

Wen, Xin; Cheng, Zhiming; Tani, Massimiliano

Working Paper

Daughters, Savings and Household Finances

GLO Discussion Paper, No. 1474

Provided in Cooperation with:

Global Labor Organization (GLO)

Suggested Citation: Wen, Xin; Cheng, Zhiming; Tani, Massimiliano (2024) : Daughters, Savings and Household Finances, GLO Discussion Paper, No. 1474, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/300891>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Daughters, Savings and Household Finances

Xin Wen*

School of Business, The University of New South Wales, Canberra, Australia

Email: xin.wen6@unsw.edu.au

Zhiming Cheng

Department of Management, Macquarie Business School, Macquarie University, Sydney,
Australia

Social Policy Research Centre, The University of New South Wales, Sydney, Australia

Global Labor Organization, Essen, Germany

Email: zhiming.cheng@mq.edu.au

Massimiliano Tani

School of Business, The University of New South Wales, Canberra, Australia

Global Labor Organization, Essen, Germany

IZA, Bonn, Germany

Email: m.tani@unsw.edu.au

Statements and Declaration

Acknowledgement

Xin Wen received the financial support received through the UNSW University International Postgraduate Award (award number: RSRE7063) and Development and Research Training Grant (grant number: RSTR9001). Zhiming Cheng received the financial support from the Macquarie University Research Acceleration Scheme (grant number: 173988730).

Conflicts of interest

Non-financial interests: none.

* Corresponding author

Daughters, Savings and Household Finances

Abstract

We explore the link between child gender and household financial decisions within a cultural environment that strongly favours having a son. Using data from the China Household Finance Survey (CHFS), we find that the presence of a daughter is associated with a lower saving rate, consistent with the hypothesis that the relative under-supply of unmarried women generates a less competitive marriage market for families with daughters vs. those with sons. As a result, such families have lower incentives to endow their daughters with bigger asset pools to enhance their marital prospects. The correlation becomes more pronounced as the daughter approaches marriageable age, and it is more common among families where the head has low financial literacy and limited education and lives in rural areas.

Keywords: daughter, household investment decisions, family savings, marriage market

JEL: D14, G11, G51, J12

1 Introduction

The preference for a son, which is common in Asian countries (Choi and Hwang, 2015; Edlund, 1999; Song and Gao, 2023), parts of Africa (Gangadharan, 2003; Milazzo, 2014; Rossi and Rouanet, 2015), and Europe (Hank and Kohler, 2000; Mills and Begall, 2010), stems from an established custom of viewing men as primary earners and carriers of the family name (Pitt et al., 2012; Asadullah et al., 2021; Lundberg, 2005). Daughters who marry typically adopt their partner's surname. This contributes to a negative bias towards them (Anukriti et al., 2022; Choi and Hwang, 2015; Kaul, 2018; Azam and Kingdon, 2013), which can take the form of lower-quality childrearing and education than comparable males in the infant and educational stages (Barcellos et al., 2014; Asadullah et al., 2021).

As children reach adulthood, parental support takes on additional forms beyond investments in education: parents may provide financial transfers that can give their offspring a substantive advantage in the quest to find a high-quality partner and help overcome the initial living costs of the married couple. In such circumstances, however, if a preference for sons has caused an environment in which the birth ratio is unbalanced in favour of males in the marriage market, the competition for a suitable partner gives parents asymmetric financial incentives according to the child's gender. Large cohorts of sons will inevitably push families to enhance their boy's marital attractiveness by endowing them with larger gifts or asset transfers (Wei and Zhang, 2011; Gao et al., 2023).¹ Parents of male offspring may have prepared for such events over the years by accumulating savings exceeding those of comparable families with daughters, who instead enjoy a relative 'scarcity' relative to the higher competitive fever for a partner that prevails among unmarried males. Is this the case?

Existing research is relatively silent about the topic, notwithstanding a large amount of literature analysing household financial choices and the broad marriage market. There are only few studies on the influence of a child's gender on his/her household's savings and investment choice (e.g. Wei and Zhang (2011), Bogan (2013), and Li et al. (2022)). Often, these have access only to cross-sectional data and hence are vulnerable to omitted variable bias because of limited controls for unobserved individual heterogeneity.²

¹ Son preference contributes to gender imbalances in populations, leading to a relative scarcity of women in certain regions, often referred to as the 'marriage-squeeze' effect.

² For example, Bogan (2013) employs the U.S. data from the National Longitudinal Survey of Youth 1979 (NLSY79), with a focus on the 2004 NLSY79 wave using a probit model specification. Her specification

We aim to answer this question by addressing whether families with a daughter save less than comparable families with a son. To do so, we use China as a case study and the China Household Finance Survey (CHFS) – a longitudinal survey – as the data source. As panel data enable us to control for time-invariant individual heterogeneity, we can measure the effect's occurrence as the child's age changes and hence integrate potential marriage age—a critical factor in the marriage market—into the analysis of how the child's gender can guide household economic decisions. This is novel in the literature. It is also relevant, as the link between a child's gender and a household's financial choices likely influences the flow of savings and the demand for investments – for example, by constraining consumption and skewing investment demand towards 'consumable' assets for the newlywed, such as real estate, whose risk profile may not be optimal for national growth targets or price inflation.

We find that households with a daughter, whether from one-child or multi-child families, tend to save less compared to families with a son. The child gender effect varies across different child age brackets. It is strongest when the child reaches marriageable age, regardless of the number of children in the family. The effect arises only on the amount of savings accumulated but does not extend to stock market participation, implying that families with a daughter save less but do not take more or less risk in their investments than similar families with a son. We find that the intention to buy a home and consumption of daily necessities are the channels through which this effect affects household saving decisions. This effect is muted in families characterised by high financial literacy, advanced educational attainment, urban residency, and a female head.

The results support the hypothesis that competitive pressures in the marriage market have widespread implications for the financial market and the broader economy and that asymmetries in the gender ratio of unmarried men and women can be a source of gender bias in the inter-generational transfer of wealth. The results also identify broad areas for possible targeted interventions, including legislative, cultural and normative revisions on matters related

incorporates a dummy variable to indicate whether respondents held stocks and bonds in 1998. Li et al. (2022) utilise 2013 CHFS, while Wei and Zhang (2011) employ 2002 Chinese Household Income Project. Our study enhances methodological approaches by utilising panel data from the 2013 and 2015 CHFS and mainly employing the RE estimator, supplemented by panel data FE estimator and G2SLS-RE-IV estimator. Additionally, we extend our analysis to include the 2017 and 2019 CHFS waves through OLS regression. This analytical strategy allows for an in-depth exploration of both observed and unobserved time-invariant heterogeneities.

to the adoption and transmission of surnames, the development of suitable models of shared responsibilities underpinning the economic and financial well-being of the new household, and educational programs to enhance parents' and offspring's financial literacy.

2 Background and literature review

After the Great Leap Forward of 1958-59, fertility rates soared as they were encouraged by the government (White, 2006), reaching an average of six births per woman in the 1960s (Banister, 1987).³ This trend was halted in the 1970s with the “*wan* (later), *xi* (longer), *shao* (fewer)” campaign, which advocated for delayed marriage and childbearing, wider birth spacing, and fewer offspring. Persistently high population growth rates led to the introduction in 1979 of a family planning program enforcing a one-child policy. This was implemented with varying stringency across regions.⁴ The policy was adjusted in 1984 to allow rural families with firstborn girls or ethnic minorities to have a second child. Policy violations faced repercussions, including job loss and fines. The policy had severe unintended consequences as the option of having only one child combined with strong cultural preferences for a male heir led to a surge in the use of technologies to anticipate the sex of the child (Ebenstein, 2010; Keyzers, 1991) as well as abortions, often under questionable medical practice, and other forms of violence such as forced sterilisation (Edlund, 1999; Li et al., 2022; Modigliani and Cao, 2004; Wei and Zhang, 2011; Shrestha and Jung, 2023). These aberrations led to smaller cohorts of newborns and a significantly skewed sex ratio in the new cohorts. New legislation prohibiting birth sex selection was eventually introduced in 1987.

The result of these measures was a progressive decrease in the size of new cohorts: while this was a welcome short-term outcome on China's resources, it seeded new problems for the medium and long term, as the shrinking population meant that fewer people in working age could produce and generate income and wealth to sustain the cost and needs of rising shares of older generations. In the 2010s, the government relaxed its population control measures to prop up declining fertility rates, workforce shrinkage, and an aging population (Ge et al., 2018). The birth limit increased to two children per couple in 2015 and three in 2021, accompanied by extended maternity leave and financial incentives.

³ The Great Leap Forward was a plan aimed at China's fast industrialisation by switching the workforce away from agriculture (Li and Yang, 2005; Meng et al., 2015).

⁴ For example, the penalties for above-quota births differed in rural areas, urban areas, and provinces.

Despite these policy changes, the preference for sons remains strong, and so do the many complex inter-generational consequences that such preference has on households' welfare and the distribution of wealth (Anderson, 2007).

In theory, cultural practices surrounding the marriage, such as dowry and traditional obligations or gifts, influence households' pre-marital savings and investment decisions differently according to their child's gender (Botticini and Siow, 2003; Grossbard-Shechtman, 2003; Platteau and Gaspart, 2007; Lundberg, 2005; Wei and Zhang, 2011).⁵ This bias, combined with a skewed sex ratio in favour of males, underpins the rise of a 'marriage-squeezing' effect (Edlund, 2000; Klinger-Vartabedian and Wispe, 1989), whereby fierce competition among unmarried males in the local marriage market is compounded by their families' varying level of material support (Edlund, 2000; Wei and Zhang, 2011; Li et al., 2022).⁶ The oversupply of eligible men leads to unfavourable marriage market conditions, with significant marriage expenses typically borne by the man's family, as per traditional norms (Li et al., 2022).⁷ As a result, families with daughters face less competition and financial pressure in finding marriage partners. This imbalance results in higher asset transfers favouring sons, perpetuating gender-based wealth disparities and contributing to intergenerational gender inequality, particularly when the gender of a child can be somewhat predetermined and the child is in the marriageable stage.

Societies in high income countries exhibit minimal gender bias in allocating resources, transferring wealth, or investing in children of different genders (Lundberg, 2005; Blau et al., 2020). This shift is partly attributed to the diminished reliance on sons for support in old age due to enhanced social welfare and public provisions in these wealthier societies. Nevertheless, a preference for sons persists to some extent in these societies⁸. Conversely, in traditional societies in middle and lower-income countries, the bias is marked by a strong preference for sons, which is ingrained in social institutions and norms. These elevate the value of sons and

⁵ Usually, bride premiums are relatively constant and transferred from the groom's parents to the bride or the bride's parents, while the dowry is at the willingness of the bride's parents (Anderson, 2007). In Chinese culture, the dowry is paid by the bride's family, whereas the bride price is provided by the groom and his family (Gao et al., 2022).

⁶ In the context of China's marriage market, there is a pronounced male marriage squeeze, attributable to a surplus of marriageable men.

⁷ Children's marriage is a vital part of the economic framework of Chinese families.

⁸ Lundberg (2005) suggests that sons may contribute to family stability and notes a tendency among U.S. fathers to spend more time with sons than with daughters. Blau et al. (2020) find that while U.S. natives exhibit no such preference, immigrants from source countries with less gender equity still tend to favour sons.

increase the perceived cost of daughters (Sen, 2003; Barcellos et al., 2014; Azam and Kingdon, 2013), influencing parental behaviours and decisions⁹, including divorce rates, birth order, fertility, childrearing, education, and resource allocation (Kabátek and Ribar, 2020; Barcellos et al., 2014; Blau et al., 2020).

In practice, it is challenging to identify with precision each factor influencing household financial decisions, as they accumulate over several layers and time. Existing research has identified variables such as children's characteristics (Love, 2010), the gender of the household head (Fonseca et al., 2012), family size, occupation (Banner and Schwarz, 2018), the level of education (Grinblatt et al., 2011), the existence of a family business (Amran and Ahmad, 2010; Ji et al., 2021), and home ownership (Vestman, 2019; Wei and Zhang, 2011; Wei et al., 2012) to name a few.¹⁰

Identifying empirically the influence of child gender on household-related financial choices is especially challenging when only cross-sectional data are available because it is not possible to control all sources of individual unobserved heterogeneity that may be at play. The estimates obtained may hence be influenced by bias that may over or under-represent the actual effect of interest and mislead the suggestion for, or type of, intervention. The availability of panel data in CHFS enable us to partly address this problem, as panel data estimation techniques can remove unobserved albeit only time-invariant influences. To highlight the possible influence of such unobserved sources of bias, we present estimates obtained from both cross-sectional, for comparison with existing work, and panel data – our preferred approach.

3 Data

3.1 The China Household Finance Survey

The CHFS is a high-quality, large, and nationally representative longitudinal survey that collects comprehensive information at the individual, household, and community levels. It is

⁹ For example, the 'social norms' hold that daughters will join their spouses' families post-marriage, so investing in them is perceived as less beneficial to their birth families (Adam, 1947; Porter, 2016). In addition, differences in the treatment of male and female offspring may arise from expectations of labour market participation and return (Choi and Hwang, 2015; Karbownik and Myck, 2017).

¹⁰ Affected financial decisions included savings (De Laiglesia and Morrisson, 2008; Curtis et al., 2015; Wei and Zhang, 2011), income and consumption (Morduch, 1995), retirement schemes (Anderson et al., 2017; Lusardi and Mitchell, 2011; Van Rooij et al., 2012), stock market participation (Almenberg and Dreber, 2015; Bogan, 2009; van Rooij et al., 2011; Zou and Deng, 2019), and investment decisions (Pahlevan Sharif et al., 2020).

publicly available in five waves: 2011, 2013, 2015, 2017, and 2019.¹¹ The sample size and coverage expanded over the waves, with the 2011 wave covering 25 provinces, 82 counties, 320 communities, and 8,438 households. Subsequent waves included more provinces, counties, communities, and households, with the wave in 2019 covering 343 counties, 1,360 communities, and 34,643 households. It has both cross-sectional and longitudinal data characteristics.

CHFS contains detailed data on financial information, such as income, assets, savings and stock market participation, as well as financial knowledge. It also includes each household member's demographic and socioeconomic characteristics, including age, gender, employment status, education, and marital status.

We limit our analysis to households with unmarried children aged 1 to 35 to mitigate any potential confounding effects of a married child's existing marriage and to capture the offspring's life cycle before and after their marriageable age (typically early to late 20s). In addition, we restrict our sample to household heads aged 25 to 65. We combine cross-sectional data for 2013, 2015, 2017, and 2019 into a pooled dataset for comparison with existing estimates. Our preferred approach, however, can only be implemented on the 2013 and 2015 waves, which are linked to generate an unbalanced panel.¹² After excluding outliers, we are left with 27,926 observations (households) in the pooled cross-sectional sample and 13,806 observations (households) in the unbalanced panel dataset.

3.2 Measures of household saving rate and financial market participation

Drawing on Modigliani and Cao (2004) and Wei and Zhang (2011), we measure household saving rate by the formula: $(\text{income} - \text{expenditure})/\text{income}$.¹³ In instances where households

¹¹ For further details and access to the CHFS datasets, please refer to the official website at <https://chfs.swufe.edu.cn/dczx.htm>.

¹² For instance, the 2011 wave does not include information on financial literacy. Financial literacy is an essential factor in understanding financial behaviours and decisions (Almenberg and Dreber, 2015; Anderson et al., 2017; Bianchi, 2018; Lusardi and Mitchell, 2011; Lusardi and Mitchell, 2017). Additionally, the financial literacy questions in 2017 were answered only by new respondents who had not been interviewed in the previous waves. Furthermore, in the 2019 wave, one financial literacy question was collected exclusively in urban areas. This limits the inclusion of 2011, 2017 and 2019 waves in the panel dataset.

¹³ A household's income includes wage income, agricultural operating income, industrial and commercial operating income, transfer income, and investment income. Expenditure includes food, clothing, housing, living goods and services, education and entertainment, transportation and communication, medical care, and other expenditures.

have negative savings, implying expenditure exceeds income, we attribute a value of zero as Lugauer et al. (2019) and Giavazzi and McMahon (2012) do.

In terms of stock market participation, following Li et al. (2022) and Wang et al. (2023), we define stock market participation through the presence of a stock account ownership. By having a stock account, families can buy and sell stocks at will. Specifically, a household is considered to participate in the stock market if it possesses a stock account, coded as one; if not, it is coded as zero. Mutual funds and equities in retirement accounts are excluded from our analysis, as the CHFS dataset does not provide the necessary information.

[Table 1 here]

The summary statistics are reported in Table 1. The households in the working sample save, on average, nearly a quarter of their income (23.8 per cent), and their average stock market participation is relatively low, at 14.2 per cent. These findings are consistent with previous studies on Chinese household finances (Wei and Zhang, 2011; Li et al., 2022).

Columns 2 and 3 of Table 1 report summary statistics for families with and without daughter. We observe significant differences in savings rates (4.1 per cent) and stock market participation (2.1 per cent) between these two groups. Specifically, families with daughter have lower saving rates and stock market participation rate than comparable families with only son.

3.3 Measures of independent variables and control variables

Table 2 presents the unconditional means of the independent variables and control variables. We use the share of daughters (SOD) to measure the children's gender composition in the family. For one-child households, this measure equals 1 if a daughter exists and 0 if not. In multi-child households, it represents the proportion of daughters among children who are alive. In these working samples, the average SOD is 0.455, and 52.8 per cent of the families surveyed have at least one daughter.

[Table 2 here]

As parental investments vary with children's stages (Wang et al., 2022), we generate five age cohorts according to children's age.¹⁴ Children in cohort 1, aged 1-12, are in the kindergarten or primary school phase, a stage unlikely to majorly impact household savings or future marriage-related decisions. The children in cohort 2, aged 13-17, might influence household savings as parents plan for their adolescents' needs and future prospects as candidates for marriage. The children in cohort 3 are aged 18-22, where parental savings are needed for marriage or studies. The children in cohort 4, aged 23-26, are the most common age at which to get married in China during the survey years. The children in cohort 5, aged 27-35, are becoming financially independent, making it less likely to require large parental resources. In our working sample, the average age of the firstborn child is 16.933 years old, and the eldest daughter is about 15.963 years old.

We use a self-reported 5-point Likert scale of risk attitudes, where a high score indicates a greater willingness to take risks, to control for the respondent's risk appetite.¹⁵ The average risk attitude is 2.186 out of 5, which suggests a prudent investment strategy.

The educational attainment of household heads is predominantly low. Specifically, 18.8 per cent of household heads have either no formal education or only primary school education. A majority, 56.6 per cent, have completed either junior or senior high school. Additionally, 23.2 per cent possess a vocational diploma, while a mere 1.4 per cent hold a university degree. 78.7 per cent of household heads are typically employed. Their level of financial literacy is low, averaging 1.034 out of 3.¹⁶ Household heads are classified into four age cohorts, and the majority (40.1 per cent) are 36 to 45 years old. A large proportion of household heads (50.8 per cent) are females. Most families (95.3 per cent) are composed of married couples. On average, families have 1.384 children. The average household income and assets are 95.241 thousand yuan and 1,181.691 thousand yuan (about 12 times annual income – a substantive amount), respectively.

¹⁴ In one-child families, this variable is the child's age, while in multi-child households, it is the firstborn child's or eldest daughter's age.

¹⁵ The risk attitudes question in the questionnaire is "Which of the choice below do you want to invest most if you have adequate money?" A. project with high-risk and high-return; B. project with slightly high-risk and slightly high-return; C. project with average risk and return; D. project with slight risk and return; E. unwilling to carry any risk". To facilitate the interpretation, we reversed the measurement scale, such that 1 represents unwillingness to carry any risk, 2 represents a project with slight risk and return, 3 represents a project with average risk and return, 4 represents a project with slightly high-risk and slightly high-return, and 5 represents a project with high-risk and high-return.

¹⁶ Exact wording of financial literacy questions in CHFS are listed in Appendix Table A1. 1.

We generate a variable with four categories to capture the residential place and *hukou* status: namely, an urban resident with non-agricultural *hukou* (50.2 per cent), an urban resident with agricultural *hukou* (24.5 per cent), a rural resident with non-agricultural *hukou* (1.3 per cent), and a rural resident with agricultural *hukou* (24 per cent).¹⁷

4 Empirical framework and identification strategy

We estimate the ‘daughter’ effect and whether its intensity varies with the progression of the daughter’s age using the following specification:

$$y_{it} = \gamma_0 + \gamma_1 SOD_{it} + \gamma_2 AC_{it} + \gamma_3 SOD_{it} \times AC_{it} + \sum_{i=4}^j \Gamma_j X_{it} + \mu_{it} + \omega_i$$

where the dependent variable, y_{it} , measures the saving rate or stock market participation of household i at time t . SOD_{it} is the share of daughters in children within a household at the time of the survey, serving as a key variable to assess child gender-specific impact on financial decisions. AC_{it} is a categorical variable that captures the age cohort of the firstborn child (daughter) in each household.¹⁸ It is interacted with the share of daughters to allow the effect of having a daughter to vary with her age.

The parameters of interest are γ_1 and γ_3 , as they help delineate gender-specific effects within different child age cohorts. In particular, γ_1 captures the daughter’s influence on the household saving rate and stock market participation, while γ_3 captures the variation in this effect across different child age groups. Together, these parameters enable us to analyse the changing impact of daughters on household finances as they progress through time.

X_{it} , a vector of control variables, which includes the characteristics of the household head and the household at large: namely, the household head’s age group, gender, education level, risk

¹⁷ The *hukou* is a household registration system that China established in the 1950s to control and manage population movements (Chan and Zhang, 1999). The type of *hukou* determines the level of access to public services, educational resources, employment opportunities, social benefits, insurance and pension rights in a locale. It is commonly viewed as a source of inequality between rural and urban areas (Song, 2014; Wang et al., 2020; Meng, 2012).

¹⁸ In one-child families, the relevant age under consideration is that of the child, whereas in multi-child households, the focus shifts to the age of the eldest child (or daughter).

attitude (Bogan, 2013; Li et al., 2022; Lugauer et al., 2019), marital status, financial literacy, employment status, number of children, residential and *hukou* status, and the logarithms of household income and assets (Bogan, 2013; Li et al., 2022; Almenberg and Dreber, 2015; Nguyen and Nguyen, 2020; van Rooij et al., 2011; Van Rooij et al., 2012). μ_{it} is an idiosyncratic error term, and ω_i captures time-invariant individual unobserved heterogeneity.

Given the influence of the local area sex ratio on household consumption, saving and investment preference, the model includes the province-level sex ratio among control variables (Li et al., 2022; Wei and Zhang, 2011; Horioka and Terada-Hagiwara, 2017). We also use province and year dummy variables to control for fixed effects at the provincial level and over time.

In one-child households, the variable SOD_{it} is effectively a dummy for the child's gender, categorised as 0 for a son and 1 for a daughter. In multi-child households, the interpretation is different as SOD_{it} becomes a continuous variable (the share of daughters among all children), and AC_{it} identifies the age group of the firstborn child (daughter). Notably, in our analysis of multi-child families with daughter, we particularly focus on the age of the eldest daughter, as determined by AC_{it} , to understand the implications related to her age group.¹⁹

We carry out analysis along two dimensions. First, we perform regressions on cross-sectional and panel data separately. In particular, the results distinguish between the ordinary least squares (OLS) estimates obtained on pooled cross-sectional data covering the years 2013, 2015, 2017 and 2019 and the random effects (RE) and fixed effects (FE) panel estimation covering the years 2013 and 2015.²⁰ Second, we first perform the empirical analysis by pooling one-child families and multi-child families together.²¹ Then, we carry out separate regressions for households with only one child versus those with two or more children.

In addition, due to the gender of the firstborn child being time-invariant, we use this factor along with a subjective preference for sons as instruments to perform a generalised two-stage least squares random-effects instrumental variables (G2SLS-RE-IV) regression to address the

¹⁹ By doing so, we focus on the daughter's effect on household savings and investment decisions.

²⁰ Given its superior interpretability, we utilise a linear probability model.

²¹ Because the child gender is time-invariant, we utilise OLS and RE regressions in the analysis for one-child families.

endogeneity of daughters' share.²² The rationale for using these instruments reflects the likely randomness and exogeneity of the firstborn child's gender, as supported by the existing literature (Li et al., 2022; Wei and Zhang, 2011; Ebenstein, 2010; Li and Wu, 2011), and the cultural and social norms favouring sons (Yang and Einstein, 2014), which influence family decisions on further childbearing, especially if the first child is a daughter (Kim and Lee, 2020; Barcellos et al., 2014; Greenhalgh, 2013).²³

It is worth noting that the FE estimator helps reduce the concerns of time-invariant omitted variables. However, the FE model performs poorly when the explanatory variable is nearly constant (Treiman, 2014). Parameter estimates are sensitive in FE estimation, particularly when there is a small variation in independent variables within units, leading to significant divergence from the true effect due to random variation, but the RE model provides more stable parameter estimates by leveraging partial information pooling across units (Clark and Linzer, 2015). In other words, the FE estimator relies on the 'sufficient' within-group variation, and it becomes unstable and may not be very reliable when the within-group change is small (Longhi and Nandi, 2015). As shown in Table A1. 2, the within standard deviations of the SOD and the age group of the firstborn child are small, suggesting that these variables do not change much within groups over time.

Furthermore, the gender of the household head is important in studying household saving and investment decisions. Suppose this variable is omitted due to its constancy over time. In that case, the analytical sample is substantively altered because the FE estimator only models changing units, making the findings applicable to a selected subgroup only (Treiman, 2014). In this case, RE estimations may be necessary to understand the population parameters (Wooldridge, 2010). Despite these limitations, we present FE estimations for reference.

5 Results

5.1 Baseline Results

Table 3 presents the estimated coefficients from the three regression models (OLS, RE and FE) for pooled households. Panel A displays the regression results, while panel B provides the

²² The 2017 and 2019 CHFS do not collect the subjective son preferences. Thus, we apply instrumental variables regressions to the panel dataset.

²³ In the mechanism analysis, the results in columns 2 and 3 of panel A in Table 14 show that son preference does not relate to household saving rate and stock account ownership.

linear combination of parameters γ_1 and γ_3 . The linear combination provides a comprehensive estimate of the daughter effect at a set age group on the family's savings rate and probability of having a stock market account for investing.

[Table 3 here]

Table 3 reveals a negative correlation between the proportion of daughters in a family and the family's savings rate as the first child approaches marriageable age (23 to 26 years in China), evidenced by columns 1, 3, and 5. This pattern is not observed in investment decisions, as indicated by columns 2, 4, and 6. Specifically, the RE model in column 3, panel A, highlights a significant negative interaction term ($SOD \times [23, 26]$), indicating a 0.051 percentage point decrease in savings rate for a percentage point increase in the share of daughters when the first child is of marriageable age, relative to those not of marriageable age. Additionally, column 3 of panel B elucidates that an increase in the proportion of daughters by a percentage point results in a 0.061 percentage point decline in the savings rate for families with a first child in the marriageable age range. This indicates that families with all sons save, on average, about 10.4% of the median income (RMB 4736.729) more than those with all daughters, when the families have a marriageable first child.

5.2 Endogeneity of Daughter Indicator

To address the endogeneity of the daughter indicator, we utilise the gender of the firstborn child and a subjective preference for sons as instruments to perform G2SLS-RE-IV regressions using the short panel based on 2013 and 2015 CHFS. The first stage results are presented in Appendix Table A1. 3. The second stage results, shown in Table 4, suggest that even when instrumented, a higher share of daughters is associated with lower household savings as the firstborn child reaches marriage age. However, such an effect does not extend to investment decisions. These results are in line with the baseline results reported in Table 3.

[Table 4 here]

5.3 One-Child Households

The results in Table 5 suggest that one-child families with a daughter save less than those with only a son (columns 1 and 3), and this effect grows non-linearly with the age of the child until about age 26. According to the RE model in panel A, the interaction term 'Daughter \times [23, 26]'

is statistically significant and negative, highlighting a decrease of 2.8 percentage points in reduction savings rates for families with a daughter aged 23 to 26, holding other factors constant. The largest observed decline in saving rates, about 4.1 percentage points, occurs when daughters reach their most eligible age for marriage, compared to families with a son in the same age bracket (panel B of column 3). This suggests that when controlling other factors, the one-child families with a son at a marriageable age save, on average, 6.9% of the median income (RMB 3681.310) more than their counterparts with a daughter.

[Table 5 here]

However, no ‘daughter’ effect emerges with reference to stock market participation (columns 2 and 4), implying that households do not vary their risk preferences to access extra earnings (at higher risk). The ‘daughter’ effect is, therefore, purely restricted to the amount of savings available but does not change the risk profile of how they are invested.

5.4 Multi-Child families

[Table 6 here]

Table 6 reveals consistent trends within multi-child households: the share of daughters inversely affects the saving rate, particularly as the first child reaches marriageable age. This indicates that holding all other factors constant, families with a marriageable firstborn child experience an additional decrease in their savings rate of 0.094 percentage points for each one percentage point increase in the share of daughters, compared to families without a marriageable first child (column 3, panel A). This means that in this age group, the family with more daughters saves less than those with more sons. The results in panel B of column 3 further delineate this effect: when the first child is in the age group 23 to 26, a one percentage point increase in the share of daughters is associated with a 0.095 percentage points decrease in saving rates, amounting to RMB 53.360 or 0.159 per cent of the median income in the multi-child families. With reference to stock market participation, the ‘daughter’ effect does not translate into any statistically discernible influence. Hence, no change in the investment risk profile emerges from such an effect.

[Table 7 here]

To provide a more comprehensive understanding of daughters' role in household savings and investment decisions, we analyse households with daughter only. The analysis for one-child households has been illustrated in Table 5. Thus, we focus on the multi-child households. Similar effects appear in Table 7, which reveals that in families where the eldest daughter is of marriageable age (23-26), an increased proportion of daughters is associated with a lower saving rate. Specifically, according to the RE model (column 3 of panel A), in families with a eldest daughter of marriageable age, each percentage point increase in the share of daughters correlates with a 0.162 percentage point reduction in the saving rate. Panel B offers a summary view among the families with a marriageable eldest daughter, as a one percentage point increase in the share of daughters results in a 0.123 percentage point reduction in the saving rate, equivalent to about 68.677 RMB or 0.207 per cent of the median income in multi-child families with daughter, as per the RE model. Given the daughter effect in one-child households, the families with more daughters will save less when the eldest daughter approaches the potential marriageable age, *ceteris paribus*, than those with sons.

Our findings reveal that the presence of a daughter in the family, whether in one-child or multi-child households, correlates with lower savings rates. This tendency is particularly marked around the daughter's and the firstborn child's prospective marital ages. These findings support the hypothesis that the competitive pressure on unmarried men motivates their families to enhance their marriage prospects by increasing their potential bride price or assets, aligning with Wei and Zhang (2011). Given the asymmetry of assets potentially transferred to the newlywed by the families of origin, this result may reveal an under-researched source of inter-generational wealth inequality by gender, whose consequences are not yet completely understood and explored.

6 Robustness Checks

We conducted several robustness tests to support the findings in the main results. First, two alternative measures of share of daughters among children are employed in the analysis: one where the presence of a daughter in a household is coded as one and the absence as zero, and another where the firstborn child being a daughter is coded as one, otherwise zero. The results are shown in Table 8. The coefficients of 'Firstborn daughter \times [23, 26]' in columns 1 and 3 of panel A reveal that as the firstborn daughter approaches marriageable age, such households exhibit a marked decrease in saving rate than their male counterparts. Additionally, the

coefficients of ‘Having a daughter × [23, 26]’ in columns 1 and 3 of panel B suggest that households with daughters demonstrate the lowest savings rate than the families with sons when the first child reaches marriageable age. However, such an effect is not presented in the stock market participation. These patterns support the main findings.

[Table 8 here]

To address potential omitted variable bias in our primary analysis, we conduct two formal tests, based on Oster (2019) and Diegert et al. (2022), respectively.²⁴ The findings, presented in Table 9, support that the results are robust against omitted variable bias. Specifically, panel A, based on Oster (2019), shows, for savings decisions, the estimates of $\delta (R^2_{max}=1.3 R^2)$, the ratio of unobserved to observed heterogeneity required to invalidate the regression coefficients of interest. This was obtained from (i) OLS utilising 2013-2019 CHFS data (columns 1) and (ii) panel FE regression using 2013 and 2015 panel data (columns 3). This ratio is consistently above the critical threshold of 1, as discussed by Oster (2019), supporting that the results are indeed robust to unobserved heterogeneity.²⁵

Panel B presents the results based on the methodology proposed by Diegert et al. (2022): the \bar{r}_x breakdown point is estimated to be 36.5% using the OLS performed on 2013-2019 CHFS data (39% for the short panel data from 2013 and 2015 CHFS). These figures indicate that the influence of unobserved variables would need to be at least 36.5% (or 39%) as large as that of observed variables to negate the observed ‘daughter’ effect on saving decisions. The omitted variables are not deemed to have such considerable impacts (Diegert et al., 2022), supporting the result presented in panel A.

[Table 9 here]

In our analysis, substituting negative savings rates with zero raises concerns about losing crucial information on households experiencing ‘real’ negative savings. In our study, a significant

²⁴ Oster’s test relies on the exogenous control assumption, while the method newly developed by Diegert et al. (2022) is under the assumption of endogenous control.

²⁵ The command ‘*psacalc*’ for the Oster test does not support testing the panel data RE regressions, so we apply it to the panel data FE regressions.

proportion (approximately 46%) of the surveyed families report a zero saving rate.²⁶ Given the differences between households with a zero savings rate and those with a positive savings rate, presented in Appendix Table A1. 4 and Table A1. 5, the possibility of sample bias exists and is a threat to the results discussed so far.

[Table 10 here]

To address this potential problem, we apply three alternative approaches. First, we follow Campbell and Mankiw (1991), Chamon and Prasad (2010), Deaton and Paxson (1994) and Wei and Zhang (2011) and alternatively use ‘log(income/consumption)’ to measure the ‘real’ saving rate. This approach allows us to retain saving rate information for samples with zero or negative saving rates, thereby avoiding their exclusion. The findings, presented in columns 1 and 3 of Table 10, align with the results on Table 3: specifically, there are statistically significant negative coefficients associated with the interaction term ‘SOD × [23, 26]’. However, within the subsample of households with a zero saving rate (columns 2 and 4), this effect, while positive, does not attain statistical significance.

[Table 11 here]

Second, we assign a dummy variable equal to one for households with positive savings rates (zero otherwise), and interact it with all the control variables and the fixed effects for province and year. The results, as shown in Table 11, are consistent with the original results in Table 3.

[Table 12 here]

Third, we follow Koné et al. (2019) and apply a Heckman-type selection model to account for the possible non-random missingness of data. In particular, we employ Heckman’s two-stage estimation. In the first stage, we examine the likelihood of having a positive saving rate on a

²⁶ The families with a zero saving rate might have a zero or negative saving rate. The literature highlights the considerable variability in household savings rates across China, attributed to pronounced income inequality and saving motivations. Gan et al. (2014) report that around half of Chinese families have a zero or negative saving rate, with distinct motivations for saving between urban and rural areas. Financing children’s education is a universal saving goal, yet urban families also emphasise housing and rural families allocate more for children’s weddings. Kong and Dickinson (2016) find that households with positive saving rates have higher incomes than those with negative or zero saving rates. Stratford and Cowling (2016) emphasise the divergent saving patterns among different income groups in China, pointing out that families, particularly those in poverty or rural areas, often grapple with insufficient income to save.

set of demographic and other characteristics to obtain the Inverse Mills Ratio (IMR). In the second stage, we add the IMR as a control and perform the regression on the sub-sample with only positive saving rates. The results, summarised in Table 12, show that the coefficients for the interaction term ‘SOD × age of firstborn child [23, 26]’ in columns 2 and 4 align with the original findings. An alternative specification - the two-part fractional model (Appendix Table A1. 6) - produces equivalent results.

To mitigate endogeneity bias from the unobserved heterogeneity correlated with explanatory variables, we perform Mundlak (1978) estimation by adding the individuals’ mean of the time-varying variables as additional regressors.²⁷ This method combines the advantages of FE and RE models. The results, shown in Table 13, are similar to RE estimations. The differences between the coefficients of the interaction term ‘SOD × [23, 26]’ obtained from RE and Mundlak methods are not statistically different.²⁸

[Table 13 here]

Finally, we assess the potential bias required to challenge the inferences drawn from a Rubin causal model (Xu et al., 2019).²⁹ By utilising the ‘*konfound*’ command on the estimate of saving rate on interaction term ‘SOD × [23, 26]’ in panel data RE regression, our analysis reveals that, to negate the obtained coefficient, 43.89% of the observed cases would need to exhibit a null effect at a 10% significance level. In FE regression, the threshold is 26.76%. Given the improbability of this high threshold, our findings appear robust and support the empirical evidence presented.

Overall, the robustness checks results support our general findings.

7 Mechanisms

²⁷ The Mundlak estimator provides consistent estimates even when the individual effects are correlated with the explanatory variables.

²⁸ The Z - statistics for the coefficient differences is 1.175, computed as $Z = (\hat{\beta}_{RE} - \hat{\beta}_{Mundlak}) / \sqrt{(s^2(\hat{\beta}_{RE}) + s^2(\hat{\beta}_{Mundlak}))}$, where $\hat{\beta}_{RE}$ and $\hat{\beta}_{Mundlak}$ are coefficient estimates from RE and Mundlak estimations, and $s^2(\cdot)$ are squared standard errors of the coefficients.

²⁹ ‘*konfound*’ quantifies the robustness of inferences to potential biases, such as unobserved variables, providing the impact of an omitted variable needed to invalidate an inference. (Xu et al., 2019).

We test four potential channels through which the share of daughters is associated with household saving decisions, particularly when the (first) child approaches the marriageable stage. Lundberg (2005) suggests that behavioural differences stem from individual preferences and constraints. Given China's social and traditional norms, we focus on the preference for a son, home-buying intention, consumption of daily necessities, and education expenditure to encapsulate the economic and sociocultural pressures families face during their first child's marriageable stage. The results are shown in Table 14.

[Table 14 here]

First, we examine the role of preference for a son. When the first child reaches marriageable age, the families with more daughters appear to have an unbalanced appreciation for the gender of their child, as suggested by the positive correlation between the interaction term 'SOD × [23, 26]' and the stated preference for a son (column 1 of panel A). However, son preference is not significantly related to savings and stock market investment decisions.

Then, we explore the role of home-buying intention. Homeownership significantly boosts unmarried men's competitiveness in Chinese marriage markets (Wei and Zhang, 2011; Wei and Zhang, 2016). Supporting this, Sun and Zhang (2020) contend that the likelihood of home acquisition peaks during the marriage year. It is a traditional norm that parents predominantly shoulder the financial responsibility of home purchases. Having more daughters when the firstborn child is of marriageable age significantly and negatively impacts home purchase intent (column 1 of panel B), consistent with the hypothesis that families who have sons face stronger financial pressures to secure additional housing to boost their son's marriage prospects. In addition, home-buying intention increases the household saving rate, as suggested by the positive and significant correlation between home-buying intention and saving rate, while the relationship between the interaction term SOD × [23, 26] and the saving rate is still significant but slightly smaller after controlling the home-buying intention.

Cultural norms and social institutions often foster the preference for a son, leading to unequal resource allocation where daughters receive fewer necessities, educational opportunities, and investment assets (Kaul, 2018; Azam and Kingdon, 2013; Lundberg, 2005; Blau et al., 2020; Pasqua, 2005). Considering marital traditions, the dowry system, which is contingent upon the discretion of the bride's family, generally incurs lower financial obligations than the bride price,

traditionally provided by the groom's side (Anderson, 2007). Consequently, families with more daughters may encounter less financial strain as the eldest child reaches marriageable age, leading to a greater likelihood of increased consumption and a focus on immediate needs. Thus, we examine the role of daily necessities expenditure and children's education expenditure. As shown in column 1 of panel C, when the firstborn child approaches marriageable age, the family with more daughters will consume more daily necessities, which is in line with our assumption. The daily necessities consumption is negatively related to the saving rate, while the estimated coefficient between the interaction term $SOD \times [23, 26]$ and the saving rate is slightly smaller but significant. Lastly, when the firstborn child approaches marriageable age, the family with more daughters will not have higher or lower educational expenditure, as suggested by the insignificant estimated coefficient (column 1 of panel D). This means that daily necessities consumption is a channel, while children's education expenditure is not.

8 Heterogeneity

We use the 2013 and 2015 CHFS panel data to explore possible variations in the results across various subgroups. The results are shown in Table 15.

[Table 15 here]

We first examine the role of financial literacy, a crucial tool for informed decision-making related to finance (Lusardi, 2008; Abreu and Mendes, 2010; Grohmann et al., 2018), to study whether better-informed families behave differently from the less informed. We hence split the sample into families with low and high financial literacy scores, respectively, and present the results in panel A.³⁰ The estimates in column 1 and column 3 suggest that the negative correlation between the ' $SOD \times [23,26]$ ' and the household's saving rate emerges exclusively in families with low financial literacy. No such daughter effect emerges in households with high financial literacy.

This result supports the possibility that improving financial literacy may raise families' likelihood of making financial decisions based on impartial market information rather than subjective cultural habits and norms. The results also reveal a possible area of target

³⁰ Respondents are classified as having high financial literacy if they correctly answer two or more out of three financial literacy questions; those who answer zero or one question correctly are deemed to have low financial literacy.

intervention: the observed bias in the financial choices (in this case, under-savings) of households with daughters (alternatively, over-saving in households with sons).

As financial literacy is a proxy of education, we re-run the regressions separately by the level of the education of the household, finding similar results (column 1 in panel B): low-education families make savings decisions based on the share of daughters and the firstborn child's age bracket.³¹ This effect is statistically nil among those with higher educational attainment (column 3 in panel B). This result is relevant, as it restricts the possible treatment to a well-defined subgroup of society: those with relatively low education attainment– a possibly vulnerable group, as it has accumulated substantial assets but does not have corresponding financial literacy and formal education to gauge investment risks.

In addition, as gender stereotypes and low income are more prevalent in rural areas than in urban areas (Lin et al., 2021; Wu and Perloff, 2005; Wei and Zhang, 2016), we re-run the analysis separately by residential location (panel C). The drop effect of the share of daughters in the household savings rate when the first child is at the marriageable age is more prominent in rural areas (column 1) than in urban areas (column 3). This suggests that the effect found is related to the educational level prevailing in the family as well as the location where it lives. This, in turn, narrows the scope for possible intervention.

Last, given the different risk preferences for men and women (Croson and Gneezy, 2009), we re-estimate the families with a male head and those with a female head (panel D). The 'daughter' effect is larger in families headed by a male head than in those headed by a female head, as evident from the estimated coefficients in columns 1 and 3. This finding further refines the target demographic for interventions. This also encourages women to take a more active role in household saving decisions, helping to reduce the savings disparity between households with sons and those with daughters.

9 Conclusion

Our study employs the pooled four waves (2013, 2015, 2017 and 2019) of CHFS, along with a panel dataset generated from the 2013 and 2015 CHFS, to investigate the impact of child

³¹ Possessing a university or college degree qualifies respondents as having high education, whereas those without are deemed to have low education.

gender on the household savings rate and participation in the stock market. Having a daughter consistently reduces households' saving rate, especially as the daughter and the first child (or daughter) approach marriageable age but does not alter the risk preferences of the household, as proxied by having a stock market account. The possible under-investment and endowments of daughters may carry undesirable consequences to the development of capital markets because it can negatively influence the supply of savings that would otherwise be available, hence conditioning the financial products that are designed and offered to retail investors.

While cultural norms are difficult to change in a short time, the results open up two possible alternatives for policy analysis and intervention. The first is a focus on the legislative environment guiding the inter-generational transmission of wealth and the associated filial responsibilities and expectations. For example, the adoption of the Law of Succession of the People's Republic of China equates inheritance rights between genders, while the Marriage Law and the Chinese Constitution mandate adult children to support their aging parents irrespective of their gender or marital status. Ensuring the enforcement of these laws and refining them are essential to address the disparities in rights and duties between adult men and women.

Related to this, are matrimonial customs. By striving for parity in the dowry and bride price, allowing the parents of marriageable sons and those of marriageable daughters to have similar financial pressures in this context, it may be possible to mitigate the disparities in the familial valuation of sons and daughters. This balance could pave the way for more equitable treatment within the family structure.

An alternative area of possible consideration is the promotion of better formal education on financial opportunities and choices, as well as gender equity among low-educated families, especially if resident in rural areas. Notwithstanding gender differences in risk preferences, better formal education and incentives can ensure that savings and investments associated with female and male children start from similar positions and do not become a source of inter-generational gender inequality.

References

- Abreu, M. & Mendes, V. 2010. Financial literacy and portfolio diversification. *Quantitative Finance*, 10, 515-528 <https://doi.org/10.1080/14697680902878105>.
- Adam, L. 1947. Virilocal and Uxorilocal. *American Anthropologist*, 49, 678-678 <https://doi.org/10.1525/aa.1947.49.4.02a00220>.
- Almenberg, J. & Dreber, A. 2015. Gender, stock market participation and financial literacy. *Economics Letters*, 137, 140-142 <https://doi.org/10.1016/j.econlet.2015.10.009>.
- Amran, N. A. & Ahmad, A. C. 2010. Family succession and firm performance among Malaysian companies. *International Journal of Business and Social Science*, 1, 193-203 <https://doi.org/10.17576/ajag-2011-2-6538>.
- Anderson, A., Baker, F. & Robinson, D. T. 2017. Precautionary savings, retirement planning and misperceptions of financial literacy. *Journal of Financial Economics*, 126, 383-398 <https://doi.org/10.1016/j.jfineco.2017.07.008>.
- Anderson, S. 2007. The Economics of Dowry and Brideprice. *Journal of Economic Perspectives*, 21, 151-174 <https://doi.org/10.1257/jep.21.4.151>.
- Anukriti, S., Bhalotra, S. & Tam, E. H. 2022. On the quantity and quality of girls: Fertility, parental investments and mortality. *The Economic Journal*, 132, 1-36 <https://doi.org/10.1093/ej/ueab035>.
- Asadullah, M. N., Mansoor, N., Randazzo, T. & Wahhaj, Z. 2021. Is son preference disappearing from Bangladesh? *World Development*, 140, 105353 <https://doi.org/10.1016/j.worlddev.2020.105353>.
- Azam, M. & Kingdon, G. G. 2013. Are Girls the Fairer Sex in India? Revisiting Intra-Household Allocation of Education Expenditure. *World Development*, 42, 143-164 <https://doi.org/10.1016/j.worlddev.2012.09.003>.
- Banister, J. 1987. *China's changing population*, Stanford University Press.
- Bannier, C. E. & Schwarz, M. 2018. Gender- and education-related effects of financial literacy and confidence on financial wealth. *Journal of Economic Psychology*, 67, 66-86 <https://doi.org/10.1016/j.joep.2018.05.005>.
- Barcellos, S. H., Carvalho, L. S. & Lleras-Muney, A. 2014. Child Gender and Parental Investments In India: Are Boys And Girls Treated Differently? *American Economic Journal: Applied Economics*, 6, 157-189 <https://doi.org/10.1257/app.6.1.157>.
- Bianchi, M. 2018. Financial Literacy and Portfolio Dynamics. *The Journal of Finance*, 73, 831-859 <https://doi.org/10.1111/jofi.12605>.
- Blau, F. D., Kahn, L. M., Brummund, P., Cook, J. & Larson-Koester, M. 2020. Is there still son preference in the United States? *Journal of Population Economics*, 33, 709-750 <https://doi.org/10.1007/s00148-019-00760-7>.
- Bogan, V. 2009. Stock Market Participation and the Internet. *Journal of Financial and Quantitative Analysis*, 43, 191-211 <https://doi.org/10.1017/s0022109000002799>.
- Bogan, V. L. 2013. Household investment decisions and offspring gender: parental accounting. *Applied Economics*, 45, 4429-4442 <https://doi.org/10.1080/00036846.2013.788782>.

- Botticini, M. & Siow, A. 2003. Why Dowries? *American Economic Review*, 93, 1385-1398 <https://doi.org/10.1257/000282803769206368>.
- Campbell, J. Y. & Mankiw, N. G. 1991. The response of consumption to income: A cross-country investigation. *European Economic Review*, 35, 723-756 [https://doi.org/10.1016/0014-2921\(91\)90033-F](https://doi.org/10.1016/0014-2921(91)90033-F).
- Chamon, M. D. & Prasad, E. S. 2010. Why Are Saving Rates of Urban Households in China Rising? *American Economic Journal: Macroeconomics*, 2, 93-130 <https://doi.org/10.1257/mac.2.1.93>.
- Chan, K. W. & Zhang, L. 1999. The Hukou System and Rural-Urban Migration in China: Processes and Changes. *The China Quarterly*, 160, 818-855 <https://doi.org/10.1017/S0305741000001351>.
- Choi, E. J. & Hwang, J. 2015. Child Gender and Parental Inputs: No More Son Preference in Korea? *American Economic Review*, 105, 638-643 <https://doi.org/10.1257/aer.p20151118>.
- Clark, T. S. & Linzer, D. A. 2015. Should I Use Fixed or Random Effects? *Political Science Research and Methods*, 3, 399-408 <https://doi.org/10.1017/psrm.2014.32>.
- Crosen, R. & Gneezy, U. 2009. Gender Differences in Preferences. *Journal of Economic Literature*, 47, 448-74 <https://doi.org/10.1257/jel.47.2.448>.
- Curtis, C. C., Lugauer, S. & Mark, N. C. 2015. Demographic Patterns and Household Saving in China. *American Economic Journal: Macroeconomics*, 7, 58-94 <https://doi.org/10.1257/mac.20130105>.
- De Laiglesia, J. R. & Morrisson, C. 2008. Household structures and savings: Evidence from Household Surveys. OECD Discussion Papers nr. 267 <https://doi.org/10.1787/245640263431>.
- Deaton, A. S. & Paxson, C. 1994. Saving, growth, and aging in Taiwan. *Studies in the Economics of Aging*. University of Chicago Press.
- Diegert, P., Masten, M. A. & Poirier, A. 2022. Assessing omitted variable bias when the controls are endogenous. <https://doi.org/10.48550/arXiv.2206.02303>.
- Ebenstein, A. 2010. The "Missing Girls" of China and the Unintended Consequences of the One Child Policy. *Journal of Human Resources*, 45, 87-115, <https://www.jstor.org/stable/20648938>.
- Edlund, L. 1999. Son Preference, Sex Ratios, and Marriage Patterns. *Journal of Political Economy*, 107, 1275-1304 <https://doi.org/10.1086/250097>.
- Edlund, L. 2000. The Marriage Squeeze Interpretation of Dowry Inflation: A Comment. *Journal of Political Economy*, 108, 1327-1333 <https://doi.org/10.1086/317688>.
- Fonseca, R., Mullen, K. J., Zamarro, G. & Zissimopoulos, J. 2012. What explains the gender gap in financial literacy? The role of household decision making. *Journal of Consumer Affairs*, 46, 90-106 <https://doi.org/10.1111/j.1745-6606.2011.01221.x>.
- Gan, L., Yin, Z., Jia, N., Xu, S., Ma, S. & Zheng, L. 2014. Data you need to know about China. *Springer Books*, <https://doi.org/10.1007/978-3-642-38151-5>.
- Gangadharan, L. 2003. Testing for Son Preference in South Africa. *Journal of African Economics*, 12, 371-416 <https://doi.org/10.1093/jae/12.3.371>.

- Gao, Y., Alessie, R. & Angelini, V. 2023. Parental housing wealth and children's marriage prospects in China—evidence from CHARLS. *Review of Economics of the Household*, 21, 615-644 <https://doi.org/10.1007/s11150-022-09608-8>.
- Gao, Z., Pang, J. & Zhou, H. 2022. The economics of marriage: Evidence from China. *Humanities and Social Sciences Communications*, 9, 399 <https://doi.org/10.1057/s41599-022-01400-4>.
- Ge, S., Yang, D. T. & Zhang, J. 2018. Population policies, demographic structural changes, and the Chinese household saving puzzle. *European Economic Review*, 101, 181-209 <https://doi.org/10.1016/j.euroecorev.2017.09.008>.
- Giavazzi, F. & McMahon, M. 2012. Policy Uncertainty and Household Savings. *The Review of Economics and Statistics*, 94, 517-531 https://doi.org/10.1162/REST_a_00158.
- Greenhalgh, S. 2013. Patriarchal Demographics? China's Sex Ratio Reconsidered. *Population and Development Review*, 38, 130-149, <http://www.jstor.org/stable/23655291>.
- Grinblatt, M., Keloharju, M. & Linnainmaa, J. 2011. IQ and Stock Market Participation. *The Journal of Finance*, 66, 2121-2164 <https://doi.org/10.1111/j.1540-6261.2011.01701.x>.
- Grohmann, A., Klühs, T. & Menkhoff, L. 2018. Does financial literacy improve financial inclusion? Cross country evidence. *World Development*, 111, 84-96 <https://doi.org/10.1016/j.worlddev.2018.06.020>.
- Grossbard-Shechtman, S. 2003. A consumer theory with competitive markets for work in marriage. *The Journal of Socio-Economics*, 31, 609-645 [https://doi.org/10.1016/S1053-5357\(02\)00138-5](https://doi.org/10.1016/S1053-5357(02)00138-5).
- Hank, K. & Kohler, H.-P. 2000. Gender preferences for children in Europe: Empirical results from 17 FFS countries. *Demographic Research*, 2, <https://www.jstor.org/stable/26347999>.
- Horioka, C. Y. & Terada-Hagiwara, A. 2017. The impact of sex ratios before marriage on household saving in two Asian countries: The competitive saving motive revisited. *Review of Economics of the Household*, 15, 739-757 <https://doi.org/10.1007/s11150-016-9344-y>.
- Ji, X., Wang, K., Xu, H. & Li, M. 2021. Has Digital Financial Inclusion Narrowed the Urban-Rural Income Gap: The Role of Entrepreneurship in China. *Sustainability*, 13, 8292 <https://doi.org/10.3390/su13158292>.
- Kabátek, J. & Ribar, D. C. 2020. Daughters and Divorce. *The Economic Journal*, 131, 2144-2170 <https://doi.org/10.1093/ej/ueaa140>.
- Karbownik, K. & Myck, M. 2017. Who gets to look nice and who gets to play? Effects of child gender on household expenditures. *Review of Economics of the Household*, 15, 925-944 <https://doi.org/10.1007/s11150-016-9328-y>.
- Kaul, T. 2018. Intra-household allocation of educational expenses: Gender discrimination and investing in the future. *World Development*, 104, 336-343 <https://doi.org/10.1016/j.worlddev.2017.12.017>.
- Keyzers, L. 1991. Where are the missing Chinese girls? *Newsletter (Women's Global Network on Reproductive Rights)*, 15, <https://pubmed.ncbi.nlm.nih.gov/12284524/>.

- Kim, S. & Lee, S.-H. 2020. Son Preference and Fertility Decisions: Evidence From Spatiotemporal Variation in Korea. *Demography*, 57, 927-951 <https://doi.org/10.1007/s13524-020-00875-7>.
- Klinger-Vartabedian, L. & Wispe, L. 1989. Age differences in marriage and female longevity. *Journal of Marriage and the Family*, 195–202 <https://doi.org/10.2307/352380>.
- Koné, S., Bonfoh, B., Dao, D., Koné, I. & Fink, G. 2019. Heckman-type selection models to obtain unbiased estimates with missing measures outcome: theoretical considerations and an application to missing birth weight data. *BMC Medical Research Methodology*, 19, 231 <https://doi.org/10.1186/s12874-019-0840-7>.
- Kong, D. & Dickinson, D. 2016. Investigating the Impact of Income on Savings Using a Chinese Household Level Dataset. *Emerging Markets Finance & Trade*, 52, 1775-1796, <https://www.jstor.org/stable/26753025>.
- Li, L. & Wu, X. 2011. Gender of Children, Bargaining Power, and Intrahousehold Resource Allocation in China. *Journal of Human Resources*, 46, 295-316 <https://doi.org/10.3368/jhr.46.2.295>.
- Li, W., Song, C., Xu, S. & Yi, J. 2022. High Sex Ratios and Household Portfolio Choice in China. *Journal of Human Resources*, 57, 465-490 <https://doi.org/10.3368/jhr.57.2.1217-9245R2>.
- Li, W. & Yang, Dennis t. 2005. The Great Leap Forward: Anatomy of a Central Planning Disaster. *Journal of Political Economy*, 113, 840-877 <https://doi.org/10.1086/430804>.
- Lin, C., Sun, Y. & Xing, C. 2021. Son Preference and Human Capital Investment Among China's Rural-Urban Migrant Households. *The Journal of Development Studies*, 57, 2077-2094 <https://doi.org/10.1080/00220388.2021.1961750>.
- Longhi, S. & Nandi, A. 2015. A Practical Guide to Using Panel Data. 55 City Road, London: SAGE Publications, Inc.
- Love, D. A. 2010. The Effects of Marital Status and Children on Savings and Portfolio Choice. *Review of Financial Studies*, 23, 385-432 <https://doi.org/10.1093/rfs/hhp020>.
- Lugauer, S., Ni, J. & Yin, Z. 2019. Chinese household saving and dependent children: Theory and evidence. *China Economic Review*, 57, 101091 <https://doi.org/10.1016/j.chieco.2017.08.005>.
- Lundberg, S. 2005. Sons, Daughters, and Parental Behaviour. *Oxford Review of Economic Policy*, 21, 340-356 <https://doi.org/10.1093/oxrep/gri020>.
- Lusardi, A. 2008. Financial literacy: an essential tool for informed consumer choice? *NBER Working Paper*, <https://doi.org/10.3386/w14084>.
- Lusardi, A. & Mitchell, O. S. 2011. Financial literacy and planning: Implications for retirement wellbeing. *NBER Working Paper*, <https://doi.org/10.3386/w17078>.
- Lusardi, A. & Mitchell, O. S. 2017. How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness. *Quarterly Journal of Finance*, 07, 1750008 <https://doi.org/10.1142/s2010139217500082>.
- Meng, X. 2012. Labor Market Outcomes and Reforms in China. *Journal of Economic Perspectives*, 26, 75-102 <https://doi.org/10.1257/jep.26.4.75>.

- Meng, X., Qian, N. & Yared, P. 2015. The Institutional Causes of China's Great Famine, 1959–1961. *The Review of Economic Studies*, 82, 1568-1611 <https://doi.org/10.1093/restud/rdv016>.
- Milazzo, A. 2014. Son preference, fertility and family structure: Evidence from reproductive behavior among Nigerian women. *World Bank Policy Research Working Paper*, <https://doi.org/10.1596/1813-9450-6869>.
- Mills, M. & Begall, K. 2010. Preferences for the sex-composition of children in Europe: a multilevel examination of its effect on progression to a third child. *Population Studies*, 64, 77-95 <https://doi.org/10.1080/00324720903497081>.
- Modigliani, F. & Cao, S. L. 2004. The Chinese Saving Puzzle and the Life-Cycle Hypothesis. *Journal of Economic Literature*, 42, 145-170 <https://doi.org/10.1257/002205104773558074>.
- Morduch, J. 1995. Income smoothing and consumption smoothing. *Journal of Economic Perspectives*, 9, 103-114 <https://doi.org/10.1257/jep.9.3.103>.
- Nguyen, T. K. T. & Nguyen, K. M. 2020. Role of Financial Literacy and Peer Effect in Promotion of Financial Market Participation: Empirical Evidence in Vietnam. *The Journal of Asian Finance, Economics and Business*, 7, 1-8 <https://doi.org/10.13106/jafeb.2020.vol7.no6.001>.
- Oster, E. 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37, 187-204 <https://doi.org/10.1080/07350015.2016.1227711>.
- Pahlevan Sharif, S., Ahadzadeh, A. S. & Turner, J. J. 2020. Gender Differences in Financial Literacy and Financial Behaviour Among Young Adults: The Role of Parents and Information Seeking. *Journal of Family and Economic Issues*, 41, 672-690 <https://doi.org/10.1007/s10834-020-09674-z>.
- Pasqua, S. 2005. Gender Bias in Parental Investments in Children's Education: A Theoretical Analysis. *Review of Economics of the Household*, 3, 291-314 <https://doi.org/10.1007/s11150-005-3459-x>.
- Pitt, M. M., Rosenzweig, M. R. & Hassan, N. 2012. Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy. *American Economic Review*, 102, 3531-3560 <https://doi.org/10.1257/aer.102.7.3531>.
- Platteau, J.-P. & Gaspart, F. 2007. The Perverse Effects of High Brideprices. *World Development*, 35, 1221-1236 <https://doi.org/10.1016/j.worlddev.2006.09.017>.
- Porter, M. 2016. How do sex ratios in China influence marriage decisions and intra-household resource allocation? *Review of Economics of the Household*, 14, 337-371 <https://doi.org/10.1007/s11150-014-9262-9>.
- Rossi, P. & Rouanet, L. 2015. Gender Preferences in Africa: A Comparative Analysis of Fertility Choices. *World Development*, 72, 326-345 <https://doi.org/10.1016/j.worlddev.2015.03.010>.
- Sen, A. 2003. Missing women--revisited. *BMJ*, 327, 1297-8 <https://doi.org/10.1136/bmj.327.7427.1297>.

- Shrestha, V. & Jung, J. 2023. Healthcare reform and gender specific infant mortality in rural Nepal. *World Development*, 161, 106018 <https://doi.org/10.1016/j.worlddev.2022.106018>.
- Song, Y. 2014. What should economists know about the current Chinese hukou system? *China Economic Review*, 29, 200-212 <https://doi.org/10.1016/j.chieco.2014.04.012>.
- Song, Y. & Gao, J. 2023. Do fathers have son preference in the United States? Evidence from paternal subjective well-being. *Review of Economics of the Household*, 21, 1083-1117 <https://doi.org/10.1007/s11150-022-09640-8>.
- Stratford, K. & Cowling, A. 2016. Chinese household income, consumption and savings. *RBA Bulletin*, 9, 31-40, <https://www.rba.gov.au/publications/bulletin/2016/sep/4.html>.
- Sun, A. & Zhang, Q. 2020. Who marries whom in a surging housing market? *Journal of Development Economics*, 146, 102492 <https://doi.org/10.1016/j.jdeveco.2020.102492>.
- Treiman, D. J. 2014. *Quantitative data analysis: Doing social research to test ideas*, John Wiley & Sons.
- Van Rooij, M. C., Lusardi, A. & Alessie, R. J. 2012. Financial literacy, retirement planning and household wealth. *The Economic Journal*, 122, 449-478 <https://doi.org/10.1111/j.1468-0297.2012.02501.x>.
- Van Rooij, M. C. J., Lusardi, A. & Alessie, R. 2011. Financial literacy and stock market participation. *Journal of Financial Economics*, 101, 449-472 <https://doi.org/10.1016/j.jfineco.2011.03.006>.
- Vestman, R. 2019. Limited stock market participation among renters and homeowners. *The Review of Financial Studies*, 32, 1494-1535 <https://doi.org/10.1093/rfs/hhy089>.
- Wang, C., Akgüç, M., Liu, X. & Tani, M. 2020. Expropriation with hukou change and labour market outcomes in China. *China Economic Review*, 60, 101391 <https://doi.org/10.1016/j.chieco.2019.101391>.
- Wang, H., Cheng, Z. & Smyth, R. 2022. Parental misbeliefs and household investment in children's education. *Economics of Education Review*, 89, 102284 <https://doi.org/10.1016/j.econedurev.2022.102284>.
- Wang, J., Zhang, D. & Wang, Z. 2023. Digital finance, stock market participation and asset allocation of Chinese households. *Applied Economics Letters*, 30, 1870-1873 <https://doi.org/10.1080/13504851.2022.2083556>.
- Wei, S. & Zhang, X. 2011. The competitive saving motive: Evidence from rising sex ratios and savings rates in China. *Journal of Political Economy*, 119, 511-564 <https://doi.org/10.1086/660887>.
- Wei, S.-J. & Zhang, X. 2016. The Competitive Saving Motive: Concept, Evidence, and Implications. *Frontiers of Economics in China*, 11, 355-366 <https://doi.org/10.3868/s060-005-016-0020-2>.
- Wei, S.-J., Zhang, X. & Liu, Y. 2012. Status Competition and Housing Prices. *NBER Working Paper*, <https://doi.org/10.3386/w18000>.
- White, T. 2006. *China's longest campaign: Birth planning in the People's Republic, 1949-2005*, Cornell University Press.

- Wu, X. & Perloff, J. M. 2005. China's income distribution, 1985–2001. *Review of Economics and Statistics*, 87, 763-775 <https://doi.org/10.1162/003465305775098206>.
- Xu, R., Frank, K. A., Maroulis, S. J. & Rosenberg, J. M. 2019. konfound: Command to quantify robustness of causal inferences. *The Stata Journal: Promoting Communications on Statistics and Stata*, 19, 523-550 <https://doi.org/10.1177/1536867x19874223>.
- Yang, X. & Einstein, G. 2014. The impact of subjective norms on the intention to select fetal sex in rural China. *Women's Studies International Forum*, 44, 69-79 <https://doi.org/10.1016/j.wsif.2013.09.004>.
- Zou, J. & Deng, X. J. 2019. Financial literacy, housing value and household financial market participation: Evidence from urban China. *China Economic Review*, 55, 52-66 <https://doi.org/10.1016/j.chieco.2019.03.008>.

Tables

Table 1. Descriptive statistics of dependent variables

	Mean			
	All	Families with daughter	Families without daughter	Difference ($H_0: \text{diff} = 0$)
	(1)	(2)	(3)	(4)
Saving rate	0.238 (0.279)	0.219 (0.271)	0.260 (0.286)	-0.041*** [0.003]
Having a stock account (Yes = 1)	0.142 (0.349)	0.132 (0.338)	0.153 (0.360)	-0.021*** [0.004]
Observations	27,926	14,731	13,195	

Notes: The data is from the 2013, 2015, 2017 and 2019 China Household Finance Survey. Standard deviations are given in parentheses. Standard errors are given in squared parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Difference = the mean of families with daughter (column 2) – the mean of families without daughter (column 3). A *t*-test is applied.

Table 2. Descriptive statistics of independent variables and control variables

	Mean			
	All	Families with daughter	Families without daughter	Difference ($H_0: diff = 0$)
	(1)	(2)	(3)	(4)
Share of daughter (SOD)	0.428 (0.442)			
The age of firstborn child:				
1 ≤ age ≤ 12	0.305 (0.460)	0.306 (0.461)	0.303 (0.460)	0.002 [0.006]
13 ≤ age ≤ 17	0.192 (0.394)	0.204 (0.403)	0.178 (0.383)	0.025*** [0.005]
18 ≤ age ≤ 22	0.230 (0.421)	0.241 (0.427)	0.218 (0.413)	0.023*** [0.005]
23 ≤ age ≤ 26	0.164 (0.370)	0.164 (0.370)	0.165 (0.371)	-0.001 [0.004]
27 ≤ age ≤ 35	0.110 (0.312)	0.086 (0.281)	0.136 (0.343)	-0.050*** [0.004]
Residential and hukou status:				
Urban resident with non-agricultural hukou	0.502 (0.500)	0.471 (0.499)	0.536 (0.499)	-0.065*** [0.006]
Urban resident with agricultural hukou	0.245 (0.430)	0.261 (0.439)	0.228 (0.420)	0.032*** [0.005]
Rural resident with non-agricultural hukou	0.013 (0.114)	0.012 (0.111)	0.014 (0.118)	-0.002 [0.001]
Rural resident with agricultural hukou	0.240 (0.427)	0.256 (0.436)	0.221 (0.415)	0.035*** [0.005]
The number of children	1.384 (0.622)	1.592 (0.711)	1.152 (0.390)	0.440*** [0.007]
Income (thousand yuan)	95.241 (213.989)	92.888 (208.666)	97.868 (219.757)	-4.980* [2.565]
Asset (thousand yuan)	1181.691 (4914.803)	1130.494 (2302.243)	1238.846 (6723.164)	-108.352* [58.908]
Age group of household head: 25 ≤ age ≤ 35	0.179 (0.383)	0.183 (0.386)	0.174 (0.379)	0.009* [0.005]
36 ≤ age ≤ 45	0.401 (0.490)	0.428 (0.495)	0.371 (0.483)	0.057*** [0.006]
46 ≤ age ≤ 55	0.336 (0.472)	0.324 (0.468)	0.350 (0.477)	-0.025*** [0.006]
56 ≤ age ≤ 65	0.084 (0.277)	0.065 (0.246)	0.105 (0.307)	-0.040*** [0.003]
Financial literacy score	1.034 (0.920)	1.016 (0.921)	1.054 (0.918)	-0.038*** [0.011]
Risk attitude	2.186 (1.207)	2.192 (1.202)	2.180 (1.212)	0.012 [0.014]
Female	0.508 (0.500)	0.506 (0.500)	0.511 (0.500)	-0.005 [0.006]
Education level: No schooling/primary school	0.188 (0.391)	0.200 (0.400)	0.175 (0.380)	0.025*** [0.006]
Junior high/senior high school	0.566 (0.496)	0.569 (0.495)	0.563 (0.496)	0.006 [0.007]
Vocational diploma	0.232 (0.422)	0.219 (0.413)	0.246 (0.431)	-0.028*** [0.006]
University degree	0.014 (0.119)	0.013 (0.113)	0.016 (0.125)	-0.003* [0.002]
Marital status (married = 1, unmarried = 0)	0.953	0.960	0.946	0.013***

	(0.211)	(0.197)	(0.225)	[0.003]
Employment status (employed = 1, unemployed = 0)	0.787 (0.409)	0.785 (0.411)	0.789 (0.408)	-0.004 [0.005]
Observations	27,926	14,731	13,195	

Notes: The data is from 2013, 2015, 2017 and 2019 CHFS. Standard deviations are given in parentheses. Standard errors are given in squared parentheses. *p < 0.1, **p < 0.05, ***p < 0.01. Difference = the mean of families with daughter (column 2) – the mean of families without daughter (column 3). A *t*-test is applied.

Table 3. Daughters, saving rate and stock market participation

	OLS		RE		FE	
	Saving rate	Having a stock account (Yes =1)	Saving rate	Having a stock account (Yes =1)	Saving rate	Having a stock account (Yes =1)
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: regression results</i>						
SOD	-0.011 (0.008)	0.010 (0.008)	-0.010 (0.008)	0.011 (0.014)	-0.013 (0.037)	-0.022 (0.026)
The age of firstborn child (ref: 1 ≤ age ≤ 12):						
13 ≤ age ≤ 17	-0.013** (0.006)	0.002 (0.009)	-0.019** (0.008)	0.002 (0.013)	-0.003 (0.019)	-0.024 (0.023)
18 ≤ age ≤ 22	-0.009 (0.007)	-0.011 (0.008)	-0.015* (0.009)	-0.020* (0.011)	0.017 (0.026)	-0.051* (0.026)
23 ≤ age ≤ 26	0.055*** (0.008)	-0.014 (0.008)	0.057*** (0.010)	-0.026** (0.011)	0.099*** (0.029)	-0.071** (0.030)
27 ≤ age ≤ 35	0.103*** (0.009)	-0.021** (0.009)	0.100*** (0.013)	-0.030** (0.012)	0.103*** (0.032)	-0.088** (0.035)
SOD × [13, 17]	-0.006 (0.008)	-0.024 (0.017)	-0.006 (0.010)	-0.022 (0.024)	-0.012 (0.027)	0.001 (0.028)
SOD × [18, 22]	-0.019** (0.009)	-0.003 (0.012)	-0.022** (0.011)	-0.006 (0.015)	-0.051 (0.037)	0.013 (0.032)
SOD × [23, 26]	-0.038*** (0.011)	-0.010 (0.015)	-0.051*** (0.015)	0.001 (0.019)	-0.100** (0.037)	0.019 (0.036)
SOD × [27, 35]	-0.009 (0.013)	0.016 (0.018)	-0.018 (0.015)	0.002 (0.025)	-0.046 (0.048)	0.026 (0.048)
The number of children	-0.015*** (0.004)	-0.017*** (0.005)	-0.014** (0.006)	-0.016*** (0.005)	-0.001 (0.016)	0.021** (0.009)
Residential and hukou status (ref: urban resident with non-agricultural hukou):						
Urban resident with agricultural hukou	-0.002 (0.006)	-0.073*** (0.007)	-0.005 (0.008)	-0.070*** (0.008)	-0.013 (0.023)	-0.035*** (0.012)
Rural resident with non-agricultural hukou	0.007 (0.015)	-0.111*** (0.012)	0.014 (0.017)	-0.121*** (0.010)	-0.009 (0.062)	-0.019 (0.016)
Rural resident with agricultural hukou	0.006 (0.009)	-0.071*** (0.008)	-0.001 (0.011)	-0.081*** (0.009)	-0.106 (0.065)	-0.013 (0.017)
Income (log)	0.056*** (0.002)	0.007*** (0.001)	0.057*** (0.002)	0.006*** (0.001)	0.052*** (0.002)	0.003** (0.001)
Asset (log)	0.002 (0.002)	0.034*** (0.004)	0.002 (0.002)	0.027*** (0.003)	0.000 (0.003)	0.015*** (0.002)
Age group of household head (ref: 25 ≤ age ≤ 35):						
36 ≤ age ≤ 45	-0.000 (0.006)	0.061*** (0.011)	-0.005 (0.007)	0.047*** (0.008)	-0.024 (0.015)	-0.001 (0.016)
46 ≤ age ≤ 55	0.002 (0.006)	0.078*** (0.014)	-0.004 (0.008)	0.063*** (0.011)	-0.005 (0.020)	0.006 (0.017)
56 ≤ age ≤ 65	0.011 (0.008)	0.095*** (0.018)	-0.003 (0.012)	0.071*** (0.015)	0.008 (0.034)	0.034 (0.027)
Financial literacy score	0.000 (0.002)	0.030*** (0.003)	-0.002 (0.003)	0.018*** (0.004)	-0.003 (0.004)	0.008* (0.004)
Risk attitude	-0.002 (0.001)	0.041*** (0.004)	0.000 (0.002)	0.029*** (0.003)	0.004 (0.003)	0.012*** (0.002)
Female (ref: male)	-0.005 (0.003)	0.004 (0.005)	-0.005 (0.004)	0.001 (0.007)		
Education level (ref: No schooling/primary school):						
Senior high/junior high	-0.008* (0.004)	-0.013** (0.005)	-0.016** (0.007)	-0.010** (0.005)	-0.033 (0.024)	0.000 (0.009)
Vocational diploma	0.009	0.073***	0.005	0.077***	-0.091**	0.021

	(0.007)	(0.010)	(0.009)	(0.014)	(0.038)	(0.040)
University degree	0.054***	0.132***	0.041***	0.148***	-0.109**	-0.076
	(0.009)	(0.013)	(0.014)	(0.017)	(0.046)	(0.068)
Married (ref: unmarried)	0.008	-0.012	0.009	-0.007	0.018	-0.014
	(0.007)	(0.008)	(0.012)	(0.011)	(0.025)	(0.020)
Employed (ref: unemployed)	0.032***	-0.015***	0.027***	-0.015**	0.025**	-0.008
	(0.004)	(0.005)	(0.005)	(0.006)	(0.010)	(0.008)
Province sex ratio	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,926	27,926	13,806	13,806	13,806	13,806
R-squared	0.244	0.240				
R2 within			0.162	0.015	0.165	0.022
R2 overall			0.246	0.231	0.196	0.063
R2 between			0.282	0.276	0.209	0.073

Panel B: a linear combination of parameters

SOD + SOD × [1, 12]	-0.011	0.010	-0.010	0.011	-0.013	-0.022
	(0.008)	(0.008)	(0.008)	(0.014)	(0.037)	(0.026)
SOD + SOD × [13, 17]	-0.017***	-0.013	-0.016**	-0.011	-0.025	-0.021
	(0.005)	(0.013)	(0.008)	(0.019)	(0.042)	(0.027)
SOD + SOD × [18, 22]	-0.030***	0.007	-0.032***	0.006	-0.063	-0.009
	(0.008)	(0.011)	(0.010)	(0.012)	(0.039)	(0.023)
SOD + SOD × [23, 26]	-0.049***	-0.000	-0.061***	0.013	-0.112***	-0.004
	(0.008)	(0.011)	(0.012)	(0.015)	(0.036)	(0.030)
SOD + SOD × [27, 35]	-0.020*	0.027	-0.027*	0.013	-0.058	0.004
	(0.010)	(0.017)	(0.014)	(0.022)	(0.057)	(0.037)

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at province level. In the panel data fixed effects regression, we drop household head's gender and province fixed effects since they are time-invariant.

Table 4. Second stage of G2SLS-RE-IV Regressions

	Saving rate (1)	Having a stock account (Yes =1) (2)
<i>Panel A: regression results</i>		
SOD	-0.006 (0.007)	0.010 (0.014)
The age of firstborn child (ref: $1 \leq \text{age} \leq 12$):		
$13 \leq \text{age} \leq 17$	-0.018** (0.007)	0.003 (0.014)
$18 \leq \text{age} \leq 22$	-0.013 (0.010)	-0.018 (0.012)
$23 \leq \text{age} \leq 26$	0.059*** (0.012)	-0.021* (0.012)
$27 \leq \text{age} \leq 35$	0.105*** (0.013)	-0.028** (0.011)
SOD \times [13, 17]	-0.007 (0.011)	-0.025 (0.027)
SOD \times [18, 22]	-0.027** (0.014)	-0.005 (0.016)
SOD \times [23, 26]	-0.051*** (0.019)	-0.007 (0.020)
SOD \times [27, 35]	-0.031* (0.017)	-0.001 (0.022)
Control variables	Yes	Yes
Province sex ratio	Yes	Yes
Province FE	Yes	Yes
Year FE	Yes	Yes
Observations	13,441	13,441
R2 within	0.162	0.014
R2 overall	0.246	0.232
R2 between	0.279	0.274
Hansen J statistic	7.586	4.961
Hansen J p-value	0.181	0.421
<i>Panel B: a linear combination of parameters</i>		
SOD + SOD \times [1, 12]	-0.006 (0.007)	0.010 (0.014)
SOD + SOD \times [13, 17]	-0.014 (0.010)	-0.015 (0.021)
SOD + SOD \times [18, 22]	-0.033*** (0.011)	0.004 (0.014)
SOD + SOD \times [23, 26]	-0.058*** (0.017)	0.002 (0.016)
SOD + SOD \times [27, 35]	-0.038** (0.015)	0.009 (0.019)

Notes: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are clustered at province level. Control variables are listed in Table 2.

Table 5. One-child households

	OLS		RE	
	Saving rate (1)	Having a stock account (Yes =1) (2)	Saving rate (3)	Having a stock account (Yes =1) (4)
<i>Panel A: regression results</i>				
Daughter (ref: son)	-0.012 (0.008)	0.010 (0.009)	-0.012 (0.008)	0.019 (0.016)
The age of child (ref: 1 ≤age≤12):				
13 ≤age≤ 17	-0.008 (0.007)	0.001 (0.012)	-0.017* (0.010)	0.007 (0.017)
18 ≤age≤ 22	-0.004 (0.009)	-0.016 (0.011)	-0.010 (0.012)	-0.033** (0.016)
23 ≤age≤ 26	0.058*** (0.010)	-0.012 (0.013)	0.061*** (0.013)	-0.032* (0.017)
27 ≤age≤ 35	0.104*** (0.011)	-0.018 (0.013)	0.095*** (0.018)	-0.040** (0.017)
Daughter × [13, 17]	-0.010 (0.009)	-0.015 (0.020)	0.003 (0.010)	-0.022 (0.029)
Daughter × [18, 22]	-0.021** (0.009)	-0.002 (0.014)	-0.024* (0.013)	-0.014 (0.018)
Daughter × [23, 26]	-0.026** (0.013)	-0.010 (0.018)	-0.028* (0.016)	-0.001 (0.025)
Daughter × [27, 35]	-0.007 (0.014)	0.019 (0.020)	0.000 (0.017)	0.007 (0.027)
Control variables	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	18,836	18,836	8,822	8,822
R-squared	0.246	0.237		
R2 within			0.168	0.022
R2 overall			0.253	0.232
R2 between			0.287	0.279
<i>Panel B: a linear combination of parameters</i>				
Daughter + Daughter × [1, 12]	-0.012 (0.008)	0.010 (0.009)	-0.012 (0.008)	0.019 (0.016)
Daughter + Daughter × [13, 17]	-0.023*** (0.006)	-0.005 (0.015)	-0.009 (0.007)	-0.003 (0.023)
Daughter + Daughter × [18, 22]	-0.033*** (0.009)	0.008 (0.013)	-0.036*** (0.013)	0.005 (0.014)
Daughter + Daughter × [23, 26]	-0.039*** (0.010)	0.000 (0.015)	-0.041*** (0.014)	0.018 (0.022)
Daughter + Daughter × [27, 35]	-0.019* (0.011)	0.029 (0.019)	-0.012 (0.014)	0.025 (0.024)

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at province level. In the panel data regression, we adopt RE regression as the variable of interest, child gender, is time-invariant. Control variables are listed in Table 2.

Table 6. Multi-child households

	OLS	Having a stock account (Yes =1)	RE	Having a stock account (Yes =1)	FE	Having a stock account (Yes =1)
	Saving rate		Saving rate		Saving rate	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: regression results</i>						
SOD	0.007 (0.016)	0.007 (0.026)	-0.001 (0.016)	-0.026 (0.027)	-0.007 (0.185)	0.059 (0.076)
The age of firstborn child (ref: 1 ≤age≤12):						
13 ≤age≤ 17	-0.022 (0.014)	0.003 (0.020)	-0.009 (0.019)	-0.019 (0.020)	0.096* (0.049)	-0.078** (0.029)
18 ≤age≤ 22	-0.016 (0.013)	-0.007 (0.019)	-0.025* (0.014)	-0.009 (0.021)	0.022 (0.053)	-0.072** (0.035)
23 ≤age≤ 26	0.064*** (0.014)	-0.021 (0.021)	0.065*** (0.022)	-0.021 (0.017)	0.121* (0.059)	-0.062 (0.042)
27 ≤age≤ 35	0.113*** (0.023)	-0.024 (0.015)	0.132*** (0.029)	-0.010 (0.029)	0.091 (0.091)	-0.029 (0.069)
SOD × [13, 17]	0.001 (0.023)	-0.047 (0.028)	-0.038 (0.027)	0.007 (0.031)	-0.154** (0.075)	0.116** (0.054)
SOD × [18, 22]	-0.020 (0.020)	-0.016 (0.026)	-0.013 (0.024)	0.009 (0.032)	-0.036 (0.086)	0.098* (0.050)
SOD × [23, 26]	-0.082*** (0.028)	-0.014 (0.030)	-0.094*** (0.036)	0.014 (0.027)	-0.140 (0.092)	0.074 (0.055)
SOD × [27, 35]	-0.030 (0.047)	-0.018 (0.026)	-0.079 (0.056)	-0.030 (0.047)	0.013 (0.164)	0.000 (0.098)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,090	9,090	4,984	4,984	4,984	4,984
R-squared	0.216	0.195				
R2 within			0.156	0.008	0.166	0.016
R2 overall			0.219	0.141	0.147	0.016
R2 between			0.239	0.161	0.137	0.017
<i>Panel B: a linear combination of parameters</i>						
SOD + SOD × [1, 12]	0.007 (0.016)	0.007 (0.026)	-0.001 (0.016)	-0.026 (0.027)	-0.007 (0.185)	0.059 (0.076)
SOD + SOD × [13, 17]	0.008 (0.015)	-0.040** (0.017)	-0.039* (0.024)	-0.019 (0.019)	-0.161 (0.161)	0.175* (0.092)
SOD + SOD × [18, 22]	-0.013 (0.012)	-0.008 (0.017)	-0.014 (0.015)	-0.016 (0.018)	-0.043 (0.169)	0.157** (0.070)
SOD + SOD × [23, 26]	-0.075*** (0.026)	-0.006 (0.012)	-0.095*** (0.034)	-0.012 (0.013)	-0.147 (0.180)	0.133* (0.072)
SOD + SOD × [27, 35]	-0.022 (0.040)	-0.010 (0.028)	-0.080 (0.056)	-0.055 (0.036)	0.006 (0.231)	0.059 (0.120)

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at province level. In the panel data fixed-effect regression, we drop household gender and province fixed effects, as they are time-invariant. Control variables are listed in Table 2.

Table 7. Multi-child households with daughter

	OLS		RE		FE	
	Saving rate	Having a stock account (Yes =1)	Saving rate	Having a stock account (Yes =1)	Saving rate	Having a stock account (Yes =1)
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: regression results</i>						
SOD	0.008 (0.022)	0.049 (0.034)	0.039 (0.024)	0.002 (0.030)	0.154 (0.227)	-0.084 (0.052)
The age of eldest daughter (ref: $1 \leq \text{age} \leq 12$):						
13 $\leq \text{age} \leq 17$	-0.008 (0.021)	0.009 (0.022)	0.054* (0.033)	0.011 (0.018)	0.112 (0.076)	-0.097** (0.045)
18 $\leq \text{age} \leq 22$	0.005 (0.025)	0.011 (0.019)	0.023 (0.030)	0.014 (0.024)	0.060 (0.092)	-0.128* (0.066)
23 $\leq \text{age} \leq 26$	0.119*** (0.026)	0.002 (0.025)	0.146*** (0.027)	-0.007 (0.023)	0.223* (0.115)	-0.120 (0.080)
27 $\leq \text{age} \leq 35$	0.100* (0.050)	0.015 (0.029)	0.128** (0.055)	-0.025 (0.032)	0.099 (0.176)	-0.127 (0.083)
SOD \times [13, 17]	-0.011 (0.029)	-0.052 (0.039)	-0.104** (0.043)	-0.027 (0.033)	-0.197* (0.103)	0.149* (0.081)
SOD \times [18, 22]	-0.041 (0.033)	-0.041 (0.032)	-0.061 (0.039)	-0.017 (0.039)	-0.110 (0.120)	0.200* (0.108)
SOD \times [23, 26]	-0.135*** (0.043)	-0.046 (0.041)	-0.162*** (0.046)	-0.006 (0.032)	-0.295* (0.155)	0.180 (0.113)
SOD \times [27, 35]	-0.018 (0.069)	-0.065 (0.043)	-0.080 (0.079)	0.003 (0.051)	-0.044 (0.267)	0.178 (0.110)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	No	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,193	7,193	3,981	3,981	3,981	3,981
R-squared	0.210	0.190				
R2 within			0.163	0.003	0.172	0.010
R2 overall			0.208	0.126	0.139	0.031
R2 between			0.225	0.150	0.125	0.035
<i>Panel B: a linear combination of parameters</i>						
SOD + SOD \times [1, 12]	0.008 (0.022)	0.049 (0.034)	0.039 (0.024)	0.002 (0.030)	0.154 (0.227)	-0.084 (0.052)
SOD + SOD \times [13, 17]	-0.003 (0.025)	-0.003 (0.023)	-0.065* (0.038)	-0.025 (0.027)	-0.043 (0.240)	0.065 (0.044)
SOD + SOD \times [18, 22]	-0.032 (0.027)	0.008 (0.022)	-0.021 (0.030)	-0.015 (0.028)	0.044 (0.257)	0.116* (0.068)
SOD + SOD \times [23, 26]	-0.127*** (0.038)	0.003 (0.020)	-0.123*** (0.042)	-0.005 (0.014)	-0.141 (0.277)	0.096 (0.075)
SOD + SOD \times [27, 35]	-0.010 (0.065)	-0.016 (0.036)	-0.041 (0.073)	0.005 (0.039)	0.110 (0.349)	0.095 (0.075)

Notes: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are clustered at province level. In the panel data FE regression, we drop household head gender and province FE, as they are time-invariant. Control variables are listed in Table 2.

Table 8. Alternative measurements of daughters

<i>Panel A: gender of firstborn child as alternative measurement</i>	OLS		RE	
	Saving rate	Having a stock account (Yes =1)	Saving rate	Having a stock account (Yes =1)
	(1)	(2)	(3)	(4)
Firstborn daughter (ref: firstborn son)	-0.009 (0.007)	0.006 (0.008)	-0.007 (0.006)	0.010 (0.012)
Age of the firstborn child (ref: 0≤age≤12)				
13 ≤age≤ 17	-0.013** (0.006)	0.002 (0.009)	-0.020*** (0.007)	0.002 (0.013)
18≤age≤ 22	-0.010 (0.008)	-0.013 (0.008)	-0.016* (0.009)	-0.020* (0.012)
23 ≤age≤ 26	0.050*** (0.009)	-0.014 (0.009)	0.050*** (0.011)	-0.022* (0.011)
27≤age≤ 35	0.104*** (0.009)	-0.020** (0.009)	0.102*** (0.012)	-0.028** (0.011)
Firstborn daughter × [13, 17]	-0.005 (0.008)	-0.024 (0.016)	-0.004 (0.009)	-0.023 (0.021)
Firstborn daughter × [18, 22]	-0.016 (0.009)	-0.001 (0.011)	-0.018* (0.011)	-0.005 (0.013)
Firstborn daughter × [23, 26]	-0.025** (0.010)	-0.011 (0.013)	-0.031** (0.013)	-0.009 (0.015)
Firstborn daughter× [27, 35]	-0.011 (0.012)	0.014 (0.015)	-0.024* (0.013)	-0.003 (0.019)
Control variables	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	27,926	27,926	13,806	13,806
R square	0.243	0.240		
R2 within			0.161	0.015
R2 overall			0.245	0.231
R2 between			0.281	0.276

<i>Panel B: having a daughter as alternative measurement</i>	OLS		RE	
	Saving rate	Having a stock account (Yes =1)	Saving rate	Having a stock account (Yes =1)
	(1)	(2)	(3)	(4)
Having a daughter (Yes = 1)	-0.012* (0.007)	0.007 (0.008)	-0.012 (0.007)	0.009 (0.013)
Age of the firstborn child (ref: 0≤age≤12)				
13 ≤age≤ 17	-0.013** (0.006)	0.005 (0.010)	-0.018** (0.008)	0.006 (0.014)
18 ≤age≤ 22	-0.010 (0.008)	-0.013 (0.009)	-0.015 (0.010)	-0.022* (0.012)
23 ≤age≤ 26	0.055*** (0.009)	-0.013 (0.010)	0.060*** (0.011)	-0.026** (0.013)
27 ≤age≤ 35	0.100*** (0.009)	-0.018* (0.010)	0.095*** (0.014)	-0.025* (0.013)
Having a daughter × [13, 17]	-0.006 (0.007)	-0.025* (0.014)	-0.007 (0.009)	-0.026 (0.020)
Having a daughter × [18, 22]	-0.014 (0.008)	-0.001 (0.010)	-0.018* (0.010)	-0.001 (0.013)
Having a daughter × [23, 26]	-0.030*** (0.009)	-0.010 (0.013)	-0.045*** (0.012)	0.000 (0.016)
Having a daughter × [27, 35]	-0.000 (0.011)	0.007 (0.015)	-0.003 (0.014)	-0.010 (0.019)
Control variables	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	27,926	27,926	13,806	13,806
R square	0.244	0.240		
R2 within			0.161	0.015
R2 overall			0.246	0.231
R2 between			0.282	0.276

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at province level. The control variables are listed in Table 2.

Table 9. Coefficients stability

Coefficients of interest	OLS		Panel data	
	Saving rate	Having a stock account (Yes = 1)	Saving rate	Having a stock account (Yes = 1)
	(1)	(2)	(3)	(4)
<i>Panel A: Exogenous controls (Oster (2019))</i>				
δ (when $R_{max}^2 = 1.3R^2$) SOD \times age of firstborn child [23, 26]	-5.647	0.436	-1.226	-2.342
<i>Panel B: Endogenous controls (Diegert et al. (2022))</i>				
\bar{r}_x (breakdown point) SOD \times age of firstborn child [23, 26]	36.5%	8.56%	39.00%	0.90%

Notes: Oster's δ is calculated using 'psacalc' Stata command, while \bar{r}_x breakdown points is calculated using 'regensitivity' Stata command. For details on these two methodologies and examples, see Oster (2019) and Diegert et al. (2022), respectively.

Table 10. Alternative measurement of saving rate: log(income/consumption)

	Income/Consumption (log)		Panel RE	
	OLS	Families with a 'zero' saving rate	Full	Families with a 'zero' saving rate
	(1)	(2)	(3)	(4)
SOD	-0.012 (0.017)	-0.002 (0.026)	-0.015 (0.021)	-0.035 (0.036)
The age of firstborn child (ref: $1 \leq \text{age} \leq 12$):				
$13 \leq \text{age} \leq 17$	-0.009 (0.019)	0.019 (0.026)	-0.011 (0.019)	0.023 (0.033)
$18 \leq \text{age} \leq 22$	-0.018 (0.018)	-0.042* (0.024)	-0.047 (0.029)	-0.075* (0.044)
$23 \leq \text{age} \leq 26$	0.061*** (0.022)	-0.067* (0.033)	0.048 (0.031)	-0.106** (0.051)
$27 \leq \text{age} \leq 35$	0.148*** (0.024)	-0.015 (0.035)	0.116*** (0.029)	-0.062 (0.049)
SOD \times [13, 17]	0.016 (0.027)	-0.006 (0.041)	0.007 (0.029)	-0.006 (0.049)
SOD \times [18, 22]	-0.034 (0.026)	0.021 (0.033)	-0.009 (0.032)	0.087** (0.041)
SOD \times [23, 26]	-0.065** (0.032)	0.025 (0.044)	-0.081* (0.044)	0.067 (0.071)
SOD \times [27, 35]	-0.044 (0.030)	-0.010 (0.060)	-0.067 (0.041)	-0.001 (0.083)
Control variables	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	27,275	12,229	13,494	6,240
R-squared	0.758	0.748		
R2 within			0.773	0.770
R2 overall			0.757	0.735
R2 between			0.751	0.728

Notes: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are clustered at province level. The OLS estimation employs the CHFS data from 2013 to 2019, whereas the panel RE model utilises a short panel dataset derived from the 2013 and 2015 CHFS.

Table 11: Controlling the difference between families with positive and zero saving rates

	Saving rate	
	OLS (1)	Panel RE (2)
SOD	-0.003 (0.004)	0.002 (0.004)
The age of firstborn child (ref: $1 \leq \text{age} \leq 12$):		
$13 \leq \text{age} \leq 17$	-0.007* (0.004)	-0.009 (0.006)
$18 \leq \text{age} \leq 22$	0.001 (0.004)	-0.000 (0.006)
$23 \leq \text{age} \leq 26$	0.029*** (0.005)	0.032*** (0.007)
$27 \leq \text{age} \leq 35$	0.048*** (0.006)	0.050*** (0.007)
SOD × [13, 17]	0.002 (0.005)	-0.001 (0.008)
SOD × [18, 22]	-0.013** (0.006)	-0.017** (0.007)
SOD × [23, 26]	-0.019*** (0.006)	-0.029*** (0.008)
SOD × [27, 35]	-0.021*** (0.007)	-0.030*** (0.010)
Control variables	Yes	Yes
Family with positive saving rate (Yes = 1)	Yes	Yes
Family with positive saving rate × control variables	Yes	Yes
Family with positive saving rate × Province sex ratio	Yes	Yes
Family with positive saving rate × Province FE	Yes	Yes
Family with positive saving rate × Year FE	Yes	Yes
Province sex ratio	Yes	Yes
Province FE	Yes	Yes
Year FE	Yes	Yes
Observations	27,926	13,806
R-squared	0.733	
R2 within		0.706
R2 overall		0.740
R2 between		0.759

Notes: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are clustered at province level. Control variables are listed in Table 2. The OLS estimation employs the CHFS data from 2013 to 2019, whereas the panel RE model utilises a short panel dataset derived from the 2013 and 2015 CHFS.

Table 12. Heckman two-stage selection model

	Pooled 2013-2019 CHFS		Panel obtained from 2013 and 2015 CHFS	
	First stage Having a positive saving rate (Yes = 1) (1)	Second stage Saving rate (2)	First stage Having a positive saving rate (Yes = 1) (3)	Second stage Saving rate (4)
SOD	-0.039 (0.036)	-0.005 (0.006)	-0.047 (0.062)	0.002 (0.010)
The age of firstborn child (ref: 1 ≤age≤12):				
13 ≤age≤ 17	0.025 (0.043)	-0.013* (0.007)	0.020 (0.069)	-0.017 (0.011)
18 ≤age≤ 22	-0.035 (0.043)	0.005 (0.008)	-0.079 (0.070)	0.000 (0.011)
23 ≤age≤ 26	0.063 (0.051)	0.059*** (0.009)	0.026 (0.083)	0.066*** (0.013)
27 ≤age≤ 35	0.202*** (0.058)	0.089*** (0.010)	0.117 (0.097)	0.092*** (0.015)
SOD × [13, 17]	0.004 (0.061)	0.002 (0.010)	-0.016 (0.098)	-0.002 (0.016)
SOD × [18, 22]	-0.054 (0.058)	-0.030*** (0.010)	-0.048 (0.093)	-0.037** (0.015)
SOD × [23, 26]	-0.086 (0.064)	-0.034*** (0.011)	-0.099 (0.107)	-0.052*** (0.017)
SOD × [27, 35]	0.028 (0.076)	-0.032*** (0.012)	0.114 (0.131)	-0.046** (0.019)
Inverse Mills Ratio		0.032*** (0.012)		0.033** (0.015)
Control variables	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	27,926	15,047	13,806	7,254

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Control variables are listed in Table 2.

Table 13. Mundlak regressions

	Saving rate		Having a stock account (Yes =1)	
	RE (1)	Mundlak (2)	RE (3)	Mundlak (4)
SOD	-0.010 (0.008)	-0.012 (0.037)	0.011 (0.014)	-0.023 (0.026)
The age of firstborn child (ref: $1 \leq \text{age} \leq 12$):				
$13 \leq \text{age} \leq 17$	-0.019** (0.008)	-0.005 (0.019)	0.002 (0.013)	-0.022 (0.023)
$18 \leq \text{age} \leq 22$	-0.015* (0.009)	0.015 (0.025)	-0.020* (0.011)	-0.048* (0.026)
$23 \leq \text{age} \leq 26$	0.057*** (0.010)	0.095*** (0.029)	-0.026** (0.011)	-0.066** (0.030)
$27 \leq \text{age} \leq 35$	0.100*** (0.013)	0.097*** (0.031)	-0.030** (0.012)	-0.082** (0.034)
SOD \times [13, 17]	-0.006 (0.010)	-0.012 (0.027)	-0.022 (0.024)	0.001 (0.028)
SOD \times [18, 22]	-0.022** (0.011)	-0.051 (0.037)	-0.006 (0.015)	0.013 (0.032)
SOD \times [23, 26]	-0.051*** (0.015)	-0.099*** (0.038)	0.001 (0.019)	0.019 (0.036)
SOD \times [27, 35]	-0.018 (0.015)	-0.045 (0.048)	0.002 (0.025)	0.026 (0.048)
Mundlak mean value	No	Yes	No	Yes
Control variables	Yes	Yes	Yes	Yes
Local sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	13,806	13,806	13,806	13,806

Notes: standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are clustered at province level. Control variables are listed in Table 2.

Table 14: Mechanism analysis

	Mediator	Saving rate	Having a stock account (Yes = 1)
	(1)	(2)	(3)
<i>Panel A: Son preference as mediator</i>			
Son preference (Yes = 1)		0.012 (0.009)	0.012 (0.009)
SOD × [23, 26]	0.045*** (0.017)	-0.056*** (0.015)	0.002 (0.019)
<i>Panel B: Home-buying intention as mediator</i>			
Willingness to buy a new apartment/house (Yes = 1)		0.010* (0.006)	0.016** (0.008)
SOD × [23, 26]	-0.066** (0.032)	-0.042*** (0.014)	0.004 (0.025)
<i>Panel C: Daily necessities consumption as mediator</i>			
Daily necessities consumption (log)		-0.040*** (0.002)	0.009*** (0.002)
SOD × [23, 26]	0.108* (0.058)	-0.042*** (0.015)	0.004 (0.018)
<i>Panel D: Children's education expenditure as mediator</i>			
Children's education expenditure (log)		-0.033*** (0.003)	0.009*** (0.002)
SOD × [23, 26]	-0.151 (0.098)	-0.028 (0.023)	0.018 (0.027)
SOD	Yes	Yes	Yes
Age group of the firstborn child	Yes	Yes	Yes
SOD × Age group of firstborn child	Yes	Yes	Yes
Control variables	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Control variables are listed in Table 2. The standard errors are clustered at the province level. Panel RE regressions are applied.

Son preference measurement: In 2013 and 2015 CHFS, respondents were asked, 'Is it better to have a boy than a girl?' to gauge their gender-bias preferences. The choices were 'A. Boy is better', 'B. Girl is better', and 'C. The same'. We recode this variable, so it has a value of 1 for A and zero for B or C.

Home-buying intention measurement: In 2013 and 2015 CHFS, the willingness to purchase a new home is recorded as one if the household intends to buy a new home and zero otherwise.

Table 15. Heterogeneity

	Saving rate	Having a stock account (Yes = 1)	Saving rate	Having a stock account (Yes = 1)
	(1)	(2)	(3)	(4)
<i>Panel A: Financial literacy</i>	<i>Low</i>		<i>High</i>	
SOD × [23, 26]	-0.063*** (0.018)	-0.013 (0.017)	-0.024 (0.027)	0.013 (0.042)
<i>Panel B: Educational attainment</i>	<i>Low</i>		<i>High</i>	
SOD × [23, 26]	-0.056*** (0.019)	-0.018 (0.022)	0.000 (0.034)	0.116 (0.071)
<i>Panel C: Resident place</i>	<i>Rural</i>		<i>Urban</i>	
SOD × [23, 26]	-0.086*** (0.025)	-0.014* (0.008)	-0.035* (0.019)	0.007 (0.025)
<i>Panel D: Gender of household head</i>	<i>Male</i>		<i>Female</i>	
SOD × [23, 26]	-0.061** (0.025)	-0.001 (0.026)	-0.039** (0.018)	0.011 (0.022)
Sex ratio	Yes	Yes	Yes	Yes
Age group of firstborn child	Yes	Yes	Yes	Yes
Sex ratio × Age group of firstborn child	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Control variables are listed in Table 2. The standard errors are clustered at the province level. Panel RE regressions are applied.

Appendix

Table A1. 1. Financial literacy questions

	2013	2015	2017	2019 ³³
Interest	Given a 4% interest rate, how much would you have after five years if you had 100 RMB at first? a. Under 120; b. Exactly 120; c. Over 120; d. Cannot figure out	Given a 4% interest rate, how much would you have in total after one year if you have 100 yuan deposited? a. Under 104; b. 104; c. Over 104; d. Cannot figure out	1. If the Annual interest rate is 4%. One saves 100 RMB in 1-year time deposit, how much could one withdraw in 1 year? a. Less than 104 RMB; b. Just 104 RMB; c. More than 104 RMB; d. Cannot figure out	1. If the Annual interest rate is 4%. One saves 100 RMB in 1-year time deposit, how much could one withdraw in 1 year? a. Less than 104 RMB; b. Just 104 RMB; c. More than 104 RMB; d. Cannot figure out
Inflation	With an interest rate of 5% and an inflation rate of 3%, after saving money in the bank for one year, can you buy more or less than last year? a. More than last year; b. The same as last year; c. Less than last year; d. Cannot figure out	With an interest rate of 5% and an inflation rate of 3%, the stuff you buy with the money you have saved in the bank for one year is a. More than last year; b. The same as last year; c. Less than last year; d. Cannot figure out	Suppose the interest rate of a 1-year time deposit is 5%, and the inflation rate is 3%. If one saves 100 RMB as a 1-year time deposit in the bank, how much could one buy with the money withdrawn from the previous 1-year time deposit? a. More than one year ago; b. The same as one year ago; c. Less than one year ago; d. Cannot figure out	Suppose the interest rate of a 1-year time deposit is 5%, and the inflation rate is 3%. If one saves 100 RMB as a 1-year time deposit in the bank, how much could one buy with the money withdrawn from the previous 1-year time deposit? a. More than one year ago; b. The same as one year ago; c. Less than one year ago; d. Cannot figure out
Risk	Do you think stocks have greater risks than equity funds? a. Yes; b. No; c. Never heard of stock; d. Never heard of equity fund; e. Never heard of neither	Which one do you think is riskier, stock or fund? a. Stock; b. Fund; c. Haven't heard about the stock; d. Haven't heard about fund; e. Neither of them has been heard about	Compare stock and fund. Which do you think is riskier? a. Stock; b. Fund; c. Never heard about stock; d. Never heard about fund; e. Never heard about any of them; f. The same	Compare Equity fund with Bond fund. Which do you think is riskier? a. Equity fund; b. Bond fund; c. Never heard about Equity fund; d. Never heard about Bond fund; e. Never heard about any of them; f. The same

³³ The risk question used in previous wave is removed from the 2019 questionnaire, so we use a similar one from the questionnaires

Table A1. 2. Panel data summarisation for time-varying variables

Variable		Mean	Std. dev.
<i>Panel A: dependent variables</i>			
Saving rate	Overall	0.232	0.278
	Between		0.231
	Within		0.162
Having a stock account (Yes = 1)	Overall	0.121	0.326
	Between		0.292
	Within		0.139
<i>Panel B: independent variables and control variables</i>			
SOD	Overall	0.427	0.435
	Between		0.431
	Within		0.051
Age group of firstborn child	Overall	1.606	1.307
	Between		1.288
	Within		0.292
Risk attitude	Overall	2.148	1.206
	Between		1.033
	Within		0.666
Age group of household head	Overall	1.320	0.830
	Between		0.814
	Within		0.221
Marital status	Overall	0.957	0.203
	Between		0.191
	Within		0.070
Financial literacy	Overall	0.974	0.890
	Between		0.729
	Within		0.526
Employed	Overall	0.800	0.400
	Between		0.357
	Within		0.197
Asset (log)	Overall	12.765	1.476
	Between		1.395
	Within		0.530
Income (log)	Overall	10.305	2.076
	Between		1.733
	Within		1.207
Education levels	Overall	2.140	0.858
	Between		0.843
	Within		0.119
Residential and <i>hukou</i> status	Overall	1.150	1.279
	Between		1.281
	Within		0.119
Number of children	Overall	1.432	0.656
	Between		0.653
	Within		0.142

Table A1. 3. First stage of G2SLS-RE-IV regression

	SOD (1)	SOD × [13, 17] (2)	SOD × [18, 22] (3)	SOD × [23, 26] (4)	SOD × [27, 35] (5)
Preference for a son	-0.019** (0.008)	0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.000)
Firstborn daughter	0.861*** (0.018)	-0.000 (0.001)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
Firstborn daughter × [13, 17]	-0.118*** (0.016)	0.738*** (0.024)	0.001 (0.001)	-0.000 (0.000)	-0.001** (0.000)
Firstborn daughter × [18, 22]	-0.126*** (0.015)	0.002** (0.001)	0.730*** (0.025)	0.000 (0.001)	-0.001** (0.000)
Firstborn daughter × [23, 26]	-0.137*** (0.017)	0.000 (0.001)	0.002* (0.001)	0.720*** (0.027)	-0.002*** (0.001)
Firstborn daughter × [27, 35]	-0.063*** (0.022)	-0.001 (0.001)	-0.001 (0.001)	-0.003*** (0.001)	0.802*** (0.031)
Preference for a son × [13, 17]	-0.005 (0.018)	-0.025 (0.018)	0.001 (0.002)	0.001 (0.001)	-0.000 (0.000)
Preference for a son × [18, 22]	0.008 (0.012)	-0.000 (0.001)	-0.009 (0.010)	-0.001 (0.001)	0.000 (0.001)
Preference for a son × [23, 26]	0.003 (0.018)	0.001 (0.001)	0.004** (0.002)	-0.022 (0.014)	0.001 (0.001)
Preference for a son × [27, 35]	0.015 (0.021)	0.002 (0.001)	0.003** (0.001)	0.002 (0.002)	-0.012 (0.019)
Age of firstborn child	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes
Province sex ratio	Yes	Yes	Yes	Yes	Yes
Province fixed effect	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. The standard errors are clustered at province level. Control variables are listed in Table 2.

Table A1. 4. Differences between families with a saving rate of zero and positive saving rate

	'0' saving rate (1)	Positive saving rate (2)	Difference (3)
Residential and hukou status:			
Urban resident with non-agricultural hukou	0.413 (0.492)	0.578 (0.494)	-0.165*** [0.006]
Urban resident with agricultural hukou	0.269 (0.444)	0.225 (0.417)	0.045*** [0.005]
Rural resident with non-agricultural hukou	0.014 (0.116)	0.013 (0.113)	0.001 [0.001]
Rural resident with agricultural hukou	0.304 (0.460)	0.184 (0.388)	0.120*** [0.005]
The number of children	1.472 (0.679)	1.310 (0.558)	0.162*** [0.008]
Income (thousand yuan)	38.506 (56.751)	143.801 (277.700)	-105.295*** [2.401]
Asset (thousand yuan)	859.796 (1740.506)	1457.206 (6486.459)	-597.410*** [55.223]
Age group of household head:			
25≤age≤35	0.174 (0.379)	0.182 (0.386)	-0.008* [0.005]
36≤age≤45	0.420 (0.494)	0.385 (0.487)	0.034*** [0.006]
46≤age≤55	0.336 (0.473)	0.336 (0.472)	0.000 [0.006]
56≤age≤65	0.070 (0.255)	0.096 (0.295)	-0.026*** [0.003]
Financial literacy score	0.914 (0.885)	1.136 (0.937)	-0.222*** [0.011]
Risk attitude	2.139 (1.221)	2.227 (1.193)	-0.088*** [0.015]
Female	0.513 (0.500)	0.504 (0.500)	0.010 [0.006]
Education level: No schooling/primary school	0.240 (0.427)	0.144 (0.351)	0.096*** [0.005]
Junior high/senior high school	0.594 (0.491)	0.542 (0.498)	0.052*** [0.006]
Undergraduate	0.159 (0.366)	0.294 (0.456)	-0.135*** [0.005]
Postgraduate	0.007 (0.082)	0.021 (0.143)	-0.014*** [0.001]
Marital status (married = 1, unmarried = 0)	0.946 (0.226)	0.959 (0.197)	-0.013*** [0.003]
Employment status (employed = 1, unemployed = 0)	0.748 (0.434)	0.821 (0.383)	-0.074*** [0.005]
Observations	12,879	15,047	27,926

Notes: The data is from 2013, 2015, 2017 and 2019 CHFS. Standard deviations are given in parentheses. Standard errors are given in squared parentheses. *p < 0.1, **p < 0.05, ***p < 0.01. Difference = the mean of families with zero saving rate (column 1) – the mean of families with positive saving rate (column 2). A *t*-test is applied.

Table A1. 5. Regression of having a positive saving rate on control variables

	Having a positive saving rate (Yes = 1) (1)
Number of children	-0.030*** (0.005)
Residential and hukou status (ref: urban resident with non-agricultural hukou):	
Urban resident with agricultural hukou	-0.025*** (0.007)
Rural resident with non-agricultural hukou	-0.001 (0.023)
Rural resident with agricultural hukou	-0.056*** (0.008)
Income (log)	0.109*** (0.001)
Asset (log)	-0.008*** (0.002)
Age group of household head (ref: 25≤age≤35):	
36≤age≤45	-0.005 (0.008)
46≤age≤55	0.031*** (0.008)
56≤age≤65	0.105*** (0.012)
Financial literacy score	0.004 (0.003)
Risk attitude	-0.010*** (0.002)
Female (ref: male)	-0.002 (0.006)
Education level (ref: No schooling/primary school):	
Senior high/junior high	0.010 (0.008)
Vocational diploma	0.029*** (0.011)
University degree	0.067*** (0.012)
Married (ref: unmarried)	0.013 (0.012)
Employed (ref: unemployed)	0.053*** (0.007)
Province sex ratio	Yes
Province FE	Yes
Year FE	Yes
Observations	27,926
R-squared	0.253

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. OLS is applied for pooled the 2013-2019 CHFS.

Table A1. 6. Two part fractional model

	Having a positive saving rate (Yes = 1) (1)	Saving rate (2)
SOD	-0.070 (0.063)	-0.011 (0.019)
The age of firstborn child (ref: $1 \leq \text{age} \leq 12$):		
$13 \leq \text{age} \leq 17$	0.035 (0.074)	-0.046** (0.023)
$18 \leq \text{age} \leq 22$	-0.060 (0.075)	0.023 (0.024)
$23 \leq \text{age} \leq 26$	0.107 (0.088)	0.183*** (0.028)
$27 \leq \text{age} \leq 35$	0.355*** (0.101)	0.275*** (0.030)
SOD \times [13, 17]	0.014 (0.105)	0.005 (0.033)
SOD \times [18, 22]	-0.086 (0.099)	-0.101*** (0.033)
SOD \times [23, 26]	-0.156 (0.111)	-0.100*** (0.035)
SOD \times [27, 35]	0.047 (0.134)	-0.096*** (0.035)
Control variables	Yes	Yes
Province sex ratio	Yes	Yes
Province FE	Yes	Yes
Year FE	Yes	Yes
Observations	27,926	15,047

Notes: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Control variables are listed in Table 2. We utilise 2013-2019 CHFS. The first stage estimates the binary component of the two-part fractional regression model using the logit function, and the second stage estimates the fractional component using a 'cloglog' function.

Appendix B

To explore the channels through which child gender affects household savings decisions, we utilise the following specification:

$$M_{it} = \beta_0 + \beta_1 SOD_{it} + \beta_2 AC_{it} + \beta_3 SOD_{it} \times AC_{it} + \sum_{i=4}^j B_j X_{it} + v_{it} + u_i \quad (1)$$

$$y_{it} = \alpha_0 + \alpha_1 M_{it} + \alpha_2 SOD_{it} + \alpha_3 AC_{it} + \alpha_4 SOD_{it} \times AC_{it} + \sum_{i=5}^j B_j X_{it} + q_{it} + p_i \quad (2)$$

where M_{it} represents mediating factors such as son preference, willingness to purchase housing, daily necessities spending, and education-related expenditure. Other variables are the same as in the baseline model. Here, the parameters of interest are β_1 , β_3 , α_1 and α_4 , which measure the extent to which the daughter's share, and its interaction with a marriageable firstborn child, mediate the impact on family savings' decisions.