NBER WORKING PAPER SERIES

FERTILITY AND SAVINGS:
THE EFFECT OF CHINA'S TWO-CHILD POLICY ON HOUSEHOLD SAVINGS

Scott R. Baker
Efraim Benmelech
Zhishu Yang
Qi Zhang

Working Paper 29856
http://www.nber.org/papers/w29856

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2022

We thank our discussant Nancy Qian as well as seminar participants at Tsinghua University and the NBER Conference on Innovative Data in Household Finance: Opportunities and Challenges. Dong Huang and Nianzu Zhang provided excellent research assistance. This research is supported by the NSFC [grant 72071117]. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Fertility and Savings: The Effect of China's Two-Child Policy on Household Savings
Scott R. Baker, Efraim Benmelech, Zhishu Yang, and Qi Zhang
NBER Working Paper No. 29856
March 2022
JEL No. D14,D31,E21,G51

ABSTRACT

China's high household savings rate has attracted great academic interest but remains a puzzle. Potential explanations include demographic, policy, and financial causes. Yet a lack of reliable microlevel data on household finances makes it difficult to assess the relative importance of each factor. This paper uses individual income and spending transactions linked to demographic characteristics and financial information on loan applications and credit availability from a large Chinese bank in Inner Mongolia. We match a large subset of bank customers to administrative records covering marriage and births and obtain a unique view into consumption and saving patterns around important life events. Our results point toward identifying income growth, financial instability, and credit access, rather than such directives as the one-child policy, as the primary causes of high levels of savings among Chinese households.

Scott R. Baker
Kellogg School of Management
Northwestern University
2211 Campus Drive
Evanston, IL 60208
and NBER
s-baker@kellogg.northwestern.edu

Efraim Benmelech
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
e-benmelech@kellogg.northwestern.edu

Zhishu Yang
School of Economics & Management
Tsinghua University
Beijing, 100084
P.R. CHINA
yangzhsh@sem.tsinghua.edu.cn

Qi Zhang
800 Dongchuan RD. Minhang District,
Shanghai,
China
zhang.qi@sjtu.edu.cn

I. Introduction

China's high household savings rate, relative to both advanced economies and developing countries, has attracted great interest and prompted a large body of research into explanations. Notably, these high savings rates have come amid decades of substantial industrialization, income growth, and economic and social policy change. Given the scale of the Chinese economy and its significant share of the global economy, Chinese households' high savings rate has played a major role in the global savings glut, affecting global interest rates and asset prices (Bernanke et al. 2011; Caballero, Farhi, and Gourinchas 2008).

Explanations for China's high household savings rate span demographic, financial, and social causes. Yet data to determine the relative importance of each potential explanation are lacking. One obstacle to research is the dearth of reliable microlevel data on household characteristics and finances. Although the accessibility of (and research using) such transaction data in the United States has increased dramatically in recent years, few papers have accessed data from Chinese households, a void which this paper fills.[2] The use of such data allows for a more granular understanding of the dynamics of household financial behavior and a cleaner identification of the drivers of spending and savings.

We use eight years (2010–17) of data from a large Chinese bank located primarily in the province of Inner Mongolia. This bank has substantial coverage of the province's population, spanning over 1.5 million retail customers and 3.5 million financial accounts. We are able to observe individual income and spending transactions from these financial accounts as well as to link additional demographic characteristics and financial information on loan applications and credit availability for these users. We can also match a large subset of these customers to administrative records covering marriage and births to give us a unique window into consumption and saving patterns around the timing of important life events.

We first demonstrate that many stylized facts about household financial behavior in developed economies are mirrored among households in China. For instance, we note that the marginal propensity to consume is decreasing in income and wealth. In addition, using an instrumental variable strategy that leverages changes in coal prices among workers in the coal

[2] See Baker and Kueng (2021) for an overview of research using such household financial data. Chen, Qian, and Wen (2020) leverage Chinese household financial transactions to investigate responses to the Covid-19 pandemic.

mining sector, we find that consumption responses are particularly sensitive to unanticipated changes in income.

In investigating the drivers of high household savings rates, we focus on two primary areas: motives linked to financial volatility and motives linked to demographics or demographic policy. We find strong evidence that financial volatility drives substantial increases in Chinese household savings rates. In contrast, while many link the imposition of the one-child policy to higher levels of saving, we find that the loosening of this policy tended to increase savings rates rather than reducing them, at least in the short run.

Aiming at the first strand of explanations, which considers economic and financial motives, we examine whether precautionary saving motives such as income volatility, income growth rates, or access to consumer credit explain Chinese households' propensity to save. Such uncertainty over the future path of income may result in households desiring substantial savings buffers in case of negative realizations (see Jappelli and Pistaferri 2014). Many papers point to the volatility of income and lack of a social safety net in China as one explanation for high rates of household saving.

Chamon, Liu, and Prasad (2013) argue that rising income uncertainty and pension reforms account for two-thirds of the increase in China's urban savings rate. In a similar vein, Choi, Lugauer, and Mark (2017) note high levels of income growth and volatility, suggesting that over 80 percent of household savings in China stems from the precautionary motive. He et al. (2018) find that precautionary savings account for about 40 percent of state-owned enterprise (SOE) households' wealth accumulation using a reform to Chinese SOEs. Several papers also document a negative relation between social security benefits and household savings rates.[3]

Our data allow us to obtain detailed insights into household income volatility over time. In general, we show that Chinese households tend to respond to growth in income, income volatility, and credit constraints in ways similar to other Western households. Across a range of specifications, households facing higher levels of past income fluctuations tend to save much more of their income.

[3] For example, Bai and Wu (2014) document a rise in consumption after the coverage of China's New Cooperative Medical Scheme, while Feng, He, and Sato (2011) find an increase in household savings rates following a pension reform in 1995–97 that reduced the target replacement ratio compared to the pre-reform period.

In conjunction with income growth and volatility, researchers have highlighted a lack of access to credit or other financial buffers as a precautionary motive for high savings rates. As one example, Coeurdacier, Guibaud, and Jin (2015) argue that growth differentials and household credit constraints can explain about a third of the divergence in aggregate saving rates across emerging and developed economies. Wen (2010) and Bussière et al. (2013) also find that borrowing constraints contribute to high household savings rates. We follow Guiso, Sapienza, and Zingales (2004) and measure credit accessibility as the probability of receiving a loan conditional on financial and demographic observables. This access to credit is highly predictive of savings rates and consistent with a precautionary savings motive.

A second area of explanations for China's high household savings rate focuses more on demographic and social factors. The sharp increase in the Chinese savings rate coincided with the implementation of the One-Child Policy (OCP) in the 1980s. According to Curtis, Lugauer, and Mark (2015), China's demographic structure can affect saving rates through three main channels: (i) fewer childcare expenses; (ii) fewer intergenerational transfers in old ages from children; and (iii) a higher share of middle-aged individuals in the population in the future.

Several studies highlight the intergenerational transfer channel. Choukhmane, Coeurdacier, and Jin (2019) construct a model and propose that the OCP explains at least 30 percent of the rise in saving. Similarly, Zhou (2014) shows that having an additional brother reduces household savings by at least 5 percentage points. Further, İmrohoroğlu and Zhao (2018) suggest that the combination of the risks faced by the elderly and the deterioration of intergenerational supports may account for half of the increase in the savings rate between 1980 and 2010.

However, Banerjee et al. (2014) argue that the negative correlation between fertility and household savings can be offset by general equilibrium forces. Also, Chamon and Prasad (2010) argue that savings are best explained by rising private expenditures on housing, education, and health care, as well as financial underdevelopment in China.[4]

[4] Apart from family sizes and age structure, Wei and Zhang (2011) argue that sex ratio imbalances and competition in the marriage market can lead parents to increase savings to improve a son's relative attractiveness for marriage. They support their results with province- and county-level data. Du and Wei (2013) provide a quantitative model for this explanation. We find evidence suggesting parents reduce their saving after their children get married, which is consistent with Wei and Zhang (2011)'s argument.

Using our high-frequency household transaction data, we investigate the dynamics of household savings rates surrounding life events: marriage and the birth of children. Childbirth differentially affects the income and spending behavior of men and women, with men tending to have increases in income after a child's birth, while women experience temporary income declines. Marriage coincides with changes in income and spending patterns, with income increasing in the quarters leading up to marriage and stabilizing afterward.

We then identify impacts of the loosening of the OCP on household savings decisions. Using a triple difference approach, we examine the difference in financial savings after the policy change between households who have zero or one child (treated group) and those who already have two children (due to preexisting allowances for subsets of the population to have more than one child). Rather than decreasing rates of savings, we find that households for whom the relaxation of the policy constituted a reduction in constraints on having additional children tended to *increase* savings rates significantly in the post-policy period.

This finding helps to extend and clarify the model proposed in Choukhmane, Coeurdacier, and Jin (2019), where savings rates are affected by the OCP through two channels, a transfer channel (fewer old age supports) and an expenditure channel (fewer childcare expenses). While the transfer channel would suggest a decrease in savings rate upon a loosening of the OCP, the expenditures channel would predict the opposite. Our results suggest that, at least during our sample window, the expenditure channel dominates the transfer channel following the relaxation of the OCP.

We also find that households who increased savings the most had the highest propensities to have additional children. Given the costs of raising additional children, these estimates would imply that the imposition of the OCP may have in fact depressed savings, at least in the short term, because households did not need to save for additional child-related expenditures. Overall, these results suggest that the primary causes of Chinese high savings are financial instability and credit access, alongside rapid income growth, and not the one-child policy.

In a decomposition exercise, we find that income growth tends to explain the lion's share of within-sample savings rate fluctuations. That is, while other factors have strong cross-sectional predictive power in savings rates, the dynamics of these other factors cannot explain changes in Chinese savings rate during our sample period. However, this decomposition is limited to our sample window (2010-2017) which may feature distinct trends from those observed in other similar papers (eg. 1970s for Banerjee et al (2014) or 1995-2005 for Chamon and Prasad (2010)). We further show that the strong relationship between income growth and savings rates is mirrored in a larger set of OECD economies.

The rest of the paper is organized as follows. In Section II, we present the data and variables used in the analysis. In Section III, we study demographic and financial factors that determine saving rates in China. Section IV evaluates the effect of the one-child policy on Chinese' saving rates. Section V demonstrates the robustness of our results to linking individuals into household units. Section VI examines the relative explanatory power of financial and demographic factors in time series. Section VII concludes.

II. Data and Sample Validation

Our main data set comes from a large Inner Mongolia commercial bank and spans the years 2010–17 and is described in detail in the Data Appendix. The bank is headquartered in Hohhot, the largest city in Inner Mongolia, and has more than 100 branches in the other seven largest cities in the province. Although it is a regional bank and is smaller than the four largest state-owned banks in China, it is one of Inner Mongolia's largest banks, with a substantial and comprehensive provincial customer base. Specifically, at the end of 2017, the bank had nearly 1.8 million retail customers with over 3.6 million separate financial accounts.

A. Household Transaction Data

Our transaction data spans over 140 million checking, savings, and credit card account transactions, each of which provides information on the amount, date, account balance, merchant information, and a textual description of the transaction. We aggregate the transaction-level data into quarterly level in the following analyses to account for regular cash withdrawals and any other periodic fluctuations within quarters.

The data set also contains detailed demographic information about each customer, including gender, date of birth, city of birth, and current city of residence. We assign customers into an industry based on their employer. We also augment the data with administrative data about marital status and childbirth. These data include customers' date of marriage, date of divorce (if any), and gender and date of birth of every child they have.

We aggregate across all accounts owned by an individual, canceling out any intra-individual transfers between accounts. As is true for much research employing data from a single large bank or financial institution (e.g., Ganong and Noel 2019; Agarwal and Qian 2014), it is possible that individuals in our sample may have accounts in other banks. To limit our sample to those who primarily use this bank, we restrict our sample to individuals who receive continuous non-trivial paycheck income during our sample period and have at least 8 quarters of data. We also drop individual with only a single account that keeps low balances or transacts very infrequently.

Our final sample consists of 571,748 quarterly observations from 37,100 individuals. Individuals are then weighted based on age, gender, and industry in order to match the provincial population distribution provided by the National Bureau of Statistics (NBS).

B. Income, Consumption, and Savings Rates

Income and consumption.—Quarterly income and expenditures are sums of all inflows and outflows to a customer's accounts, excluding intra-individual transfers. Most of income is driven by paycheck income (~80%), with cash income making up most of the remainder.

In terms of expenditures, cash withdrawals amount for most account outflows and hence are the largest category of expenditure in the data. Measuring consumption is made more difficult by the presence of high cash withdrawals, and we show robustness to the exclusion of such spending. We define consumption by excluding some categories from our measure of spending such as transfers to other individuals that do not mention consumption activities, investment transactions, or insurance payments.

Savings rate.—Quarterly savings rates of individuals are defined as one minus the ratio of consumption to income. Figure 1 depicts the trend of aggregate household savings rates from several data sources, with our bank data represented by the dark blue solid line. As Figure 1 demonstrates, the average household savings rate increases from around 22 percent in 2010 to nearly 32 percent in 2017. We compare the household savings rate in our sample with other sources: NBS, China Family Panel Studies (CFPS), and CHFS. Given the relatively small sample sizes in the surveys, we include in our samples all urban households in the ten Central and Western China provinces that are similar to Inner Mongolia in terms of economic development.

Figure 1 shows that the magnitudes of estimated aggregate household saving rates are quite similar across these data sources. In addition, three of our four measures (except the CHFS) show a similar upward trend during the sample period. The estimates of China's savings rates, ranging from 20 percent to 35 percent, are much higher than those in other major countries. For example, according to the OECD, in 2013 the savings rate was 0.7 percent in Japan, 3.1 percent in the United Kingdom, 5.2 percent in South Korea, and 6.6 percent in the United States.[5]

C. Other Key Variables

To further examine the relation between individual consumption and personal characteristics, we construct additional financial variables. For instance, we follow Carroll (1992) and Choi, Lugauer, and Mark (2017) and calculate individual income volatility as the portion of variance of income over the previous four quarters unexplained by individual trends and characteristics.

We proxy for credit access using information on loan approvals and denials from our bank. We see the set of information observed *Approval* by the bank when making their decisions and can construct probabilities of approval of credit for all individuals in our sample.

We are also able to observe marital status and number of children (and the dates of marriage and births) using linked administrative data. We can proxy the individual's child's marital status with a dummy variable, I($ChildAge_{i,t}$ > 30). Last, we have a dummy variable *Has2ndChild* that equals one if an individual has a second child. This variable describes households' decision to have a second child and will capture any changes in outcome variables after that decision.

D. Summary Statistics

Panel A of Table 1 provides summary statistics on the financial and demographic variables in our sample, aggregated at the individual-quarter level. The average quarterly income is 22,873.18 yuan, the average consumption expenditure is 16,465.60 yuan, and the average household quarterly saving rate is 19.06%. Income Variability differs greatly across individuals

[5] See https://data.oecd.org/hha/household-savings.htm, last accessed April 3, 2021.

and quarters, with an interquartile range of 0.0435 to 0.3167. The average Credit Accessibility measure is 15.52, indicating a 94 percent probability of receiving a loan.[6]

To assess the representativeness of our sample, we compare mean income and consumption in our sample to the information provided by official statistics. Panel B of Table 1 presents average annual income and consumption in our sample and two household surveys in China. Also, we include in the survey sample all urban households in 10 Central and Western China provinces that are similar to Inner Mongolia to maintain a large sample size.

Table 1, panel B, demonstrates that our data are similar to the two surveys but feature two important advantages over the survey data. Our sample is much larger and observed more frequently and also the view of consumption and income are likely more accurate. We have a continuous panel consisting of over 160,000 individual-year observations, whereas the surveys are conducted every two years with fewer than 10,000 observations in each wave.

III. Factors Affecting Chinese Household Savings Rates

In this section, we analyze the determinants of the savings rate in China focusing on factors that the literature shows are important determinants of saving. We focus first on income, which has been shown to be one of the most important factors affecting saving. We then examine the impacts of such other factors as income variability, credit accessibility, marital status, and number of children and childbirth in a unified framework. Importantly, by using the implementation of the so-called two-child policy in 2014 as an exogenous shock and using a difference-in-difference (DID) setting, we demonstrate that the opportunity of having an additional child increases the savings rates of treated individuals, suggesting that the number of children leads to higher savings rates in China.

[6] Using an odds ratio of 15.52, the corresponding probability of receiving a bank loan is calculated as 15.52 / (1 + 15.52) = 94 percent.

A. The Relation between Consumption and Income

It is widely documented that the income elasticity of consumption is lower than one and is decreasing in income, which leads to an increasing savings rate as income grows.[7] We verify this relation using our bank sample in this section. Following Baker (2018), we start by estimating the quarterly income elasticity of consumption with a panel fixed effects model, as shown in the following equation:

$$\Delta Log\left(Consumption_{i,t}\right) = \beta_1 \Delta Log\left(Income_{i,t}\right) + \alpha_i + \gamma_t + \epsilon_{i,t}, \tag{1}$$

where $Income_{i,t}$ and $Consumption_{i,t}$ are quarterly income and consumption of individual i in quarter t as defined in Section II.C. α_i and γ_t are individual fixed effects and calendar-quarter fixed effects, respectively. By definition, β_1 is the average income elasticity of consumption.

To examine the cross-sectional differences in savings rates that are due to income level variation, we further define the rank of average quarterly income, $IncQuintile_i$. Specifically, we sort all individuals evenly into five quintiles based on their average quarterly income throughout the sample period and assign them an integer 1 (lowest) to 5 (highest), accordingly. We then include the interaction term between $\Delta Log(Income_{i,t})$ and $IncQuintile_i$ in equation (1) as an additional explanatory variable. Equation (2) illustrates the specification where the coefficient of the interaction term, β_2, captures the relation between the elasticity of consumption and income quintiles.

$$\begin{aligned} &\Delta Log\left(Consumption_{i,t}\right) \\ &= \beta_1 \Delta Log\left(Income_{i,t}\right) + \beta_2 \Delta Log\left(Income_{i,t}\right) \times IncQuintile_i + \alpha_i + \gamma_t + \epsilon_{i,t} \end{aligned} \tag{2}$$

The first two columns of Table 2 report the regression results from estimating equations (1) and (2), respectively. Column 1 shows that the average income elasticity of consumption is around 0.448, which indicates that a 1 percent increase in income will lead to a 0.448 percent increase in consumption. The magnitude of the elasticity is higher than that in the United States, implying that Chinese households are more sensitive to short-term income fluctuation than US households.[8] Column 2 demonstrates that there is a negative relationship between elasticity and income; a one-

[7] See, for instance, Dynan, Skinner, and Zeldes (2004) and Fagereng, Holm, and Natvik (2021), which demonstrate reductions in marginal propensity to consume (MPC) as income increases across households.

[8] Baker (2018) gives an estimation of 0.295 on the income elasticity of consumption.

quintile increase in average quarterly income is associated with a 0.068 decline in the elasticity. This finding suggests that the marginal saving rate is increasing in income. The overall average saving rate, therefore, is also increasing in income.

Columns 3 and 4 of Table 2 reproduce the results in the first two columns by replacing quarterly income with the corresponding quarterly wage. The estimated coefficients are similar to those found in the first two columns but are smaller in magnitudes as wages account for 81.8 percent of total income.

One concern with any regression of spending on income is that changes in desired or required spending may anticipate or actually induce changes in household labor supply and earned income in general. As such, the coefficients would be biased estimates of the true causal impact of changes in income on spending behavior. To mitigate this concern in the above specification, we use changes in coal prices as instrumental variables to isolate exogenous shocks to income.

We first classify all the 21 industry sectors into three groups according to the relation between their profits and coal prices. Specifically, for each industry, if the overall profit is positively (negatively) correlated with coal prices directly, it is classified as a positively (negatively) correlated industry.[9] If the industry is not generally related to the coal industry directly, we classify it as "coal-neutral" and use it as the benchmark group. For each individual i in quarter t, we then define two dummy variables indicating which industry the individual belongs to. $CoalPos_{i,t}$ ($CoalNeg_{i,t}$) equals one if an individual i's working industry is a positively (negatively) correlated industry. Finally, the instrumental variables for quarterly income are constructed as the interaction term between the industry type dummies and $\Delta Log(CoalPriceIndex)$.

We use a two-stage least squares regression (2SLS) to estimate the income elasticity of consumption with instrumental variables:

$$\Delta Log\left(Consumption_{i,t}\right) = \beta_1\, \Delta Log\left(\widetilde{Income_{i,t}}\right) + \alpha_i + \gamma_t + \epsilon_{i,t}, \tag{3}$$

[9] Specifically, positive-correlated industries (or companies) include coal mining industry, manufacturers of coal-related equipment (such as drillers, trucks, and other coal mining machineries), coal trading companies, railroad and highway transportation industry, natural resources investment companies, and environment technology companies. In contrast, negative-correlated industries (or companies) are those that use coals as inputs, including metallurgical industry (such as steel industry) plants and heating providers.

where the variable $\Delta Log\left(\widetilde{Income_{i,t}}\right)$ represents the fitted value from the first-stage equation:

$$\begin{aligned}\Delta Log\left(Income_{i,t}\right) = \beta_1 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t)\\ +\beta_2 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) + \alpha_i + \gamma_t + \epsilon_{i,t},\end{aligned} \tag{4}$$

where $CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t)$ and $CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t)$ are the instruments for quarterly income, and α_i and γ_t are individual and calendar-quarter fixed effects, respectively. Appendix Table A2 reports results from the first-stage regressions.

For the income quintiles specification, we obtain the 2SLS estimators using the same instruments and their interactions with the income quintiles:

$$\begin{aligned}\Delta Log\left(Consumption_{i,t}\right) = \beta_1\, \Delta Log\left(\widetilde{Income_{i,t}}\right) + \beta_2 \Delta Log\left(\widetilde{Income_{i,t}}\right) \times IncQuintile_i + \alpha_i +\\ \gamma_t + \epsilon_{i,t},\end{aligned} \tag{5}$$

where $\Delta Log\left(\widetilde{Income_{i,t}}\right)$ and $\Delta Log\left(\widetilde{Income_{i,t}}\right) \times IncQuintile_i$ represent the fitted values from the following first-stage equation:

$$\begin{aligned}\Delta Log\left(Income_{i,t}\right) = {}& \beta_1 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t)\\ & +\beta_2 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t)\\ & +\beta_3 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i\\ & +\beta_4 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i + \alpha_i + \gamma_t + \epsilon_{i,t}\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta Log\left(Income_{i,t}\right) \times IncQuintile_i = {}& \beta_1 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t)\\ & +\beta_2 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t)\\ & +\beta_3 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i\\ & +\beta_4 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i\\ & \quad +\alpha_i + \gamma_t + \epsilon_{i,t}.\end{aligned} \tag{7}$$

Columns 5 and 6 in Table 2 present the results from estimating equations (3) and (5), respectively.

We find that the income elasticity of consumption is greater when isolating income fluctuations driven by exogenous coal price changes. This observation indicates that individuals are more sensitive to unanticipated income shocks. Also, as column 6 demonstrates, the cross-sectional variation of income elasticities across different income quintiles becomes stronger. Individuals with high income are likely able to better smooth consumption when faced with unexpected income shocks. Columns 7 and 8 replicate the results in columns 5 and 6 by replacing quarterly income with quarterly wages, finding similar results. F-statistics of the four specifications are all greater than the critical value, mitigating the weak instrument concern.

B. Financial Constraints, Income Growth, Financial Volatility, and Savings

Next, we examine the relation between savings rates and expected or realized financial constraints, especially income variability and credit accessibility. Income variability measures the volatility of transitory labor income and can constrain individuals following low realizations of income (see Jappelli and Pistaferri 2014). According to the precautionary saving theory, the greater income volatility is, the more an individual will save.

Similar to Guiso, Sapienza, and Zingales (2004), we measure credit accessibility using the predicted probability of receiving a loan from our sample bank. Individuals with a higher predicted probability face fewer credit constraints, in expectation. Details about the definitions of the two variables can be found in data appendix. We test the effects of income variability and financial constraints by estimating the following panel data model with demographic control variables,

$$SavingRate_{i,t} = \beta_1\, Log\left(Income_{i,t}\right) + \beta_2 IncomeVariability + \beta_3 CreditAccessibility + \beta_4 \boldsymbol{X}_i + \gamma_t + \epsilon_{i,t}, \tag{8}$$

where subscript i and t represent individual and quarter. The key explanatory variables are the logarithm of quarterly income, income variability, and our measure of credit accessibility. Control variables include a gender dummy, age, age squared, industry dummies, and the interaction terms between age and industry dummies. When we exclude individual fixed effects, we are able to examine these individual-specific characteristics, γ_t.

Table 3, columns 1–4, report the results from estimating regression (8). Column 1 shows the baseline result in which the saving rate is highly positively correlated with income. Generally speaking, a 1 percent increase in quarterly income is associated with a 0.119 percentage point increase in the savings rate. This result is consistent with our findings in Section IV.A.

In addition, savings rates vary across different demographic groups. First, the average savings rate among males is around 5.95 percentage points lower than that among females. Second, there is a strong positive correlation between age and savings rate after controlling for income. Notably, this positive coefficient does not necessarily imply a universally upward-sloping age-saving profile. Because the influence of income level is dominating, the shape of the age-saving profile is determined primarily by the age-income profile. Without controlling for income, we get an inverted U-shaped age-savings profile, consistent with Coeurdacier, Guibaud, and Jin (2015).

In columns 2 and 3, we add income variability and credit accessibility variables and find results consistent with theoretical predictions. The savings rate is higher among individuals with more volatile income. These individuals use saving to buffer against transitory income fluctuations and smooth their consumption. A one standard deviation (0.49) increase in quarterly income variability leads to a 1.4 percentage point increase in the savings rate.

The negative coefficient on credit accessibility suggests that individuals with easier access to bank credit have lower savings rates. Because they can use bank loans as an alternative buffer against unfavorable income shocks, they do not have to save as much of their income. A one standard deviation (7.46) increase in credit accessibility is associated with a decrease in the savings rate of approximately 7 percentage points. The magnitudes and significance of both coefficients remain almost unchanged when both are included together, indicating that the two variables capture different elements of an individual's financial status.

Columns 5–8 replace the personal characteristics controls $\boldsymbol{X}_i$ in columns 1–4 with individual fixed effects α_i, which absorb any time-invariant individual-specified characteristics. The impacts of income and income variability become somewhat stronger, whereas the impact of credit accessibility decreases slightly. Because fixed-effects models can better account for unobservable characteristics, we use them in the remainder of the analysis.

C. Marriage and the Savings Rate

Wei and Zhang (2011) argue that households increase their savings rate in order for their children to have a competitive position in the marriage market. In this section, we test this explanation at the microlevel by using our administrative data about the date of marriage.

As a first glimpse at the administrative data, we show the dynamics of income, consumption, and savings rates around marriage and childbirth. We introduce two sets of dummies indicating the time (in quarters) around the quarter of marriage and childbirth separately:

$$Y_{i,t} = \sum_{\tau=-8}^{32} \beta_\tau \times DMarry_{i,t,\tau} + \sum_{\tau=-8}^{32} \delta_\tau \times DChild_{i,t,\tau} + \mu + \alpha_i + \gamma_t + \epsilon_{i,t}, \tag{9}$$

where subscript i and t represent individual and quarter, respectively. $Y_{i,t}$ is one of the outcome variables—namely, the logarithm of quarterly income, the logarithm of consumption expenditure, or the saving rate. $DMarry_{i,t,\tau}$ is a dummy variable for the timing of marriage. $DMarry_{i,t,\tau}$ equals one if quarter t is τ quarters away from the quarter of marriage for individual i, and zero otherwise. The β_τs capture the changes in $Y_{i,t}$ around marriage. $DChild_{i,t,\tau}$ is a dummy variable for the timing of childbirth. $DChild_{i,t,\tau}$ equals one if quarter t is τ quarters away from the quarter of childbirth for individual i, and zero otherwise. The δ_τs capture the changes in $Y_{i,t}$ around childbirth. α_i and γ_t are individual fixed effects and calendar-quarter fixed effects, respectively.

Figure 2 depicts the average changes in $Log(Income)$, $Log(Consumption)$, and savings rate (that is, the coefficients of the β_τs) in panels A–C, respectively. Panel A shows that the quarterly income reaches a peak in the quarter of marriage and quickly drops to an almost constant level afterward. When we divide the sample by gender, we find that this income pattern is driven mainly by males. From panel B, quarterly consumption expenditures have a similar pattern around the quarter of marriage, but the increase of consumption around marriage is greater than that of income, leading to a sharp decline in the savings rate in the quarter of marriage, as shown in panel C. The savings rate rebounds to its previous level in the following quarter. These findings suggest that marriage is an event that requires large one-time expenditures in China. This fact is the basis of the competitive marriage market savings motive proposed by Wei and Zhang (2011).

Next, we test how marriage affects individuals' saving rate in the fixed-effects model framework by including two dummies that capture marital status. In Table 3, column 9, we add a dummy, *GettingMarried,* in the regression, which equals zero for single individuals and becomes one after they get married. As column 9 shows, the coefficient is negative but is not significantly different from zero, consistent with a temporary effect on savings rates.

In column 10, we further include a proxy for children's marital status, I(*ChildAge*>30). It equals one if any one of the individual's children is older than 30, after which more than half of the children have gotten married. As their children age, parents' savings rates decrease by 2.43 percentage points (t = 1.84) on average. This finding suggests that parents exhibit some competitive saving motives. Consistent with Wei and Zhang (2011), parents in China may save extra money for their children in order to build up greater advantages in the marital market and reduce savings rates after their children marry.

D. Heterogeneity Tests

In the previous sections, we have shown how financial constraints and family status affect savings rates. To examine the heterogeneous effects of these factors among individuals of different demographics, we divide the sample into several groups based on their personal characteristics and separately estimate fixed-effects models for each group.

Table 4 displays the results of several heterogeneity tests. In columns 1–4, we group individuals by age. For groups "40<Age≤50" and "Age≥50," we include all five explanatory variables, as in the last column of Table 3. Columns 5–8 separately examine individuals in different income quartiles. Finally, in columns 9 and 10, we estimate the model for females and males, respectively.

Overall, we find that our results hold broadly across all of these different groups, with some variation in the magnitudes of the results. For instance, older individuals tend to become less sensitive to income fluctuations but the impact of credit accessibility increases in age. Our factors of interest are more influential among individuals of middle-income quartiles than those of extreme groups and the magnitudes of males' coefficients are generally greater. One thing to note is that the dummy of "getting married" is significant at the 10 percent level among females, indicating a mild decline in savings rates after they marry. Another interesting observation is that males significantly reduce their saving rates (by 5.69 percentage points [t = 2.83]) after their children are older than 30.

Finally, we examine the employees of the bank that provided us the data in column 11. These specific bank employees are unlikely to have major bank accounts in other banks, and thus we can have more accurate information on their income and spending. The signs and significance of quarterly income, income variability, and credit accessibility remain unchanged. This finding alleviates some of the concerns about individuals' unobservable savings and spending in other, unobserved bank accounts.

IV. The Relation between the One-Child Policy and the Savings Rate

A number of previous work has pointed to the one child policy as a driver of higher savings rates among Chinese households.[10]. In this section, we analyze the relation between the one-child policy and the savings rate by using a policy change implemented in China in 2014 that eliminated the one-child policy. By using this shift in policy as an exogenous shock to the number of children a household can have, we can test the interaction between family structure and savings rate using a DID design.

A. The Two-Child Policy

After almost thirty years in place, the one-child policy in China began to be relaxed. This adjustment aimed to increase the nation's fertility rate, which had fallen to unprecedentedly low levels. The two-child policy was implemented in three phases, with the first phase starting in 2008. Because of concerns that the relief of population control might trigger a baby boom, only couples who were both the sole children of their parents were permitted to have a second child.

This policy adjustment proved to be too conservative—only about 4 million couples (out of 118 million couples with one child in 2014, or 3.4 percent) were eligible (Zhai, Li, and Chen 2016)—and total fertility rate remained lower than desired by national policy makers. To this end, in early 2014 the government initiated the second phase of the two-child policy, which allowed parents to have a second child if *either* parent was an only child. This policy change expanded the number of eligible couples by around 14 million, or 9.5 percent, of all couples with one child (Zhai, Li, and Chen 2016; Zhang and Wang 2014).

Despite this expansion, worries about low fertility rates grew, and in late 2015 the government decided to abolish the one-child policy. Beginning in January 2016, all couples could have two children. Population experts in China estimate that an additional 91 million couples can benefit

[10] For instance, Choukhmane, Coeurdacier, and Jin (2019) argue that the one-child policy induced parents to save more due to the expected lower support from only one child and that this explains at least 30 percent of China's high savings rate. Zhou (2014) and İmrohoroğlu and Zhao (2018) also point to the one-child policy as driving savings rates.

from this policy, and 17.2 million more children were projected to be born in the five years following the removal of all one-child restrictions (Zhai, Li, and Chen 2016).

Considering the limited influence of the first-phase two-child policy, we use the second phase starting in 2014 as the shock to fertility constraints. The exact policy implementation date varies across provinces. In Inner Mongolia, the provincial government announced the second-phase policy on January 3, 2014, stating that the policy would take effect before mid-2014.[11] The policy was implemented on March 31, 2014, three months after this announcement.[12] We thus regard 2014Q1 as the quarter of policy implementation and define the policy dummy in the DID framework accordingly.

Figure 3 shows the effectiveness of the two-child policy, especially in the last two phases, using national statistics (panel A) and our analysis sample (panel B). In both panels, we calculate the share of second births in total births following Choukhmane, Coeurdacier, and Jin (2019). The dashed red line marks the year of the policy change, 2014. In panel A, we show that the share of second births surged from 30 percent to almost 50 percent after implementation of the policy. Our analysis sample shows a similar pattern in panel B, indicating a significant impact of the policy after 2014.

We also examine the effectiveness of the policy by age since older couples have less willingness and ability to bear more children. Figure 4 presents urban women's second-child fertility rates by ages ranging from 22 to 45 around 2014. For each age group, the fertility rate is calculated as the number of second children borne by women in the corresponding age divided by the number of women of the same age. The dashed red line marks the year of the policy change.

The overall pattern in Figure 4 is similar to Figure 3. More important, the rise in fertility is more significant in women younger than 41. The fertility rates of women ages 42 to 45 are quite low in both the pretreatment and treatment periods. To formalize this observation, we run a series of regressions of fertility rate responses to the policy. Specifically, we estimate the regression:

$$\log\left(Fertility\ Rate_{a,t}\right) = \beta \cdot I(a \leq k) \times I(t \geq 2014) + \gamma_a + \mu_t + \epsilon, \quad (k = 23, 24, \ldots, 44) \tag{10}$$

[11] See http://www.chinanews.com/df/2014/01-03/5696615.shtml (in Chinese).
[12] See http://www.gov.cn/xinwen/2014-04/20/content_2663058.htm (in Chinese).

where subscripts a and t represent age and year, respectively. We use the fertility rate in logs as the dependent variable since the treatment effect is likely to be proportional. For each integer k between 23 and 44 (inclusive), we estimate the response difference between women no older than k and the remaining, which is captured by β. γ_a and μ_t are age fixed effects and time fixed effects, respectively. We then compare the Akaike information criteria (AICs) of these models as displayed in Appendix Figure A.2. The graph shows that the AIC reaches the minimum at $k = 41$, indicating that 41 is indeed the optimal age break to separate the responsive populations.[13]

B. Income, Consumption, and Savings around Childbirth

Paralleling Figure 2, Figure 5 presents the average changes in the three outcome variables. Panel A plots the estimated changes in log income around the quarter of childbirth. In contrast to the upward trend before marriage, the sample's overall income declines about three quarters before the child's birth. This fluctuation comes mainly from females, who probably reduce their labor supply temporarily during pregnancy. In contrast, males' income seems slightly lower within two years after the child's birth but later reverts back.

As panel B demonstrates, individual consumption declines about one year before the birth of the child, especially among females. This trend is mostly a natural response to the lower income during pregnancy. Panel C shows that individuals also react to lower income around childbirth by increasing their savings rate. Once again, these changes are more prominent among females. The impacts of childbirth on males' financial status are mild.

C. The Impacts of the Policy on Savings Rates

As mentioned, we first exclude individuals older than 41 because they are less responsive to the two-child policy. We then assign individuals with no child or one child into the treatment group and those with two or more children into the control group as the latter have already had two children and are not eligible to have another child. The grouping dummy, *ZeroOrOneChild*, equals one if the individuals belong to the treatment group, and zero otherwise.

Since individuals with different numbers of children may act differently even before the policy is implemented, we construct a matched sample in which the treated and the controlled are similar

[13] Evaluating the models using the Bayesian information criterion (BIC) or R-squared generates the same result since the models are linear with a fixed number of explanatory variables.

in observable characteristics in 2013. Specifically, we estimate a logit model of group assignments on financial variables (income, spending, account balance, income variability, and credit accessibility) and demographic variables (age, sex, marital status, minority ethnics, and industry dummies). We then match each treated individual with the nearest neighbor in the control group based on computed propensity score. This helps to control flexibly for larger differences in observable characteristics between groups.[14] Individuals in the control group are weighted by their frequency of matches since they can be matched several times with treated individuals.

We estimate the saving rate response to the two-child policy using a DID specification:

$$\begin{gathered}SavingRate_{i,t} = \beta_1 Policy_t \times ZeroOrOneChild_i + \beta_2 ZeroOrOneChild_i \times Log(CoalPriceIndex_t) + \\ \beta_3 \boldsymbol{X}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t},\end{gathered} \tag{11}$$

where subscript i and t represent individuals and calendar-quarters, respectively. $Policy_t$ is a dummy indicating the implementation of the policy, which equals one if quarter t is in or after 2014. The coefficient of interest is β_1, which captures the savings rate response of eligible individuals after the implementation of the two-child policy.

Although we are using a matched sample, the two groups may still react differently to exogenous income shocks such as coal price fluctuations. To this end, we add the interaction between the group assignment dummy and the logarithm of the quarterly coal price index to capture the heterogeneous responses to other contemporary fluctuations. We also include the explanatory variables $\boldsymbol{X}_{i,t}$ as in the previous models, such as logged quarterly income, income variability, credit accessibility, and marital status. Individual fixed effects, α_i, and calendar-quarter fixed effects, γ_t, are also included to account for unobserved individual-specified characteristics and common trends in savings rates.

We further split the treated population into two groups, one with no children and the other with one child, and estimate the following model:

$$\begin{gathered}SavingRate_{i,t} = \beta_1 Policy_t \times ZeroChild_i + \beta_2 Policy_t \times OneChild_i \\ +\beta_3 ZeroChild_i \times Log(CoalPriceIndex_t)\end{gathered} \tag{12}$$

[14] Using a linear control regression yields qualitatively similar and statistically significant results as compared to the propensity score approach taken here.

$$+\beta_4 OneChild_i \times Log(CoalPriceIndex_t) + \beta_5 \boldsymbol{X}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t},$$

where the indicator *ZeroChild* (*OneChild*) equals one if the individuals have no (one) child, and zero otherwise. This allows us to examine heterogeneous responses within the treatment group.

Table 5 presents the results from the full matched sample, females, and males separately. From the first column, we find that the savings rate of the treatment group (where the individuals have fewer than two children by 2013) increases by 8.17 percentage points (t = 4.23) after the implementation of the two-child policy, compared to the control group (where the individuals have two or more children by 2013). Column 2 refines the treatment group assignment and shows a 1.68 percentage point (t = 1.76) difference between the savings rate responses of the zero-child and one-child groups.

Columns 3–6 estimate the models separately for males and females. We find that the savings rate patterns shown in the first two columns are driven mainly by females, whereas the responses among males are moderate and barely significant. These results echo our finding in the savings rate dynamics that females' financial behaviors are more responsive to the birth of a child.[15]

In the theoretical model of Choukhmane, Coeurdacier, and Jin (2019) one-child policy causes high saving rate by two channels, a transfer channel that expectation of fewer supports in old age leads households to save more, and an expenditure channel that rearing fewer children reduces expenditure and rises saving rate. Our DID results suggest that, instead of considering the additional child as providing insurance and support against the parents' future consumption needs, parents worry more about the costs of raising a second child and therefore save more immediately after the implementation of the policy.

Unfortunately, we cannot observe longer term impacts of this policy where the immediate desire for savings and liquidity may be superseded by the insurance effects of having a child. However, the fact that the savings responses are larger for individuals with no children than those with one child may indicate that this potential insurance explanation is untrue. That is, in the case

[15] To further address potential differences between treatment and control groups, we test whether the parallel trend assumption holds in our specification by replacing the policy dummy with a set of annual dummies spanning the distance between the observation year and the policy year. Appendix Figure A.3 plots the estimated annual coefficients along with the 95 percent confidence levels. The results confirm that the parallel trend assumption is valid since the estimated δ_τs with negative τ do not significantly deviate from zero in all three graphs.

that savings rates were driven upwards by a lack of children to act as insurance for parents, the household with no children would likely already possess higher rates of savings than those with one child and would have to increase savings by less than those that already have one child.

We assert that there are at least two reasons why the two-child policy is more influential for the savings rate of individuals with no child. First, the two-child policy offers them a large decision set (i.e., zero, one, or two children) than that of their one-child counterparts (i.e., only an option of having one more child). Second, one-child individuals are less flexible in adjusting their saving plans because they already have one child that requires them to spend a larger share of their income in the present.

D. Increasing Savings Rates and the Effects on Fertility

To provide further evidence that the rise in the savings rate is due to the two-child policy instead of other shocks that coincided with the policy change, we check the relation between the policy and the savings rate from the other direction. Specifically, we want to show that individuals with increasing savings rates one year after the policy implementation year are more likely to have a new child in the following years. If the savings rate increase is otherwise unrelated to future fertility, then it would be hard to argue that the savings rate increase results from the population policy.

To do so, we use a subsample of individuals having either no child or one child at the end of 2013 and estimate the following cross-sectional model,

$$\begin{aligned} HasNewChild = {} & \beta_1 \Delta SavingRate_{2014} + \beta_2 \Delta SavingRate \times OneChild \\ & + \beta_3 \Delta SavingRate \times I(Age_{2014} \leq 41) \\ & + \beta_5 \Delta SavingRate \times OneChild \times I(Age_{2014} \leq 41) \\ & + OneChild + I(Age2014 \leq 41) + OneChild \times I(Age2014 \leq 41) \\ & + Log(AvgIncome) + MaleDummy \\ & + MinorityPercentage + IndustryDummies + \epsilon, \end{aligned} \tag{13}$$

where subscript *i* that indicates individuals in the subsample is omitted for conciseness. The dependent variable, *HasNewChild*, is a dummy indicating whether the individual has a new child between 2015 and 2017 (inclusive)—one to three years after the two-child policy is implemented. Our sample period ends in 2017, so we are unable to extend the analysis horizon further.

The key explanatory variable in this regression is $\Delta SavingRate_{2014}$ (or $\Delta SavingRate$ for simplicity), which is the difference in individuals' savings rate between the end of 2014 and 2013. It measures the change in the savings rate in the policy year. We thoroughly interact $\Delta SavingRate$ with the *OneChild* dummy and the age group dummy, $I(Age_{2014} \leq 41)$ in order to capture the explanatory power of saving rate changes in different demographic groups. We use the age of 41 as the threshold since it is the upper bound for the policy to be significantly effective. Individuals older than 41 are regarded as the base group. The results are also robust to other different age group specifications.

Other control variables are the average log income in 2013–14, a male dummy, and a set of industry dummies. We also include the percentage of minority ethnic population in the individual's residential city. This is a proxy for whether the individual belongs to a minority group that needs to be controlled since minority populations in China generally face a less restrictive childbirth policy and tend to have higher fertility rates.

Table 6 reports the results from fitting the model in equation (13). In columns 1–3, we estimate the model with logit regressions since they are easy to interpret. The results indicate that a savings rate increase can predict the probability that an individual will have an additional child in the future among those younger than 42 with one child. Specifically, the estimated coefficients of the three-way interaction term, $\Delta SavingRate$ ×OneChild× $\mathrm{I}(\mathrm{Age}_{2014} \leq 41)$, are around 1.2, which remain significant after controlling for income and demographic variables.

Note that the effect of individuals' fertility ends in 2017 due to the limited length of our sample period. To account for the time censoring problem, in columns 4–6 of Table 6, equation (13) is estimated again with Cox regressions in which the timing variable is the time elapsed between the first quarter of 2014 and the quarter of having a new child in the following years. From columns 4–6, we find that Cox regressions improve the significance of coefficients, especially for individuals who have one child, in both age groups.

Overall, these results show that an increased savings rate after the implementation of the two-child policy is associated with a greater tendency to have a new child in the following years, which is the direct effect of the policy.

V. Savings Behavior Among Linked Households

Most literature on Chinese savings rates, including Choukhmane, Coeurdacier, and Jin (2019) and Wei and Zhang (2011), focuses on household rather than individual behavior. While this results in part from the nature of the survey data employed by many such papers, savings rate decisions are often made at a household level. We find similar impacts to our main results when restricting to accounts for which we can link individuals to household units.

A. Household Sample

The household sample derives from two sources of data discussed above in Section II: the transaction data from our sample bank, and the marriage and childbirth data from the administrative system. For the household transaction data, similar practices aimed at reducing the risk of having samples with unobservable income and consumption are taken as in Data Appendix. For marriage data, we choose couples who have no divorce records.

We match individuals' transaction records by marriage registry data. A family record is constructed by combining the husband's and wife's record for this quarter. For some couples, their transaction records happen before the marriage date. In such cases, we link individual records up to one year before marriage, enabling us to check the effect of marriage on saving. Limiting to one year of pre-marriage linkage also reduces the risk of overestimating the intensity of a relationship. After matching, the household sample includes 2,049 families and 11,874 records.

B. Household Results

In order to test the effect of income variability and financial constraints on household saving, we adopt the same framework as in Section III.B.

To make our results more representative, households are weighted by the individual sample weight of the husband, mentioned in Data Appendix. Using the husband's sample weight assumes that the difference in marriage rates across industry and age groups among males is comparable to the difference in industry and age group structure between our sample and the whole population of China. The results are robust against changing the weighting method to using the wife's sample weight or not using sample weights at all.

The regression results are reported in Table 7. We find strong significance for the income, income variability, and credit constraints variables on their effects on saving, consistent with our previous findings in Section III.B. Moreover, the magnitude of response in the household sample is very close to that in the individual sample. However, family saving seems to be more sensitive to changes in household income compared with our results for individuals. Since household income is likely a more accurate gauge of one's financial status than individual income, this may indicate that the individual results are attenuated by measurement error.

VI. The Relative In-Sample Explanatory Power of Financial and Demographic Factors

In this paper, we test several factors that might influence saving rates across individuals in China. In this section, we try to quantify how these factors influence the evolution of savings rates over time in our sample period. Specifically, we examine the variables that we observe covary strongly with savings rates in equation (8): income, income variability, credit accessibility, marriage status, and having children.

Equation (8) identifies how saving rates respond to changes in these five factors at the individual or household level. As a back-of-the-envelope exercise, we examine the drivers of the aggregate time series variation in savings rates during our sample window by linearly combining the changes in average levels of the five factors in the whole population with the coefficients from the regression.[16] This leads to a definition of a factor's contribution to the predicted change in savings rates, which is the product of the factor-level change and the relevant regression coefficient:

$$ctr_i = chg_i * coef_i. \tag{14}$$

Where ctr_i is the contribution of factor i, chg_i is the change of factor level across the sample period, and $coef_i$ is the regression coefficient of factor i. Further, we define the proportion of a factor's contribution to the predicted change in savings rates as its relative explanatory power:

$$rep_i = \frac{ctr_i}{\sum_j ctr_j}. \tag{15}$$

[16] Similar calculation has been used in Wei ang Zhang (2011) to estimate the contribution of marriage market competition in explaining high Chinese saving rate.

Table 8 presents the result of this exercise, contrasting 2010 and 2017, the endpoints of our sample. The results are similar when we include demographic characteristics or individual fixed effects in the regression. Changes in income over time is seen to be the only driver of the increases in savings rates during this period, while income variability and credit accessibility, though significantly different from zero, have somewhat negative effects. Marriage status and having children also have zero or negative impacts on the aggregate changes in savings rate we observe.

Overall, we find that savings rates are responsive to financial attributes such as income volatility and credit accessibility as well as marriage and demographic policy in the cross-section. However, in the time series, these variables cannot explain the increase in savings rates across our sample window. For instance, even though income volatility may induce increases in savings rates for individuals, aggregate income volatility does not increase in tandem with savings rates across all households in our sample. That said, this exercise only considers changes occurring during our sample window (2010-2017) and cannot speak to aggregate changes in financial or demographic changes that occurred during other periods of Chinese history.

Income growth seems to explain much of the aggregate trend in savings rate changes in China during our sample window. We also show some evidence that income growth is a powerful predictor of savings rates across a range of countries. In Table 9, we regress changes in savings rate on changes in income growth across 33 OECD countries. We vary the sample period to match our own and to extend further back in time, seeing consistent global evidence that national income growth covaries strongly with national savings rates.

VII. Conclusion

High levels of household savings are a consistent feature of the Chinese economy. These high levels of savings have both profound implications for domestic growth and investment as well as impacts on the wider global economy, and many explanations, spanning demographic, financial, and political factors, have been proposed.

This paper employs transaction-level financial data across thousands of individuals to provide a unique view of savings decisions in China over eight years. Moreover, we are able to link this transaction data to other financial information regarding credit access as well as to administrative records on such demographic factors as children, marriage, and household formation.

We show that households in China tend to respond similarly to financial shocks when compared to Western households. We then demonstrate that such financial factors as income growth and lower income volatility are primary predictors of high savings rates. Moreover, access to credit tends to depress savings rates among Chinese households, just as is seen among consumers in other countries. While such factors explain cross-sectional heterogeneity in savings rates in our sample, aggregate trends are unable to be explained by aggregate shifts in financial access or volatility. Rather, high levels of income growth likely drove up savings rates, a relationship that we show holds true not only in China but in OECD countries across the world.

A second potential driver of high savings rates concerns demographics or politics. We investigate whether the relaxation of the one-child policy had any substantial effects on savings rates of Chinese citizens, following research that proposed the one-child policy as one factor causing high rates of savings in past decades. Using a difference-in-difference strategy, we find little to support this view, at least in the short run. In fact, the relaxation of the policy tends to significantly *increase* savings rates in the following years.

References

Agarwal Sumit, and Wenlan Qian. 2014. "Consumption and Debt Response to Unanticipated Income Shocks: Evidence from a Natural Experiment in Singapore" *A.E.R.* 104 (12): 4205–4230.

Bai, Chong-En, and Binzhen Wu. 2014. "Health Insurance and Consumption: Evidence from China's New Cooperative Medical Scheme." *J. Comparative Econ.* 42 (2): 450–69.

Baker, Scott R. 2018. "Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data." *J.P.E.* 126 (4): https://doi.org/10.1086/698106.

Baker, Scott R., and Lorenz Kueng. 2021. "Household Financial Transaction Data." Working Paper.

Banerjee, Abhijit, Xin Meng, Tommaso Porzio, and Nancy Qian. 2014. "Aggregate Fertility and Household Savings: A General Equilibrium Analysis Using Micro Data." Working Paper no. 20050, NBER, Cambridge, MA.

Bernanke, Ben S., Carol Bertaut, Laurie Pounder DeMarco, and Steven Kamin. 2011. "International Capital Flows and the Returns to Safe Assets in the United States, 2003–2007." International Finance Discussion Papers, Board of Governors of the Federal Reserve System, no. 1014.

Bussière, Matthieu, Yannick Kalantzis, Romain Lafarguette, and Terry Sicular. 2013. "Understanding Household Savings in China: The Role of the Housing Market and Borrowing Constraints." Working Paper no. 44611, MPRA.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2008. "An Equilibrium Model of 'Global Imbalances' and Low Interest Rates." *A.E.R.* 98 (1): 358–93.

Carroll, Christopher D. 1992. "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence." *Brookings Papers on Economics Activity* 2: 61–159.

Chamon, Marcos, Kai Liu, and Eswar Prasad. 2013. "Income Uncertainty and Household Savings in China." *J. Dev. Econ.* 105 (C): 164–77.

Chamon, Marcos D., and Eswar S. Prasad. 2010. "Why Are Saving Rates of Urban Households in China Rising?" *American Econ. J.: Macroeconomics* 2 (1): 93–130.

Chen, H., W. Qian, and Q. Wen. 2020. "The Impact of the Covid19 Pandemic on Consumption: Learning from High Frequency Transaction Data." Working Paper.

Choi, Horag, Steven Lugauer, and Nelson C. Mark. 2017. Precautionary Saving of Chinese and U.S. Households. *J. Money, Credit and Banking* 49 (4): 635–61.

Choukhmane, Taha, Nicholas Coeurdacier, and Keyu Jin. 2019. "The One-Child Policy and Household Savings." Working Paper.

Coeurdacier, Nicolas, Stéphane Guibaud, and Keyu Jin. 2015. "Credit Constraints and Growth in a Global Economy." *A.E.R.* 105 (9): 2838–81.

Curtis, Chadwick C., Steven Lugauer, and Nelson C. Mark. 2015. "Demographic Patterns and Household Saving in China." *American Econ. J.: Macroeconomics* 7 (2): 58–94.

Dynan, Karen E., Jonathan Skinner and Stephen P. Zeldes 2004. "Do The Rich Save More?" *J.P.E* 112 (2): 397-444.

Du, Qingyuan, and Shang-Jin Wei. 2013. "A Theory of the Competitive Saving Motive." *J. International Economics* 91 (2): 275–89.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik. 2021. "MPC Heterogeneity and Household Balance Sheets." *American Econ. J: Macro* 13 (4): 1-54.

Feng, Jin, Lixin He, and Hiroshi Sato. 2011. "Public Pension and Household Saving: Evidence from Urban China." *J. Comparative Econ.* 39 (4): 470–85.

Ganong, Peter, and Pascal Noel. 2019. "Liquidity vs. Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession." *A.E.R.* 110 (10): 3100–3138.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. "Does Local Financial Development Matter?" Q.J.E. 119(3): 929–969.

He, Hui, Feng Huang, Zheng Liu, and Dongming Zhu. 2018. "Breaking the 'Iron Rice Bowl': Evidence of Precautionary Savings from the Chinese State-Owned Enterprises Reform." *J. Monetary Econ.* 94 (C): 94–113.

İmrohoroğlu, Ayşe, and Kai Zhao. 2018. "The Chinese Saving Rate: Long-Term Care Risks, Family Insurance, and Demographics." *J. Monetary Econ.* 96 (C): 33–52.

Jappelli, Tullio, and Luigi Pistaferri. 2014. "Fiscal Policy and MPC Heterogeneity." *AEJ: Macroeconomics.* 6 (4): 107-136.

Wei, Shang-Jin, and Xiaobo Zhang. 2011. "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China." *J.P.E.* 119 (3): 511–64.

Wen, Yi. 2010. "Saving and Growth under Borrowing Constraints: Explaining the 'High Saving Rate' puzzle." Federal Reserve Bank of St. Louis Working Paper No. 2009-045C.

Zhai, Zhenwu, Long Li, and Jiaju Chen. 2016. "Target Population and Extra Births under the Universal Two-Child Policy." *Population Research* 40 (4): 35–51.

Zhang, Liping, and Guangzhou Wang. 2014. "Target Population of the 'Selective Two-Child Policy' and the Analysis of Related Issues." *Sociological Studies* 1: 25–39. [In Chinese.]

Zhou, Weina. 2014. Brothers, household financial markets and savings rate in China. *J. Development Econ.* 111 (C): 34–47.

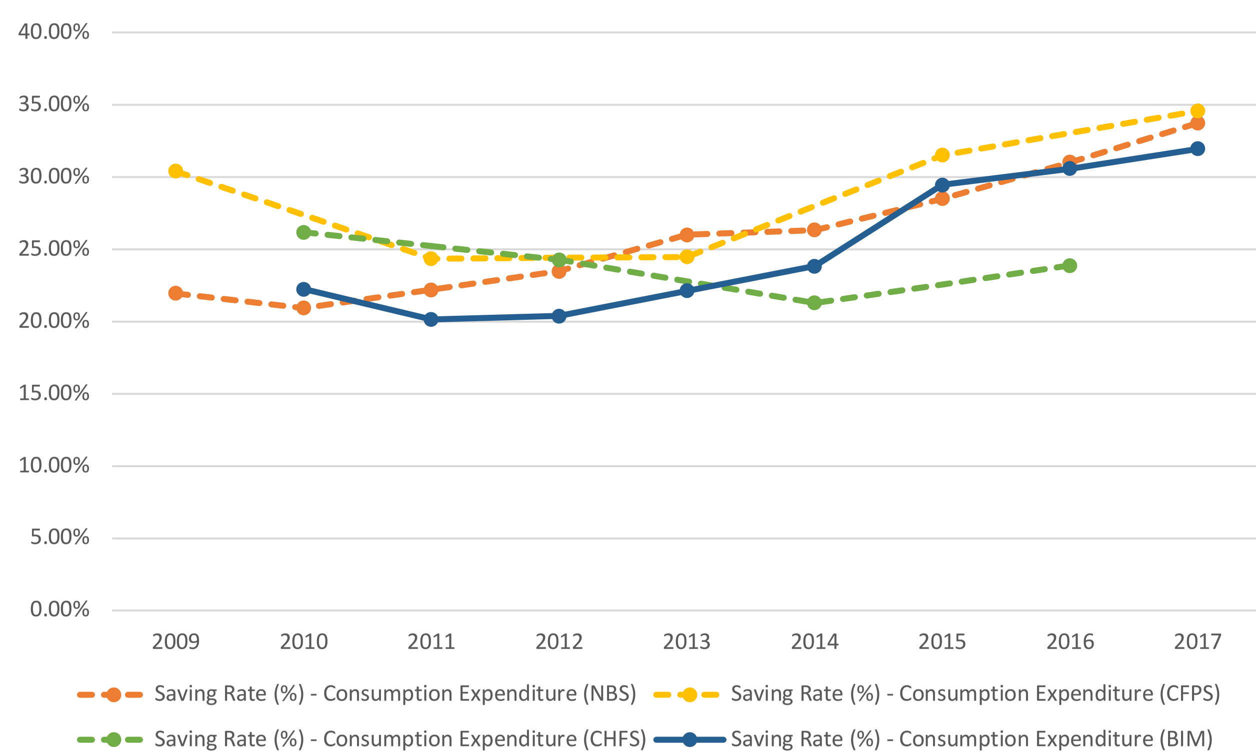
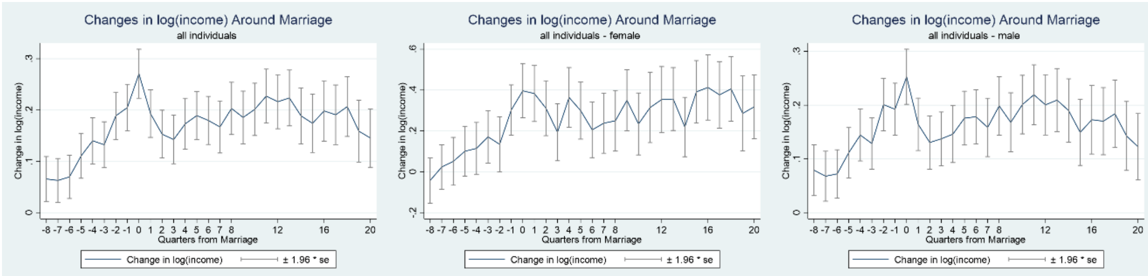
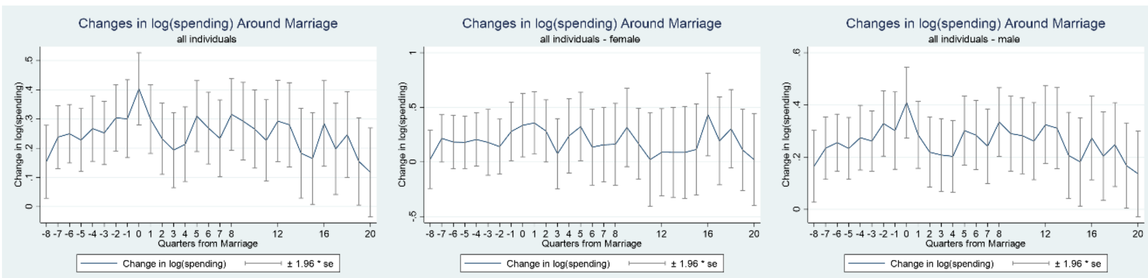


FIG. 1.—Comparison of aggregate savings rates in the bank sample, official statistics, and other surveys. This figure shows aggregate savings rates calculated using data from the bank sample, NBS and several other household surveys in China. The aggregate savings rate is defined as 1 minus the ratio of average consumption to average income. The dark blue solid line plots the aggregate savings rates in our bank sample from 2010 to 2017, while the red, yellow and green dashed lines plot the rates in the NBS, CFPS and CHFS, respectively. NBS savings rates are calculated with official statistics about Inner Mongolia urban households every year. The CFPS and CHFS are conducted every two years. Due to the small sample sizes, urban households in provinces that are similar to Inner Mongolia in geographic location and economic development are used in the calculation.

Panel A. Changes in income around marriage



Panel B. Changes in consumption around marriage



Panel C. Changes in savings rate around marriage

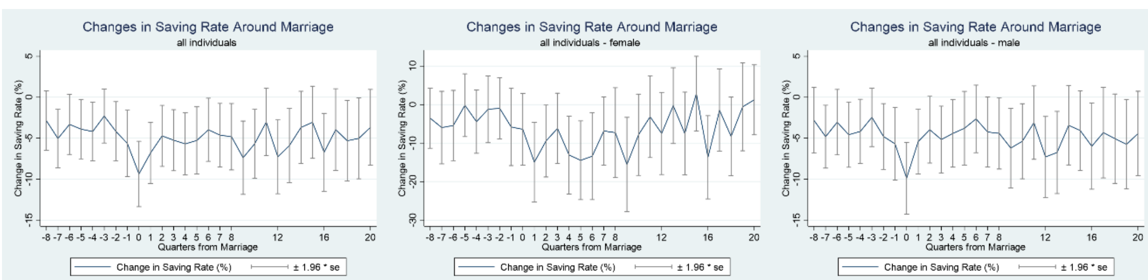
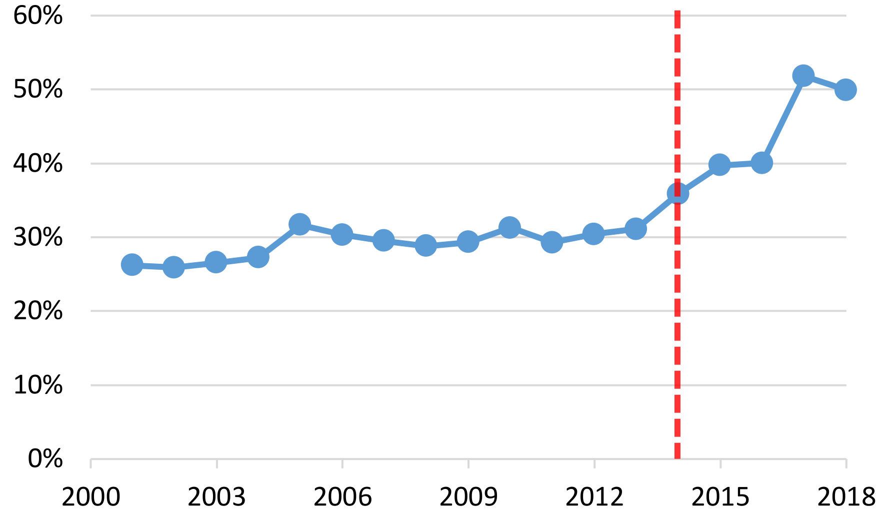


FIG. 2.—Changes of income and consumption around marriage. Changes of income around marriage and childbirth are estimated by the following equation:

$$\log\,(income)_{i,t} = \alpha_i + \gamma_t + \mu + \sum_{\tau=-8}^{32} \beta_\tau \times DMarry_{i,t,\tau} + \sum_{\tau=-8}^{32} \delta_\tau \times DChild_{i,t,\tau} + \varepsilon_{i,t}$$

$DMarry_{i,t,\tau}$ is a dummy variable for marriage. $DMarry_{i,t,\tau}$ equals 1 if for individual i quarter t is τ quarters away from the quarter of marriage, and 0 otherwise. $DChild_{i,t,\tau}$ is a dummy variable for childbirth. $DChild_{i,t,\tau}$ equals 1 if for individual i quarter t is τ quarters away from the quarter of childbirth, and 0 otherwise. α_i and γ_t are individual fixed effects and quarter fixed effects, respectively. Panel A shows the estimations of β_τ for the whole sample, the female subsample, and the male subsample, respectively. Panel B shows the estimations of β_τ on consumption by replacing $\log(income)$ with $\log(consumption)$. Panel C shows the estimations of β_τ on consumption by replacing $\log(income)$ with savings rate. Some control variables in table 3, including income variability and financial constraint proxy are also included in the savings rate regression.

Panel A. Percentage of second child in total births: National statistics



Panel B. Percentage of second child in total births: Bank sample

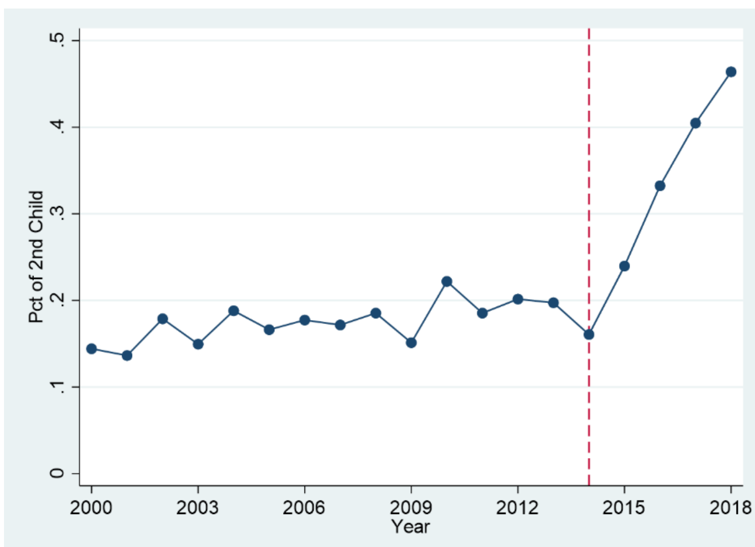


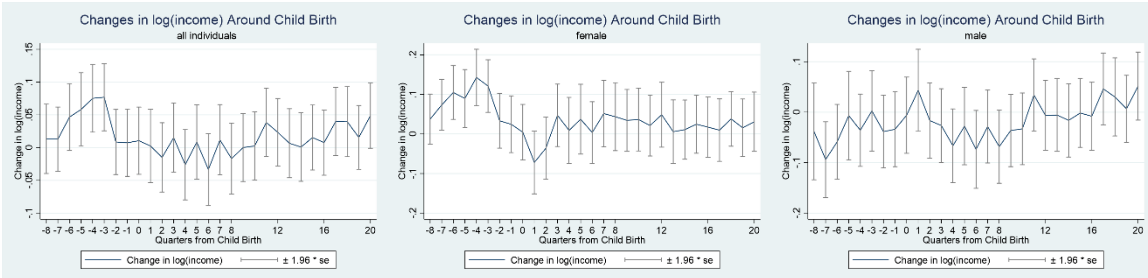
FIG. 3.—The effectiveness of the two-child policy. This figure shows the percentage of the number of family's second child born in the year among all children born in the year. Panel A shows the nationwide statistics from the NBS, and Panel B shows the percentages calculated from our matched Bank sample.

Panel A. Age group 22~33

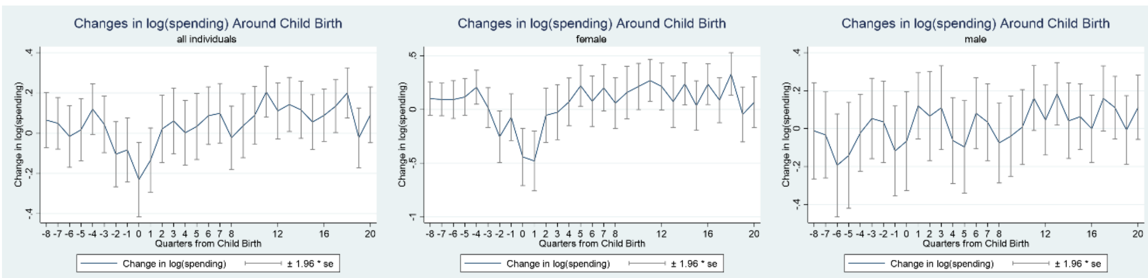
Panel B. Age group 34~45

FIG. 4.—Urban women second-child fertility rates in different age groups around the two-child policy. This figure shows the city women fertility rates in different age groups around the implementation of the two-child policy. For each age group in the year, the fertility rate is calculated as the number of second-child born by women in the age group divided by the population of the women in the same age group. The red dashed line indicates year 2014, the implementation of the two-child policy. Panel A shows the figures for age groups 22 to 33. Panel B shows the figures for age groups 34 to 45. The data are from NBS.

Panel A. Changes in income around childbirth



Panel B. Changes in consumption around childbirth



Panel C. Changes in savings rate around childbirth

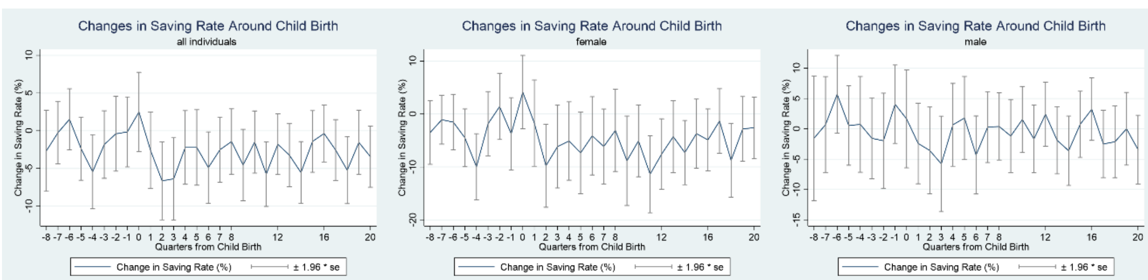


FIG. 5.—Changes of income and consumption around childbirth. Changes of income around marriage and childbirth are estimated by the following equation:

$$\log\,(income)_{i,t} = \alpha_i + \gamma_t + \mu + \sum_{\tau=-8}^{32} \beta_\tau \times DMarry_{i,t,\tau} + \sum_{\tau=-8}^{32} \delta_\tau \times DChild_{i,t,\tau} + \varepsilon_{i,t}$$

$DMarry_{i,t,\tau}$ is a dummy variable for marriage. $DMarry_{i,t,\tau}$ equals 1 if for individual i quarter t is τ quarters away from the quarter of marriage, and 0 otherwise. $DChild_{i,t,\tau}$ is a dummy variable for childbirth. $DChild_{i,t,\tau}$ equals 1 if for individual i quarter t is τ quarters away from the quarter of childbirth, and 0 otherwise. α_i and γ_t are individual fixed effects and quarter fixed effects, respectively. Panel A shows the estimations of δ_τ for the whole sample, for the female subsample, and for the male subsample, respectively. Panel B shows the estimations of δ_τ about consumption by replacing $\log(income)$ with $\log(consumption)$. Panel C shows the estimations of δ_τ about consumption by replacing $\log(income)$ with savings rate. Some control variables in table 3, including income variability and financial constraint proxy are also included in the savings rate regression.

TABLE 1

SUMMARY STATISTICS

Panel A: Quarterly bank sample

	Mean	Standard Deviation	25% Percentile	Median	75% Percentile	Observations
Income (yuan)	22,873.18	47,678.92	7,457.87	11,319.70	19,160.60	571,748
Wage (yuan)	12,351.95	12,137.58	6,771.00	9,738.42	14,408.75	571,748
Spending (yuan)	19,225.37	42,119.93	5,083.00	9,900.00	17,400.00	571,748
Consumption (yuan)	16,465.60	30,907.97	5,000.00	9,500.00	16,500.00	571,748
Savings Rate	19.06%	44.89%	-8.43%	12.84%	53.87%	571,748
Income Variability	0.2776	0.4891	0.0435	0.1167	0.3167	540,136
Credit Accessibility	15.52	7.46	10.17	13.82	19.08	571,748
Age (in 2014)	42.00	12.59	31	42	51	37,100
Male Dummy	0.5651	0.4984				37,100
GettingMarried Dummy	0.2023	0.4017				32,987
I(ChildAge>30) Dummy	0.0249	0.1560				32,987
Has2ndChild Dummy	0.0543	0.2265				32,987
Aggregate Savings Rate	28.01%					

Panel B: Comparison with other survey data

	Mean	Standard Deviation	25% Percentile	Median	75% Percentile	Observations
Annual Bank Sample						
Income (yuan)	81,305.58	139,935.40	27,225.36	44,989.07	77,672.39	160,151
Consumption (yuan)	58,529.01	83,508.63	21,200.00	37,582.00	63,279.58	160,151
Savings rate	13.39%	64.25%	-0.46%	9.80%	36.94%	160,151
Aggregate savings rate	28.01%					
CFPS 2014 (China Family Panel Survey)						
Income (yuan)	71,293.36	58,898.83	40,000	60,000	86,000	1,004
Consumption (yuan)	53,853.75	35,013.36	30,140	44,720	65,068	1,004
Savings rate	13.80%	39.84%	-11.51%	20.02%	42.00%	1,004
Aggregate savings rate	24.46%					
CFPS 2016						
Income (yuan)	97,848.71	171,028.2	45,000	70,000	102,600	887
Consumption (yuan)	67,012.82	57,556.8	33,780	52,552	79,820	887
Savings rate	10.28%	50.98%	-14.27%	20.69%	46.16%	887
Aggregate savings rate	31.51%					
CHFS 2013 (China Household Finance Survey)						
Income (yuan)	62,756.22	76,964.87	25,000	48,000	80,400	4,990
Consumption (yuan)	47,530.31	44,145.10	24,760	38,000	56,560	4,990

Savings rate	13.57%	55.94%	-10.17%	28.19%	53.25%	4,990
Aggregate savings rate	24.26%					
CHFS 2015						
Income (yuan)	67,721.83	100,706.9	24,580	50,087	85,850	7,518
Consumption (yuan)	53,312.45	55,046.25	26,956	40,950	61,140	7,518
Savings rate	10.99%	57.60%	-14.10%	25.97%	51.67%	7,518
Aggregate savings rate	21.28%					
CHFS 2017						
Income (yuan)	74,415.84	78,729.38	30,000	58,800	94,567	7,785
Consumption (yuan)	56,653.05	51,838.83	28,902	45,102	67,938	7,785
Savings rate	12.32%	54.51%	-10.62%	26.51%	50.68%	7,785
Aggregate savings rate	23.87%					

NOTE.—This table shows the summary statistics of our sample from the bank and other widely used household finance surveys in China. Panel A shows the summary statistics of the quarterly bank sample. Income and Spending are calculated according to the account transaction records, excluding transfers between the accounts of the same individual. Wage and Consumption are identified according to the brief transaction descriptions. The savings rate equals 1 minus the ratio of consumption to income. "Income Variability" is calculated using wage according to Carroll (1992) and Choi, Lugauer, and Mark (2017). "Credit Accessibility" is the fitted odds ratio of a logit regression model between loan approvals and personal characteristics estimated with loan application data. "GettingMarried" is a dummy that equals 1 if the individual is married in the quarter, and 0 otherwise. "I(ChildAge>30)" is a dummy and equals 1 if the individual has a child older than 30 in the quarter, and 0 otherwise. "Has2ndChild" is a dummy that equals 1 if the individual has the second child in the quarter, including the time of pregnancy, and 0 otherwise. The aggregate savings rate is 1 minus the ratio of average consumption to average income. See Section I.D and I.E for additional details. For demographics and family status dummies, the latest value of each individual is summarized. Panel B shows the distributions of income, consumption, and savings rate, along with the aggregate savings rate in the bank sample and other commonly used survey data in China, including the CFPS and CHFS. The quarterly Bank sample has been aggregated to annual frequency so that it has the same horizon as the surveys. In the survey data, the sample consists of urban households in Central and Western China provinces that are similar to Inner Mongolia in terms of geographic location and economic development.

TABLE 2

THE INCOME ELASTICITY OF CONSUMPTION AND ITS RELATIONSHIP WITH INCOME QUANTILES

	Dependent Variable: ΔLog(Consumption)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
ΔLog(Income)	0.448***	0.708***			1.603***	0.956***		
	(0.004)	(0.012)			(0.128)	(0.074)		
ΔLog(Income)×IncQuintile		-0.068***				-0.093***		
		(0.003)				(0.016)		
ΔLog(Wage)			0.208***	0.387***			1.527***	0.930***
			(0.008)	(0.020)			(0.137)	(0.077)
ΔLog(Wage)×IncQuintile				-0.049***				-0.123***
				(0.005)				(0.017)
Observations	534,577	534,577	534,577	534,577	534,577	534,577	534,577	534,577
- Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of individuals	37,083	37,083	37,083	37,083	37,083	37,083	37,083	37,083
F-test					97.51	50.61	97.80	66.92

NOTE.—IncQuintile is the income quintile ranging from 1 (minimum) to 5 (maximum). Robust standard errors clustered by individual are reported in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

TABLE 3

THE IMPACT OF INCOME LEVEL, INCOME VARIABILITY, AND FINANCIAL CONSTRAINTS ON THE SAVINGS RATE

	Dependent Variable: Savings Rate									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Income)	0.119***	0.121***	0.141***	0.144***	0.150***	0.156***	0.161***	0.167***	0.167***	0.167***
	(0.00166)	(0.00167)	(0.00187)	(0.00188)	(0.00184)	(0.00185)	(0.00198)	(0.00198)	(0.00198)	(0.00198)
Income volatility		0.0282***		0.0288***		0.0381***		0.0377***	0.0377***	0.0377***
		(0.00115)		(0.00114)		(0.00108)		(0.00108)	(0.00108)	(0.00108)
Credit accessibility			-0.00936***	-0.00964***			-0.00833***	-0.00797***	-0.00797***	-0.00797***
			(0.000463)	(0.000463)			(0.000419)	(0.000418)	(0.000418)	(0.000418)
GettingMarried									-0.000180	-0.000411
									(0.00772)	(0.00772)
I(ChildAge>30)										-0.0243*
										(0.0132)
Male Dummy	-0.0595***	-0.0595***	-0.0701***	-0.0705***						
	(0.00297)	(0.00298)	(0.00301)	(0.00302)						
Age	0.145	0.156	0.621***	0.646***						
($\times 10^{-2}$)	(0.132)	(0.132)	(0.133)	(0.133)						
Age Squared	0.452***	0.443***	-0.132	-0.159						
($\times 10^{-4}$)	(0.150)	(0.151)	(0.152)	(0.152)						
Age×Indust. FE	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Observations	493,818	493,818	493,818	493,818	493,648	493,648	493,648	493,648	493,648	493,648
R-squared	0.048	0.051	0.050	0.053	0.202	0.205	0.203	0.206	0.206	0.206
- Individual FE	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
- Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTE.—“Income Variability”, “Financial Constraints”, “Married”, and “I(ChildAge>30)” are as defined in table 1. Robust standard errors clustered by individual are reported in parentheses. * Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

TABLE 4

THE IMPACT OF INCOME LEVEL, INCOME VARIABILITY, AND FINANCIAL CONSTRAINTS ON SAVINGS RATE: HETEROGENEITY TESTS

	DEPENDENT VARIABLE: SAVINGS RATE										
	Age				Income Quartiles				Gender		Bank Employee
	Age≤30	30<Age≤40	40<Age≤50	Age>50	Q1 (Min)	Q2	Q3	Q4 (Max)	Female	Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log(Income)	0.142***	0.166***	0.194***	0.180***	0.180***	0.182***	0.175***	0.158***	0.163***	0.170***	0.130***
	(0.00336)	(0.00374)	(0.00398)	(0.00471)	(0.00592)	(0.00469)	(0.00413)	(0.00294)	(0.00271)	(0.00287)	(0.00493)
Income variability	0.0385***	0.0431***	0.0487***	0.0246***	0.0303***	0.0404***	0.0426***	0.0380***	0.0345***	0.0406***	0.0126**
	(0.00184)	(0.00211)	(0.00217)	(0.00315)	(0.00211)	(0.00199)	(0.00214)	(0.00239)	(0.00163)	(0.00143)	(0.00519)
Credit accessibility	-0.00707***	-0.00733***	-0.00934***	-0.00997***	-0.00311***	-0.00950***	-0.0142***	-0.00845***	-0.00688***	-0.00923***	-0.00564***
	(0.000699)	(0.000779)	(0.000857)	(0.00137)	(0.00110)	(0.00107)	(0.00112)	(0.000670)	(0.000579)	(0.000610)	(0.00118)
GettingMarried	-0.00550	0.0212	-0.0477	0.0149	0.00350	0.00885	-0.00969	0.00296	-0.0320*	0.00824	0.00351
	(0.00932)	(0.0204)	(0.0438)	(0.0524)	(0.0190)	(0.0132)	(0.0157)	(0.0151)	(0.0191)	(0.00849)	(0.0187)
I(Childage>30)			-0.00711	-0.00896	-0.00189	-0.0558*	-0.0458*	0.0103	0.00228	-0.0569***	
			(0.0315)	(0.0148)	(0.0189)	(0.0289)	(0.0269)	(0.0389)	(0.0171)	(0.0201)	
Observations	120,100	127,392	140,298	104,180	145,983	115,407	115,623	116,635	234,010	259,638	22,561
R-squared	0.257	0.234	0.224	0.205	0.160	0.178	0.198	0.259	0.199	0.210	0.278
- Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
- Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

NOTE.—"Income Variability", "Financial Constraints", "Married", and "I(ChildAge>30)" are as defined in table 1. For the groups "Age ≤30", "30<Age≤40" and "Bank Employee", there is no individual who has a child over 30 in the sample period, so the dummy variable "I(ChildAge>30)" is omitted in these three groups. Robust standard errors clustered by individual are reported in parentheses

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

TABLE 5

THE IMPACT OF THE SECOND-CHILD POLICY ON THE SAVINGS RATE: DID ANALYSIS

	DEPENDENT VARIABLE: SAVINGS RATE					
	All		Female		Male	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Income)	0.131***	0.131***	0.133***	0.133***	0.125***	0.125***
	(0.00992)	(0.00993)	(0.0120)	(0.0120)	(0.0116)	(0.0116)
Income Variability	0.0761**	0.0764**	0.0585	0.0589	0.0865**	0.0867**
	(0.0322)	(0.0322)	(0.0471)	(0.0471)	(0.0380)	(0.0380)
Credit Accessibility	-0.00384**	-0.00385**	-0.00424*	-0.00426*	-0.00401	-0.00402
	(0.00184)	(0.00184)	(0.00245)	(0.00245)	(0.00281)	(0.00281)
GettingMarried	0.0119	0.00994	-0.0261	-0.0276	0.0219	0.0201
	(0.0144)	(0.0146)	(0.0305)	(0.0302)	(0.0175)	(0.0178)
Policy×ZeroOrOneChild	0.0817***		0.0950***		0.0604*	
	(0.0193)		(0.0193)		(0.0322)	
Policy×ZeroChild		0.0863***		0.101***		0.0636*
		(0.0197)		(0.0201)		(0.0326)
Policy×OneChild		0.0695***		0.0796***		0.0514
		(0.0202)		(0.0212)		(0.0332)
ZeroOrOneChild× Log(CoalPriceIndex)	0.188*		0.335***		0.0165	
	(0.113)		(0.103)		(0.157)	
ZeroChild× Log(CoalPriceIndex)		0.182		0.333***		0.00766
		(0.114)		(0.105)		(0.158)
OneChild× Log(CoalPriceIndex)		0.202*		0.341***		0.0391
		(0.115)		(0.106)		(0.161)
Observations	130,865	130,865	67,289	67,289	63,576	63,576
R-squared	0.217	0.217	0.275	0.275	0.192	0.192
- Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
- Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of individuals	6,745	6,745	3,537	3,537	3,208	3,208

NOTE.—$Policy_t$ equals 1 if quarter t is in year 2014 or later, and 0 otherwise. $ZeroChild_i$ equals 1 if individual i had no child in 2014, and 0 otherwise. $OneChild_i$ equals 1 if individual i had only one child in 2014, and 0 otherwise. Other explanatory variables are defined in the same way as before. Robust standard errors are reported in parentheses.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

TABLE 6

THE RELATION BETWEEN SAVINGS RATE INCREASE AND THE TENDENCY OF HAVING A NEW CHILD AFTER THE TWO-CHILD POLICY

	DEPENDENT VARIABLE: DUMMY: HAVING A NEW CHILD BETWEEN 2015 AND 2017					
	Logit Regression			Cox Regression		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta SavingRate_{2014}$	0.102	0.0928	0.0798	0.101	0.0928	0.0799
	(0.0732)	(0.0697)	(0.0730)	(0.0727)	(0.0692)	(0.0724)
$\Delta SavingRate_{2014} \times OneChild$	-1.073**	-1.058**	-1.003**	-1.068**	-1.052**	-0.998**
	(0.473)	(0.462)	(0.482)	(0.464)	(0.453)	(0.472)
$\Delta SavingRate_{2014} \times I(Age2014 \leq 41)$	-0.111	-0.0753	-0.0469	-0.110	-0.0753	-0.0477
	(0.0924)	(0.0916)	(0.0924)	(0.0908)	(0.0901)	(0.0903)
$\Delta SavingRate_{2014} \times OneChild \times I(Age2014 \leq 41)$	1.206**	1.194**	1.199**	1.197**	1.185**	1.190**
	(0.518)	(0.512)	(0.539)	(0.503)	(0.497)	(0.523)
$OneChild$	0.300	0.225	0.154	0.300	0.228	0.160
	(0.782)	(0.781)	(0.784)	(0.781)	(0.781)	(0.783)
$I(Age2014 \leq 41)$	4.005***	4.073***	4.013***	3.980***	4.042***	3.981***
	(0.504)	(0.503)	(0.515)	(0.504)	(0.503)	(0.514)
$OneChild \times I(Age2014 \leq 41)$	0.158	0.144	0.172	0.137	0.124	0.136
	(0.789)	(0.788)	(0.791)	(0.788)	(0.787)	(0.790)
$Log(Avg. Income\ in\ 2013\text{~}2014)$		-0.409***	-0.287***		-0.393***	-0.267***
		(0.0490)	(0.0695)		(0.0473)	(0.0660)
Observations	14,232	14,232	14,232	14,232	14,232	14,232
- Industry Dummies	NO	NO	YES	NO	NO	YES

NOTE.—We use a subsample of individuals having no child or one child at the end of 2013. The dependent variable is a dummy indicating if the individual had a child between 2015 and 2017 (inclusive). The key independent variable, $\Delta SavingRate$, is the change of individual's savings rate between the end of 2013 and 2014. Also included are 'male' dummies and "Minority percentage", the percentage of minority ethnic population in the individual's residential city, is also included. Robust standard errors are reported in parentheses.

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

TABLE 7

THE IMPACT OF INCOME LEVEL, INCOME VARIABILITY, AND FINANCIAL CONSTRAINTS ON THE SAVINGS RATE

VARIABLES	DEPENDENT VARIABLE: SAVINGS RATE					
Log(income)	0.211***	0.216***	0.219***	0.222***	0.222***	0.222***
	(0.0113)	(0.0116)	(0.0118)	(0.0119)	(0.0119)	(0.0119)
Income variability		0.131***		0.134***	0.135***	0.134***
		(0.0461)		(0.0453)	(0.0451)	(0.0451)
Credit accessibility			-0.00479***	-0.00445***	-0.00450***	-0.00451***
			(0.00171)	(0.00168)	(0.00168)	(0.00168)
GettingMarried					0.0214	0.0213
					(0.0291)	(0.0291)
Haschd30						-0.0242
						(0.117)
Observations	10,567	10,225	10,567	10,225	10,225	10,225
R-squared	0.406	0.415	0.407	0.416	0.416	0.416
- Household FE	YES	YES	YES	YES	YES	YES
- Quarter FE	YES	YES	YES	YES	YES	YES
Number of HHs	1580	1558	1580	1558	1558	1558

NOTE.—“Income variability”, “Credit accessibility”, “GettingMarried”, and “HasChild30” are defined as in table 7. Robust standard errors are reported in parentheses.
* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

TABLE 8

RELATIVE EXPLANATORY POWER OF FACTORS

VARIABLES	(1) Savings rate change	(2) Savings rate change
Log(income)	1.296***	1.200***
	(0.0250)	(0.0205)
Income variability	-0.00461***	-0.00678***
	(0.000779)	(0.000591)
Credit accessibility	-0.243***	-0.189***
	(0.0226)	(0.0164)
Getting Married	-0.0294***	0.00867
	(0.00649)	(0.00974)
HasChild30	-0.0191***	-0.0125***
	(0.00392)	(0.00354)
Observations	483,360	483,360
- Individual Characteristics	YES	NO
- Individual FE	NO	YES
- Quarter FE	YES	YES

NOTE.—We decompose the predicted savings rate change between year 2010 and year 2017, the beginning and end year of our sample, into contributions from five explanatory variables based on regression equation (8). The table presents the proportion of each explanatory variable's contribution to the predicted savings rate change, which are their relative explanatory power. The standard error is calculated using nlcom command in Stata. Standard errors in parentheses. * Significant at the 10 percent level. ** Significant at the 5 percent level. *** Significant at the 1 percent level.

TABLE 9

INCOME GROWTH AND SAVINGS GROWTH ACROSS COUNTRIES

	(1) ΔSaving Rate	(2) ΔSaving Rate	(3) ΔSaving Rate	(4) ΔSaving Rate
ΔLog(Income)	0.158***	0.358***	0.265***	0.385***
	(0.0422)	(0.0522)	(0.0598)	(0.0924)
Observations	841	841	263	263
R-squared	0.052	0.361	0.139	0.324
- Country FE	NO	YES	NO	YES
- Year FE	NO	YES	NO	YES
Time Horizon	1970-2017	1970-2017	2010-2017	2010-2017
Num. Countries	33	33	33	33

NOTE.— This tables regresses the year-by-year increase in household savings rate on that in logarithm of household disposable income, for 33 OECD countries. Dependent variable is the year-by-year increase in household savings rate. Independent variable is the year-by-year increase in the logarithm of household disposable income. Fixed effects are included in the regressions as noted. Robust standard errors clustered by country are reported in parentheses.

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level

Data Appendix

A. Household Transaction Data

Bank customers completed more than 142 million checking and savings account transactions during our sample period. For each transaction, we obtain the account number, transaction amount, transaction date, account balance, and, importantly, a short textual description of the transaction. This description allows us to identify the source and purpose of the corresponding transaction. For a paycheck income transaction, the description reveals the employer's name so that we can clearly identify for whom and in which industry the account holder works.[17] For a credit card or debit card transaction, the description includes the merchant's name and category. For a transfer transaction, the description records the account numbers and the names of both the transferor and the transferee.

A short description also lists the payment method—that is, whether the corresponding transaction is settled with a cash payment or as an online transfer. During our sample period, especially in the early years, credit cards and debit cards were not widely used in China. According to the 2014 Chinese Household Finance Survey (CHFS), 96.7 percent of households use cash as their major payment method, and only 6.3 percent of households have credit cards. Similarly, more than 75 percent of account outflows in our data consists of cash withdrawals at automated teller machines (ATMs) or bank counters, especially in the early years. Many customers tend to withdraw cash monthly or weekly to cover immediate expenses. We aggregate the transaction-level data into quarterly level in the following analyses to account for regular cash withdrawals and any other periodic fluctuations within quarters.

The data set also contains detailed demographic information about each customer, including gender, date of birth, city of birth, and current city of residence. As mentioned above, we assign customers into an industry based on their employer. We also augment the data with administrative data about marital status and childbirth. These data include customers' date of marriage, date of

[17] We manually infer the main business of employer companies from their names and information on the Internet, and classify them into 19 industries following the definitions used by National Bureau of Statistics of China. The 19 industries are agriculture, mining, manufacturing, public utility, construction, wholesales, retails, transportation, accommodation and catering, IT, finance, real estate, commercial service, technology, environment, household service, education, health care and medical, and entertaining.

divorce (if any), and gender and date of birth of every child they have. These data allow us to test directly how marriage and fertility affect the dynamics of household savings rates.

B. Sample Construction

Because some customers have more than one account in our sample, we aggregate all transactions in all accounts by category and cancel out transfers between accounts owned by the same customer. We aggregate transactions activity at the quarterly level to alleviate potential seasonality within quarters. We drop the first quarter of new customers and the last quarter of customers closing all accounts to avoid large abnormal deposits or withdrawals around the opening or closing of an account.

As is true for much research employing data from a single large bank or financial institution (e.g., Ganong and Noel 2019; Agarwal and Qian 2014), it is possible that individuals in our sample may have accounts in other banks. In such a case, we would not observe the transactions in those accounts for a given household. To alleviate this concern, we employ two criteria to identify accounts in our sample that are likely primary bank accounts. First, since paycheck income constitutes the greatest source of total personal income in China, we restrict our sample to individuals who receive continuous paycheck income during our sample period.[18] Since these individuals receive most of their income in the accounts in our sample, they are more likely to use the same accounts for their major expenditure needs. This criterion also ensures that we can clearly identify employer information for every individual in our sample. Second, we filter customers based on the number of accounts they hold and the frequency of their transactions. We keep individuals who have more than one checking account or have at least one savings account in our sample. We exclude individuals with only one checking account and no savings account who generally keep a low account balance with infrequent transactions.[19]

---

[18] The amount of paycheck income should also be at a reasonable level. If the annual paycheck income is less than the 10th percentile of the entire sample (roughly 1,500 yuan), it would not be considered the primary paycheck income. The corresponding individuals are then excluded from our analysis.

[19] Specifically, we regard an individual as having a low account balance if the sum of balances in all accounts is below 100 yuan. We then calculate the average time that the individual takes to reach a low balance and the average number of outflow transactions during the time. If the average time is less than 100 days (the third quartile among all individuals) and the average number of outflow transactions is less than three (roughly once every month), or the average time is over 100 days but there is less than one outflow transaction every 50 days, the individual would be considered as inactive and excluded from our analysis sample.

Last, we restrict our analysis sample to individuals having at least eight quarters (two years) of data. Our final sample consists of 571,748 quarterly observations from 37,100 individuals between 2010 and 2017.

Individuals are then weighted based on age, gender, and industry in order to match the provincial population distribution provided by the National Bureau of Statistics (NBS).[20] For consistency with the aggregation approach used by the NBS, we calculate the distribution in two steps. First, we obtain the gender-industry distribution of employees in the province from the Inner Mongolia Statistical Yearbook. This distribution is then multiplied by the nationwide age distribution within each gender-industry group, which is extracted from the Chinese Statistical Yearbook, to obtain the age-gender-industry distribution of the employed population in the province. All observations in an age-gender-industry group are weighted equally such that the sum of weights equals the group density in the NBS statistics. We apply the same adjustment to every year in our sample period. For retired individuals, their "paycheck incomes" are essentially retirement benefits paid by the Bureau of Social Security and do not reveal industry information. We thus regard retired individuals in our sample as a whole and weight them equally such that the sum of weights matches the proportion of retired people in the entire population.[21]

C. Income, Consumption, and Savings Rates

Income and spending.—Quarterly income (expenditure) of a customer is defined as the sum of all inflows (outflows) in all of a customer's checking and saving accounts within a quarter, excluding transfers between accounts of the same individual. The lion's share of individual income comes from regular paycheck payments, which amount to about 81.8 percent of total income, on average. Cash deposits constitute around 13.8 percent of total income. Other income sources include transfers from other accounts (4.1 percent) and miscellaneous income (0.3 percent) such as interest income, subsidies, and tax refunds.

[20] Before such weighting, our sample is quite similar to the distribution of population in the NBS statistics along these dimensions.

[21] Specifically, according to the 2012 Hohhot Statistical Yearbook, the average household size is 2.64, the average number of employed individuals in a household is 1.43, and the average number of retired individuals in a household is 0.51. Therefore, the sum of weights of retired individuals is set to be 35.7 percent (= 0.51 / 1.43) of the sum of weights of employed individuals, and then the weights are evenly assigned to each retired individual every year.

In terms of expenditures, cash withdrawals amount for most account outflows and hence are the largest category of expenditure in the data. Spending through other payment methods, including debit card payments, mobile payments, and online transfers, have been growing over time but remain small in terms of shares of total withdrawals during our sample period.

Consumption.—Measuring consumption using financial transaction data does present a number of pitfalls. Issues of the observation of cash spending are most pressing in our sample. Spending using cash can be observed through cash withdrawals from bank accounts, but the category of outflow for such withdrawals cannot be determined with certainty. For this reason, we present a robustness table wherein we test our primary empirical results using a definition of consumption that excludes cash spending. In addition, spending on goods financed with credit may be conflated with financing charges. Further, some transfers out of accounts may go toward investments (e.g., real estate or real or financial assets) or to intrahousehold transfers.

To be consistent with the definition of consumption in national statistics and most household surveys, we define consumption by excluding a number of categories of outflows from spending. First, transfers to other individuals are excluded if the transaction description left by the transferors during transfers say nothing about consumption activities. Second, debit card and mobile payments on real estate, financial assets, investment goods (e.g., gold, silver, and antiques), and commercial insurance are regarded as investments rather than consumption. Finally, outflows categorized as lawsuit fees and fines are also subtracted from spending.

D. Loan Application and Approval Data

We assess customers' access to credit in empirical tests by exploiting data on loan applications and approvals in the same bank. The data set contains all loan applications, more than 25,000, submitted to the bank during our sample period. The bank provides such retail loans as mortgages, car loans, and smaller-denomination consumer term loans for personal consumption. Mortgages and car loans account for most of the applications before 2014, but the share of consumer term loans increases in the following years. Because the credit card industry is less developed in China, mortgages and consumer term loans are the major sources of household credit during our sample period.

Each loan application contains the applicant's demographics, standardized financial information, and tentative contract terms. The demographic information includes gender, age, ethnic group, marital status, educational level, job position, postal code, and employer's name and industry. The financial information includes the applicant's annual income, annual household income, collateral assets, and housing status—that is, whether the applicant owns a house or an apartment, rents one, or shares one with other family members. The data also include the proposed terms of the loan, including the loan type, amount, term, interest rate, overdue rate, and repayment method, along with type of collateral or guarantee. For each application we know whether the loan was approved or denied. The unconditional mean approval rate in our dataset is 91 percent, which is consistent with CHFS survey data.

E. Other Key Variables

To further examine the relation between individual consumption and personal characteristics, we construct the following demographic and financial variables.

Income variability.—Following Carroll (1992) and Choi, Lugauer, and Mark (2017) we calculate individuals' income variability. Specifically, we first detrend individuals' quarterly paycheck income by dividing it by the sample average across all individuals within the corresponding quarter. We then partial out the variance in the detrended income that can be explained by demographics and life cycle, including sex, age, age squared, and industry dummies. This step is conducted by regressing detrended income on the above variables and obtaining the residuals, which are income adjusted for observable characteristics. We then take the average adjusted income across all observations for each individual. This average is the permanent component of income, and the difference is the transitory income. Finally, we define income variability for each individual in each year as the variance of the four quarterly transitory incomes within the individual-year pair.

Credit accessibility.—We also construct a measure of an individual's likelihood of obtaining mortgages or consumer loans from the bank. Specifically, we calculate the measure as the fitted odds ratio of getting a loan estimated by the following logit regression model (equation (16)) using loan applications data from the bank.

$$Approved = \beta_1 MaleDummy + \beta_2 Age + \beta_3 AgeSquared + \beta_4 \log(AnnualIncome) + \beta_5 BankEmployeeDummy + IndustryDummies + QuarterDummies + \gamma X + \epsilon \tag{16}$$

The subscript *i*, which indicates different loan applications, is omitted in every variable for brevity. The dependent variable *Approved* is an indicator for loan approval. It takes the value of one if the loan application was approved by the bank, and zero otherwise. *X* is a set of control variables that are available only in the loan application data set, and not the transactional one. *X* includes personal characteristics, such as educational level, housing status, and marital status, as well as several loan terms, such as lending amounts, maturity, and collateral information. Appendix Table A.2 reports the regression results for this specification.

After estimating equation (16), we calculate credit accessibility as the fitted odds ratio for every individual-quarter in our transaction data. Specifically, we use equation (17), where subscript i indicates individuals and t indicates calendar quarters. The hatted betas are the corresponding estimated coefficients from equation (16). The greater the value, the higher the probability that the individual will receive the loan conditional on application and can be used as a measure of financial constraints.

$$Credit\ Accessibility_{i,t} = \exp\{\hat{\beta}_1 MaleDummy_i + \hat{\beta}_2 Age_{i,t} + \hat{\beta}_3 AgeSquared_{i,t} + \hat{\beta}_4\, log(YearIncome_{i,t}) + \hat{\beta}_5 BankEmployeeDummy_i + \widehat{IndustryDummies_i} + \widehat{QuarterDummies_t}\} \tag{17}$$

Family status.—We use three variables to describe an individual's family status: the individual's marital status, the marital status of the individual's child or children, and whether the individual has one or two children. First, $GettingMarried_{i,t}$ is a dummy that equals one if individual *i* has been married in quarter *t*, and zero otherwise.

Second, we proxy the individual's child's marital status with a dummy variable, $I(ChildAge_{i,t} > 30)$. It equals one if individual *i* has a child older than 30 in quarter *t*; otherwise, it equals zero.[22] Specifically, we estimate the following Cox hazard model of marital decision using our administrative data,

[22] The threshold of 30 is motivated by the results from the following Cox hazard regression on marriage as shown in figure A.1.

$$\lambda(t) = \lambda_0(t)\exp\{\beta_1 \log(Income) + \beta_2 SavingRate + \beta_3 X\}, \tag{18}$$

where the timing variable t represents age and the control variables include income level, income variability, sex, industry dummies, and calendar-quarter dummies. $\lambda(t)$ is the hazard function of marriage at age t, and $\lambda_0(t)$ is the baseline hazard function of marriage. These are displayed in appendix Figure A.2, where the first panel plots the estimated smoothed baseline hazard function, and the second panel shows the estimated Kaplan-Meier survival function. These figures demonstrate that the peak of marriage probability is around age 27 and that the portion of single individuals decreases quickly to nearly 50 percent by age 30. Therefore, 30 is a reasonable threshold for a proxy of a child's marital status.

Last, we have a dummy variable *Has2ndChild* that equals one if an individual has a second child in a quarter, and zero otherwise. This variable describes households' decision to have a second child and will capture any changes in outcome variables after that decision.

F. Linked Household Variables

We construct four kinds of variables on the household sample where we link individuals within the same household:

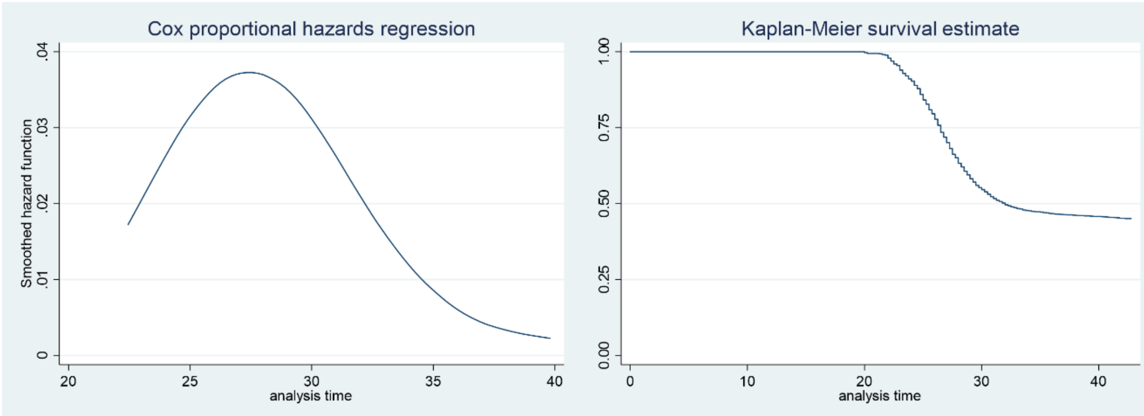
Income, consumption and saving.—The quarterly income of a family is the sum of the husband's and wife's income in that quarter, where the individual income used is defined as in Section II.C. The quarterly paycheck income and consumption of a family are also defined using the measurement in Section II.C. The quarterly savings rate of a family is defined as one minus the ratio of family consumption to income in that quarter.

Income variability.—The construction of family income variability is analogous to that of individual income variability mentioned in Section II.E. The sole difference is that we focus on the family paycheck income instead of individual paycheck income and regress the detrended family paycheck income on the age, age squared, and interaction of age and industry of husband and wife while strictly following the rest of the procedure.

Credit accessibility.—The family credit accessibility is defined here as the sum of the husband's and wife's individual credit accessibility as constructed in Section II.E. Results on savings behavior are robust to different calculations of household credit accessibility, such as including husband's and wife's credit accessibility separately or including only the larger one of the two credit accessibilities.

Family Status.—As in Section II.E, we depict family status using two dummy variables: *GettingMarried* and *HasChild30*. *GettingMarried* describes the marriage status, taking the value one after or in the quarter of marriage event. *HasChild30* takes the value of one if the oldest child, if there is a child in the family, is at least 30 years old.

The summary statistics for this sample are detailed in Table 7. The aggregate savings rate of the household sample is 27.4 percent, moderately lower than the 28 percent in the whole sample. In addition, we find a similar trend of increasing household savings rates across our sample period, mirroring the trend observed among individuals displayed in Figure 1.



(a) (b)

FIG. A. 1.—The timing of marriage. This figure shows the results from a Cox hazard regression about marital decision using the loan application data set. The timing variable is age and the control variables include income level, income variability, sex, industry dummies and calendar-quarter dummies. Panel (a) shows the estimated smoothed baseline hazard function—that is, the marginal probability of getting married at a given age conditional on being single after controlling for the other variables. Panel (b) shows the estimated Kaplan-Meier survival function—that is, the portion of single individuals at a given age.

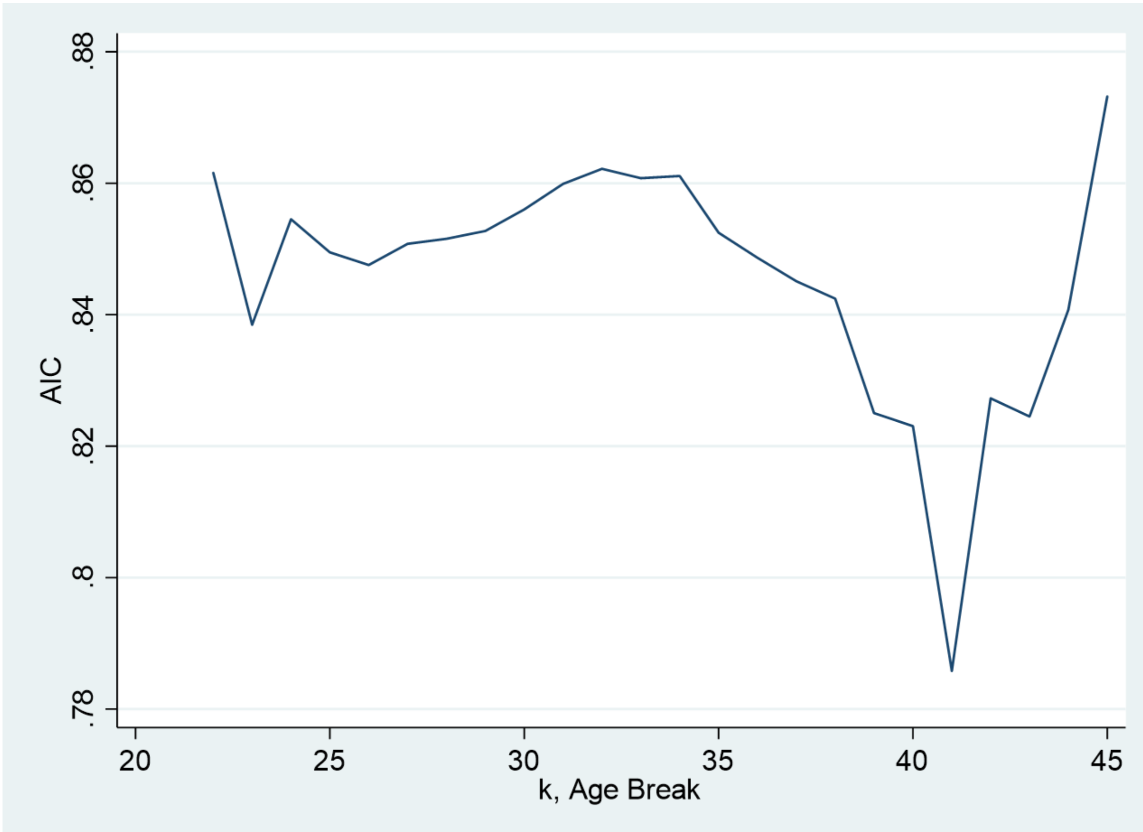
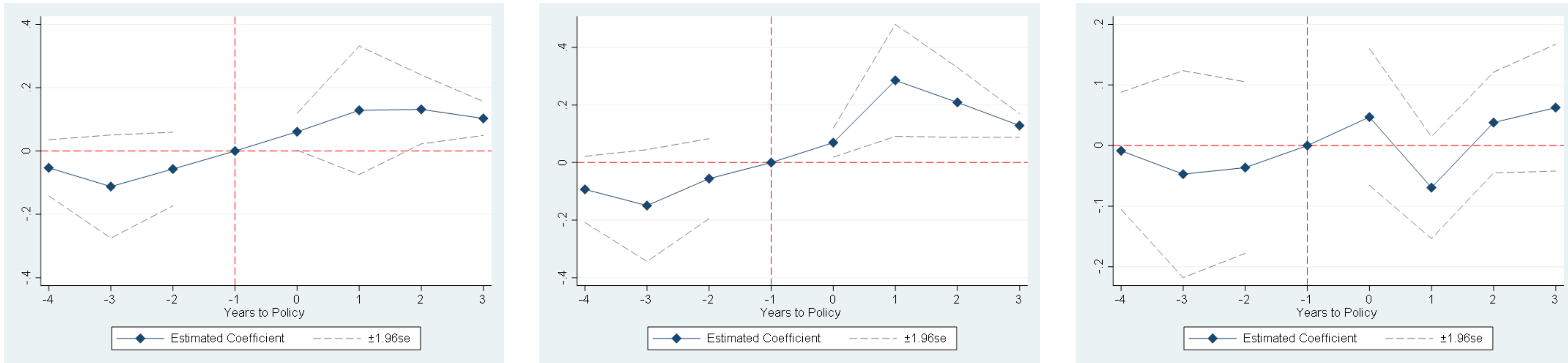


FIG. A. 2.—The performance of the DID models about second-child fertility rate with different age breaks. This figure displays the performance of the DID models about second-child fertility rate, as discussed in Section IV.A. Different age breaks (represented by symbol *k* in the equations) are used in the models and the corresponding AICs are plotted.



(a) (b) (c)

FIG. A. 3.—Tests of the parallel trend assumption in DID Analysis. This figure shows the results from equation (13) which tests the validity of the parallel trend assumption in our DID specifications. The graphs plot the estimated δ_τs along with the 95 percent confidence levels, which capture the savings rate impacts of the two-child policy in the τth year after the implement of the policy. $\tau = -4, -3, -2, -1, 0, 1, 2, 3$ is the distance between observation year and 2014—the policy year. Year -1 is regarded as the base year and the corresponding coefficient, δ_{-1}, is set to 0. Panels (a)—(c) present the results for the whole matched sample, the female matched subsample, and the male matched subsample, respectively.

TABLE A.1

THE INCOME ELASTICITY OF CONSUMPTION: FIRST-STAGE REGRESSIONS

	(1) ΔLog(Income)	(2) ΔLog(Wage)
CoalPos× ΔLog(CoalPriceIndex)	0.363***	0.355***
	(0.068)	(0.063)
CoalNeg× ΔLog(CoalPriceIndex)	-0.614***	-0.661***
	(0.051)	(0.055)
Observations	534,577	534,577
R-squared	0.026	0.029
- Quarter FE	Yes	Yes
- Individual FE	Yes	Yes

NOTE.—This table shows the results of first-stage regressions in table 2, Columns 5~8, where coal prices are used as instrument variables for individual income. The dependent variable is the first-order difference of quarterly income in logarithm, ΔLog(Income). CoalPos and CoalNeg are dummy variables indicating the type of industry in which the individual works in terms of his or her relationship with the coal industry. CoalPos equals 1 if the industry's profits are generally positively correlated with coal prices, and 0 otherwise. CoalNeg equals 1 if the industry's profits are generally negatively correlated with coal prices, and 0 otherwise. ΔLog(CoalPriceIndex) is the log-difference of the China Coal Price Index. Quarter fixed effects and individual fixed effects are also included in the regressions. Robust standard errors clustered by individual are reported in parentheses.

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

TABLE A.2

ESTIMATIONS OF INDIVIDUALS' CREDIT ACCESSIBILITY

	DEPENDENT VARIABLE: LOAN APPLICATION APPROVED DUMMY				
	(1)	(2)	(3)	(4)	(5)
Male Dummy	-0.067	-0.065	-0.074	-0.074*	-0.074
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
Age	2.558*	2.413*	1.886	1.896	2.073
	(1.409)	(1.415)	(1.557)	(1.559)	(1.556)
Age Squared	-3.219*	-3.006*	-2.417	-2.423	-2.607
	(1.683)	(1.691)	(1.817)	(1.820)	(1.815)
Log(Year Income)	0.187**	0.187**	0.187**	0.186**	0.190**
	(0.094)	(0.094)	(0.094)	(0.094)	(0.094)
Missing Year Income Dummy	1.963*	1.966*	1.977*	1.981*	2.017*
	(1.064)	(1.065)	(1.065)	(1.066)	(1.067)
Bank Employee Dummy	0.249***	0.248***	0.247***	0.239***	0.245***
	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)
Personal Characteristics Dummies:					
Education: Undergraduate or Above		0.107	0.109	0.108	0.106
		(0.099)	(0.099)	(0.099)	(0.100)
Education: High School or College		0.105	0.106	0.104	0.102
		(0.093)	(0.093)	(0.093)	(0.093)
Education: Primary School or Illiterate		-0.565***	-0.555***	-0.556***	-0.562***
		(0.177)	(0.177)	(0.177)	(0.177)
Education: Unknown		0.102	0.105	0.108	0.102
		(0.098)	(0.098)	(0.098)	(0.098)
Marriage Status: Single			-0.062	-0.060	-0.087
			(0.074)	(0.074)	(0.076)
Marriage Status: Widow			-0.356	-0.360	-0.359
			(0.276)	(0.276)	(0.275)
Marriage Status: Divorced			-0.155	-0.154	-0.158
			(0.106)	(0.106)	(0.106)
Marriage Status: Unknown			-0.176	-0.179	-0.202
			(0.254)	(0.254)	(0.255)
Minority Ethnic Dummy				0.003	-0.003
				(0.093)	(0.093)
Housing Status: Self-Owned					-0.133
					(0.085)
Housing Status: Shared					0.088
					(0.136)
Housing Status: Rental					-0.308**
					(0.154)

Industry Dummies	Yes	Yes	Yes	Yes	Yes
Loan Characteristics:					
Log(Loan Amount)	-0.276***	-0.280***	-0.280***	-0.281***	-0.283***
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Terms (in months)	0.003***	0.004***	0.004***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Loan Characteristics Dummies	Yes	Yes	Yes	Yes	Yes
Observations	24,614	24,614	24,614	24,614	24,614
- Quarter FE	Yes	Yes	Yes	Yes	Yes
- Branch FE	Yes	Yes	Yes	Yes	Yes

NOTE.—This table shows the results of logit regressions analyzing individuals' financial constraints, that is, their accessibility to bank loans. The dependent variable is an indicator of loan approval. For each loan application, the indicator equals 1 if the loan was approved by the bank, and 0 otherwise. Independent variables are personal and loan characteristics, including sex, age, income, bank employee indicator, educational level, marital status, minority ethnics indicator, housing status, working industry dummies, loan amounts, loan terms, and collateral type dummies. Quarter and bank branch fixed effects are included. The models are fitted using loan application data from the bank during 2010—17. Robust standard errors are reported in parentheses.

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.