group

VARIABLES	(1) Nonlinear model	(2) College dummy	(3) Young children	(4) College children
<i>Children</i>		-0.0212*** (0.00293)		
<i>College</i>		-0.0586*** (0.00856)		
<i>One Child</i>	-0.0454*** (0.00600)			
<i>Two Children</i>	-0.0598*** (0.00787)			
<i>Three+ Children</i>	-0.0674*** (0.0132)			
<i>Children(Age<=16)</i>			-0.0146*** (0.00323)	
<i>Children(16<Age<=25)</i>				-0.0385*** (0.00489)
Observations	21,861	21,861	21,861	21,861

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Education Expenditures by Age of Child, Sector of Employment, and Area of Residence

Variable	Obs	Average	Share of Income
Child Age<=16	8975	3931 (7547)	-
16<Child Age<=25	3741	7842.14 (23256)	-
Public	4054	4696 (16546)	0.0713 (0.13)
Nonpublic	17807	2579 (11605)	0.0592 (0.13)
Urban	15033	3359 (14789)	0.0565 (0.12)
Rural	6828	2118 (5774)	0.0724 (0.15)

Standard Deviation in parenthesis

Table 4 reports education expenditures broken up by the age of the oldest dependent child. The summary statistics show that families with dependent children aged 16-25 spend twice as much on education (7842 RMB) compared to families with children under the age of 16 (3931 RBM). Families with children receiving overseas education (not separately reported) spend about 20 times more on education than families with children under the age of 16. The structural life-cycle model from Section 2 implies that the saving rate decreases with the degree to which parents care for their children (over the relevant range of parameter values). Sending children to college (or education expenditures more generally) could act as a proxy for the degree of parental care (or ability to care) for dependent children. That is, education expenditures might contain information on household specific preferences (we return to the idea of household fixed effects below, with the panel data).

Table 3 columns (2) - (4) report regression results aimed at further testing whether households with college age children save less. The second column uses a dummy variable equal to one if the household contains a dependent child aged 16 to 25 and the parents pay the child's education fees. As the results show, the households with children aged 16 to 25 have significantly lower saving rates. The last two columns report the regression results with indicator dummy variables for each age group. We break the children groups at 16 as this is the average age that children enter college or similar education institutions. The regression results show that families with children aged 16 to 25 reduce their saving rate (-0.039) more than households with children aged less than 16 (-0.015).

Households working in the public sector have higher saving rates, while households residing in rural areas have significantly lower saving rates. Those employed in the nonpublic sector or living in rural areas have lower income, possibly due to the history of the Hukou and social security systems. These groups also may have different education (or other) expenditures related to children. In the data, education is the fourth largest component of family expenditures (after food, housing, and transportation). Table 4 reports education expenditures by sector of employment (public versus nonpublic) and location of residence (urban versus rural), both as a total and as a percent of household income. Education spending as a share of income is higher for the public (7.1 percent) group than for the nonpublic (5.9 percent) group. The rural group (7.2 percent) actually spends more on education as a share of income than the urban group (5.7 percent). Although, total education expenditures are much higher for households employed in the public sector or living in urban areas, on average.

Table 5 reports the coefficient estimate for β in Equation 5, with the data broken up into the urban (column 1) and rural (column 2) sub-samples and, separately, into the public (column 3) and nonpublic (column 4) sub-samples. The regressions indicate that an additional child lowers the saving rate more for urban households, possibly reflecting the higher (absolute) cost of raising children in urban areas. Similarly, an additional child reduces the saving rate more for families working in the public sector. These estimates hint at a tradeoff between the quantity and quality of children. The single-child families employed in the public sector invest more in the one child compared to nonpublic sector families, which typically have more children.

Table 5: The Effect of Dependent Children on Saving Rate by Area and Sector

VARIABLES	(1) Urban Tobit	(2) Rural Tobit	(3) Public Tobit	(4) Nonpublic Tobit
<i>Children</i>	-0.0285*** (0.00409)	-0.0235*** (0.00412)	-0.0313*** (0.00863)	-0.0236*** (0.00302)
Observations	15,033	6,828	4,054	17,807

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Finally, Table 6 reports the results using the matched 2011 and 2013 CHFS data. The panel allows us to address the potential bias from household specific missing variables. Table 6 reports model specifications using random and fixed effects. Note that some families change status for many of the variables. For instance, the number of dependents can change, or rural status changes if the family migrates to the city. Therefore, we can estimate the coefficients on these variables. While smaller in magnitude, the coefficient (β) on the number of dependent children is still large and significant at the 1 percent level.

Table 6: The Effect of Dependent Children on Household Saving in the Panel Data (2011-2013)

VARIABLES	(1) Random Effects	(2) Fixed Effects
<i>Children</i>	-0.0192*** (0.00170)	-0.0141*** (0.00402)
<i>Observations</i>	13,120	13,120
<i>R-squared</i>		0.029
<i>Number of sid</i>	8,752	8,752

Robust standard errors in parentheses,*** p<0.01, ** p<0.05, * p<0.1.

To summarize, the empirical findings support the main theoretical prediction that the number of dependent children is negatively correlated with household saving rates. Saving rates are significantly higher for urban households or those employed in the public sector. Households residing in urban areas have higher saving rates in part because they have a lower ratio of education expenditures to income. However, having one more child reduces saving rates more for households in urban areas than in rural areas. Using the education expenditure data, we also identified a group of households spending more on education expenditures – those with children in college. These households have significantly lower saving rates. Conversely, households with fewer children in college save more.

6. Conclusion

This paper studies the impact of family size on household saving in China. We first present a 2-period structural model that includes finite lifetimes, saving for retirement, and in which parents care about the consumption of their dependent children. For plausible parameter values, the model implies a negative relationship between the number of dependent children in the family and the household's saving rate. Then, we test the model's implications using a new data set on household finances. The strong response of fertility rates to family-planning policies (e.g. the one-child policy) in China allows us to address the endogeneity between saving and birth decisions. The enforcement of the family-planning policies has varied across geographic regions. Thus, we instrument for the number of dependent children in the household with the county level number of births in a two-stage regression analysis.

We find that Chinese households with fewer dependent children have significantly higher saving rates. This finding supports the implications from the life-cycle model and provides additional evidence supporting the idea that the decline in fertility rates has contributed to the increase in aggregate household saving over time. We also find that saving rates vary with age and tend to be higher for the households with more workers, higher education, better health, and more assets.

As mentioned, a few related theories could help explain our main empirical results. In particular, Choukhmane, Coeurdacier, and Jin (2017) emphasize children as an investment for old-age support and Imrohorglu and Zhao (2017) focus on children as insurance against health shocks. We have emphasized the theory of Curtis, Lugauer, and Mark (2015), which models children as a consumption good. However, each theory could imply the observed negative correlation between the number of children and household saving rates. Disentangling the relative importance of these three potential mechanisms lies beyond the scope of our analysis, and, so, we leave it to future research. Most likely, each has had a significant effect on saving by Chinese households.

These papers connecting family size to saving decisions contribute to the rapidly evolving literature on the interaction between family economics and the macro-economy (see Greenwood, Guner, and Vandenbroucke (in press) for an overview) and the related health and family planning policies. While our analysis does not focus on policy changes, the findings may be relevant for the recent relaxation of the one-child rule. Broadly speaking, our results indicate that additional children could decrease aggregate household saving rates in the near term. Families with more dependent children save less.

Notes

- + We thank seminar participants at the University of Notre Dame, University of Nebraska at Omaha, Zhejiang University, Jinan University, the 2015 Chinese Economists Society Conference (Ann Arbor), the 2015 World Congress of Comparative Economics (Rome), the 2016 Chinese Economists Society Conference (Sacramento), and the 2016 ASSA meetings (San Francisco) for their helpful comments.
1. The household saving rate in China generally stayed below 5 percent prior to 1980. Today, Chinese households save nearly 30 percent of their income. Several recent papers have focused on the aggregate Chinese household saving rate, including Bannerjee et al. (2014), Chamon and Prasad (2010), Chao et al. (2011), Choukhmane, Coeurdacier, and Jin (2017), He, Lei, and Zhu (2017), Horioka and Wan (2007), Lugauer and Mark (2013), Rosenzweig and Zhang (2014), and Song et al. (2015).
 2. These papers build off of Braun, Ikeda, and Joines (2009) and Chen, Imrohorglu, and Imrohorglu (2007), which focus on Japan, and Curtis and Mark (2011), which examines the application of standard macroeconomic models to China. Auerbach and Kotlikoff (1987) lay out the foundation for these types of models.
 3. In our simple framework, household outcomes are certain. Chamon, Liu, and Prasad (2013) and Choi, Lugauer, and Mark (2017) study how idiosyncratic income shocks affect the saving behavior of Chinese households.
 4. Curtis, Lugauer, and Mark (2015) consider a richer environment featuring a formal social security system, informal intergenerational transfers, taxes, and time-varying wages and interest rates. Since the results derived in this section also hold in the richer environment, we do not include these features for the sake of simplicity.
 5. Clearly, we are abstracting from many features of China, such as the transition to a market orientated economy (see Song, Storesletten, and Zilibotti (2011), Berkowitz, Ma, and Nishioka (2015), Chang et al. (2015), and Curtis (2016) for more on this topic). Instead of building an all-inclusive model, our intention is to highlight the dependent children mechanism.
 6. Our simple set-up implicitly assumes away the dependence of the saving rate on income. More general models (with more periods or a different utility function) may have the saving rate depend on income. For now, we omit the possible income effect, but we return to this point below.

7. In 2011, the CHFS randomly selected 80 counties among the total 2,585 counties in China. In each county, 4 communities were selected, 320 communities in total.
8. Most of the literature on Chinese saving has concentrated on household rather than public or corporate saving. Ma and Yi (2010) and Yang (2012) are notable exceptions.
9. We omit income level from the main regressions, as it does not enter into the structural model's equation for the saving rate. Assets and education proxy, somewhat, for a possible income effect. Although, our results remain unchanged whether or not these control variables are included.
10. A few related papers have documented and accounted for differences across provinces, including Zhang, Zhang, and Zhang (2015), Qian (2009), and Wei and Zhang (2011).
11. Many papers, such as Rosenweig and Wolpin (1980) and Li, Zhang, and Zhu (2008), have used the occurrence of twins as an exogenous shock to family size. Angrist and Evans (1998) use parental preference for mixed-sex siblings, and Wu and Li (2012) consider the changing enforcement of the one-child policy over time. None of these studies focuses on saving behavior, however.
12. We drop 176 observations that do not have county information.
13. As discussed above, our empirical findings are not necessarily inconsistent with some of the other theories for why the number of children affect household saving decisions.
14. The results are presented without the instrumental variable approach. The two-stage estimates of the coefficient of interest are larger, as in Table 2.
15. A few recent papers have explored the link between the number of siblings (for example, Zhou 2014) and / or the gender of dependent children (for example, Wei and Zhang 2011) and household saving behavior.

References

- Angrist, Joshua and William Evans, 1998. "Children and their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 88(3), 450-477.
- Auerbach, Alan J. and Laurence J. Kotlikoff, 1987. "Evaluating Fiscal Policy with a Dynamic Simulation Model," *American Economic Review*, 77(2), 49-55.
- Banerjee, Abhijit, Xin Meng, Tommaso Porzio and Nancy Qian, 2014. "Aggregate Fertility and Household Savings: A General Equilibrium Analysis using Micro Data," NBER Working Paper 20050.
- Barro, Robert J. and Gary S. Becker, 1989. "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57(2), pp. 481-501.
- Berkowitz, Daniel, Hong Ma and Shuichiro Nishioka, 2015. "Recasting the Iron Rice Bowl: The Reform of China's State Owned Enterprises," Working Paper.
- Braun, R. A., D. Ikeda and D. H. Joines, 2009. "The Saving Rate in Japan: Why it has Fallen and Why it will Remain Low," *International Economic Review*, 50, 291-321.
- Chamon, Marcos and Eswar S. Prasad, 2010. "Why are Saving Rates of Urban Households in China Rising?," *American Economic Journal: Macroeconomics*, 2, 91-130.
- Chamon, Marcos, Kai Liu and Eswar S. Prasad, 2013. "Income Uncertainty and Household Savings in China," *Journal of Development Economics*, 105, 164-177.
- Chang, Chun, Kaiji Chen, Daniel Waggoner, and Tao Zha, 2015. "Trends and Cycles in China's Macroeconomy," NBER Working Paper 21244.
- Chao, Chi-Chur, Jean-Pierre Laffargue and Eden Yu, 2011. "The Chinese saving puzzle and the life-cycle hypothesis: A revaluation," *China Economic Review*, 22, 108-120.
- Chen, K., A. Imrohoroglu and S. Imrohoroglu, 2007. "The Japanese Saving Rate Between 1960 and 2000: Productivity, Policy Changes, and Demographics," *Economic Theory*, 32, 87-104.
- Choi, Horag, Steven Lugauer and Nelson Mark, 2017. "Precautionary Saving of Chinese and US Households," *Journal of Money, Credit and Banking*, 49, 635-661.

Choukhmane, Taha and Nicholas Coeurdacier, and Keyu Jin, 2017. "The One-Child Policy and Household Savings," Working Paper.

Coeurdacier, Nicholas, Stephane Guibaud, and Keyu Jin, 2014. "Fertility Policies and Social Security Reforms in China," *IMF Economic Review*, 62(3), 371-408.

Curtis, Chadwick C., 2016. "Economic Reforms and the Evolution of China's Total Factor Productivity," *Review of Economic Dynamics*, 21, 225-245.

Curtis, Chadwick C., Steven Lugauer and Nelson C. Mark, 2015. "Demographic Patterns and Household Savings in China," *American Economic Journal: Macroeconomics*, 7(2), 58-94.

Curtis, Chadwick C., Steven Lugauer and Nelson C. Mark, 2017. "Demographics and Aggregate Household Savings in Japan, China and India," *Journal of Macroeconomics*, 51, 175-191.

Curtis, Chadwick C. and Nelson C. Mark, 2011. "Business Cycles, Consumption and Risk-Sharing: How Different is China," *Frontiers of Economics and Globalization Vol. 9: The Evolving Role of Asia in Global Finance*. Yin-Wong Cheung, Vikas Kakkar, and Guonan Ma, editors.

Gan, Li, Zhichao Yin, Nan Jia, Shu Xu, Shuang Ma, Lu Zheng, 2014. "Data you need to know about China: Research Report of China Household Finance Survey", Springer Berlin Heidelberg.

Greenwood, Jeremy, Nezih Guner, and Guillaume Vandenbergue, *In Press*. "Family Economics Writ Large," *Journal of Economic Literature*.

He, Hui, Ning Lei, and Dongming Zhu, 2017. "The Impact of Rapid Aging and Pension Reform on Savings and the Labor Supply: the Case of China," Working Paper.

Horioka, C. and J. Wan, 2007. "The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data," *Journal of Money, Credit and Banking*, 39, 2077-2096.

Imrohorglu, Ayse and Kai Zhao, *In Press*. "Intergenerational Transfers and China's Social Security Reform," *The Journal of the Economics of Ageing*.

Imrohorglu, Ayse and Kai Zhao, 2017. "The Chinese Saving Rate: Long-Term Care Risks, Family Insurance, and Demographics," Working Paper, USC Marshall.

Kane, Thomas J., and Cecilia Elena Rouse, 1995. "Labor-Market Returns to Two- and Four-Year College," *The American Economic Review*, 85(3), 600-614.

Li, H, J Zhang, and Y Shu, 2008. "The Quantity-Quality Tradeoff of Children in a Developing Country: Identification using Chinese Twins," *Demography*, 45(1), 223-243.

Lugauer, Steven and Nelson C. Mark, 2013. "The Role of Household Saving in the Economic Rise of China," in the CES-ifo seminar series (MIT Press), *Global Interdependence, Decoupling and Recoupling*, Cheung and Westermann editors.

Ma, Guonan and Wang Yi, 2010. "China's High Saving Rate: Myth and Reality," *International Economics*, 122, 5-40.

Modigliani, Franco and Shi Larry Cao, 2004. "The Chinese Saving Puzzle and the Life-Cycle Hypothesis," *Journal of Economic Letters*, 42, 145-170.

Qian, Nancy, 2009. "Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrollment in Rural China," NBER Working Paper, no. 14973.

Rosenzweig, Mark R. and K Wolpin, 1980. "Testing the Quantity-Quality Fertility Model: the use of Twins as a Natural Experiment," *Econometrica*, 48(1), 227-240.

Rosenzweig, Mark R. and Junsen Zhang, 2014. "Co-residence, Life-Cycle Savings, and Intergenerational Support in Urban China," NBER Working Paper, no. 20057.

Song, Zheng Michael, Kjetil Storesletten, Yikai Wang and Fabrizio Zilibotti, 2015. "Sharing High Growth Across Generations: Pensions and Demographic Transition in China," *American Economic Journal: Macroeconomics*, 7(2), 1-39.

Song, Zheng Michael, Kjetil Storesletten and Fabrizio Zilibotti, 2011. "Growing Like China," *American Economic Review*, 101(1), 196-233.

Song, Zheng Michael and Dennis Tao Yang, 2010. "Life Cycle Earnings and Saving in a Fast-Growing Economy," Working Paper.

Wei, Shang-Jin and Xiaobo Zhang, 2011. "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," *Journal of Political Economy*, 119, 511-564.

Wooldridge, Jeffrey, *Introductory Econometrics: A Modern Approach*, 5/e, Southwestern, 2013.

Wu, Xiaoyu and Lixing Li, 2012. "Family Size and Maternal Health: Evidence from the One-Child Policy in China," *Journal of Population Economics*, 25, 1341-1364.

Yang, Dennis Tao, 2012. "Aggregate Savings and External Imbalances in China," *Journal of Economic Perspectives*, 26(4), 125-146.

Zhang, Haifeng, Hongliang Zhang and Junsen Zhang, 2015. "Demographic Age Structure and Economic Development: Evidence from Chinese Provinces," *Journal of Comparative Economics*, 43, 170-185.

Zhou, Weina, 2014, "Brothers, Household Financial Markets and Savings Rate in China", *Journal of Development Economics*, 111, 33-47.

Appendix: Derivation of the Saving Rate (Equation 4)

The household utility maximization problem (Equation 3) can be rewritten by using the linear budget constraints (Equations 1 and 2) to replace C_1 and C_2 . The new problem is to choose S and C^c to maximize Equation A.1.

$$U = (1-\sigma)^{-1} [\mu n^\eta (C^c)^{1-\sigma} + (I-S-nC^c)^{1-\sigma}] + \delta (1-\sigma)^{-1} S^{1-\sigma} \quad (\text{A.1})$$

The optimal choices solve the following two first order conditions.

$$0 = \mu n^\eta (C^c)^{-\sigma} - n(I-S-nC^c)^{-\sigma} \quad (\text{A.2})$$

$$0 = -(I-S-nC^c)^{-\sigma} + \delta S^\sigma \quad (\text{A.3})$$

Solving Equation A.3 for C^c and substituting into A.2 gives the following.

$$0 = \mu n^{\sigma+\eta} (I-S-\delta^{-1/\sigma}S)^{-\sigma} - n \delta S^\sigma \quad (\text{A.4})$$

Solving Equation A.4 for S/I using simple algebra leads to Equation 4 in the main text.

Vitae:

Steven Lugauer is in the Department of Economics at the University of Kentucky. He received a PhD and MS in Economics at Carnegie Mellon University, an MBA in Economics and Finance at the University of Miami, and a BA in Economics at the University of Chicago. His primary research interests are in demographics, saving, and China.

Zhichao Yin is a professor in the School of Finance at the Capital University of Business and Economics. Dr. Yin received a PhD in Economics at Southwestern University of Finance and Economics. He is the Deputy Director of Survey and Research Center for China Household Finance. His research interests are Household Finance and Applied Microeconomics.