Fertility and Savings: The Effect of China's Two-Child Policy on Household Savings

SCOTT R. BAKER, EFRAIM BENMELECH, ZHISHU YANG and JACKY ZHANG[1]

October 2021

ABSTRACT

China's high household saving rate has attracted great academic interest but still remains a puzzle. Potential explanations for China's high saving rate include both demographic, policy, and financial causes. However, lack of reliable micro-level data on household finances make it difficult to assess the relative importance of each explanation. This paper utilizes individual income and spending transactions linked to demographic characteristics and financial information on loan applications and credit availability from a large Chinese bank located in Inner Mongolia. We match a large subset of these customers to administrative records covering marriage and births, giving us a unique window into consumption and saving patterns around important life events. Our results point towards financial instability and credit access, rather than directives like the one-child policy, being the primary causes of higher levels of savings among Chinese households.

[1] Scott R. Baker and Efraim Benmelech are with the Kellogg School of Management and NBER, Dong Huang and Zhishu Yang with Tsinghua University, and Jacky Zhang with Durham University.

1. Introduction

China's high household saving rate, both relative to advanced economies as well as to developing countries, has attracted great academic interest and has prompted a large body of research into potential explanations. Notably, these high savings rates have come amidst substantial industrialization and income growth, as well as major economic and social policy changes, over the past decades. The effects of China's high saving rates are not confined only to its own borders. Given the scale of the Chinese economy and its significant share in the global economy, China's high savings rates have played a major role in the global saving glut affecting global interest rates and asset prices (Bernanke et al. (2011), Caballero, Farhi and Gourinchas (2008)).

There is a wide array of potential explanations for China's high saving rate, spanning demographic, financial, and social causes. However, there is no conclusive evidence as to the relative importance of each of the potential explanations. One of the difficulties in the research on China's saving rates is the lack of reliable micro-level data on household characteristics and finances. This paper attempts to fill the void by bringing detailed new household financial transaction data to bear on this question. While there has been a dramatic increase in the accessibility of (and research using) such transaction data in the U.S. in recent years, few papers have been able to leverage data from Chinese households.[2] Using such data allows for a more granular understanding of the dynamics of household financial behavior and cleaner identification of the causes for fluctuations in spending and savings.

We utilize eight years (2010-2017) of data from a large Chinese bank located primarily in the province of Inner Mongolia. This bank has substantial coverage of the population of the province, 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 and are also able to link additional demographic characteristics and financial information on loan applications and credit availability for these users. We are also able to match a large subset of these customers

[2] See Baker and Kueng (2021) for an overview of research utilizing such household financial data. Chen et al (2020) is one example of a paper that leverages Chinese household financial transaction to investigate responses to the COVID-19 pandemic.

to administrative records covering marriage and births, giving 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 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 have linked the imposition of the one-child policy to higher levels of savings, 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 considers economic and financial motives, we examine whether precautionary saving motives such as income volatility and access to consumer credit can 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)). The volatility of income and lack of social safety nets in China has been pointed to by many as one explanation for high rates of household saving.

Chamon, Liu, and Prasad (2013) argue that the rising income uncertainty and pension reforms account for two-thirds of the increase in China's urban saving rate. In a similar vein, Choi, Lugauer, and Mark (2017) suggest that over 80% of household savings in China arises from the precautionary motive while He et al. (2018) identify that precautionary savings account for about 40 percent of SOE households' wealth accumulation using a reform to Chinese stated-own enterprises (SOE). Several papers also document a negative relationship between social security benefits and household saving rates.[3]

[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 saving rates following a pension reform in 1995-1997 which reduce the target replacement ratio compared to the pre-reform period.

Our data allows us detailed insights into the volatility of household income over time and also gives us a picture of the availability of credit. Across a range of specifications, households facing higher levels of past income fluctuations tend to save much more of their income. Given higher levels of volatility of income in China as a whole, we view these results as supportive micro-evidence for precautionary savings motives playing a non-trivial role in explaining high savings rates in China.

In conjunction with income volatility, a lack of access to credit or other financial buffers has also been highlighted as one potential precautionary motive for high savings rates. As one example, Coeurdacier, Guibaud, and Jin (2015) argue that the interaction between growth differentials and household credit constraints can explain about a third of the divergence in aggregate saving rates across emerging (e.g., China) and developed (e.g., the U.S.) economies. Specifically, the fall in global interest rates due to the integration of global economy and rapid growth of emerging economies induce more borrowing by the young in less credit constrained economies and greater saving of the middle-aged in more constrained ones. Differences in credit constraints across economies leads to a divergence in saving rates. Wen (2010) and Bussière et al. (2013) also incorporate borrowing constraints in their models and find they contribute to high household saving rates.

We follow Guiso, Sapienza, and Zingales (2004) and measure credit accessibility as the probability of receiving a loan from our bank conditional on financial and demographic observables. This predicted access to credit is highly predictive of (lower) savings rates. Given lower access to consumer credit, such results are again consistent with a precautionary savings motive explaining higher savings rates.

A second area of explanations for its high savings rate has focused more on demographic and social factors that may distinguish China from other countries, both developed and emerging. The sharp increase in the Chinese savings rate coincided with the implementation of the one-child policy in the 1980s, which led several researchers to emphasize the role of China's unique demographic structure in its high saving rate. According to Curtis, Lugauer, and Mark (2015), Chine'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 highlighted the intergenerational transfer channel. Choukhmane, Coeurdacier, and Jin (2019) construct a model with intergenerational transfers and human capital accumulation and propose that the one-child policy explains at least 30% of the rise in aggregate saving. Similarly, Zhou (2014) shows that having an additional brother reduces an individual's household savings rate by at least 5 percentage points. Children help households by sharing risks and the costs of supporting parents in the future. In addition, İmrohoroğlu and Zhao (2018) suggest that the combination of the risks faced by the elderly and the deterioration of intergenerational supports due to the one-child policy may account for approximately half of the increase in the saving rate between 1980 and 2010.

However, Banerjee et al. (2014) argue that the negative correlation between fertility and household saving can be substantially offset by general equilibrium forces. Also, Chamon and Prasad (2010) find that demographic shifts are not able to account for the increase in household savings. Instead, they argue that savings are best explained by rising private expenditures on housing, education, and health care, as well financial underdevelopment in China.[4]

Using our high frequency household transaction data, we investigate some of the dynamics of household savings rates surrounding events: marriage and the birth and marriage of children. Events such as births differentially affect the income and spending behavior of men and women, with men tending to have increases in income after the birth of a child, while the income of women declines, albeit temporarily. Marriage also coincides with some changes in income and spending patterns, with income increasing in the quarters leading up to marriage and then stabilizing afterwards. We also see a notable spike in spending among men in the quarter the marriage takes place in.

We formalize this analysis by identifying impacts of the loosening of the One-Child Policy (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 had two children (due to pre-existing allowances for subsets of the population to have more than one child) or who were likely too old to have additional

[4] Apart from family sizes and age structure, Wei and Zhang (2011) argue that the sex ratio imbalances and competition in the marriage market can lead parents to increase savings to improve their 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.

children. Rather than decreasing rates of savings, we find that households for whom the relaxation of the OCP constituted a reduction in constraints on having additional children tended to *increase* savings rates significantly in the post-policy period.

We are also able to utilize our linked administrative records to examine the response of fertility associated with the post-policy increases in savings. Households that increased savings the most in response to the relaxation of the OCP tended to have 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, as households did not need to save for additional child-related expenditures. Overall, these results point towards financial instability and credit access, alongside rapid income growth, being the primary causes of higher levels of savings among Chinese households rather than the one-child policy.

The rest of the paper is organized as follows. In Section 2, we present the data and variables used in the analysis. In Section 3, we study demographic and financial factors that determine saving rates in China. Section 4 evaluates the effect of the one-child policy on Chinese' saving rates. Section 5 demonstrates the robustness of our results to linking individuals into household units. Section 6 concludes.

2. Data and Sample Validation

The main dataset we use comes from a large commercial bank in the Inner Mongolia province and spans the years 2010 to 2017. 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 the largest banks in the province with a large and comprehensive customer base in the province. Specifically, the bank has nearly 1.8 million retail customers with over 3.6 million separate financial accounts at the end of 2017.

2.1 Household Transaction Data

Customers of the bank completed more than 142 million transactions with their checking and saving accounts during our sample period. For each transaction, we have information about the account number, the transaction amount, transaction date, account balances and, importantly, a short textual description of the transaction. This description allows us to identify the source and the 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 is working.[5] 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% of the households used cash as their major payment method, and only 6.3% of the households have credit cards. Similarly, more than 75% of account outflows in our data are cash withdrawal at automatic teller machines (ATMs) or bank counters, especially in the early years. Many customers tend to withdraw cash on a monthly or weekly basis to cover their daily expenses in the following days. 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 dataset 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 each customer 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), as well as gender and date of birth of every child they have.

[5] 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 & catering, IT, finance, real estate, commercial service, technology, environment, household service, education, healthcare & medical, and entertaining.

These data allow us to test directly how marriage and fertility affect the dynamics of household saving rates.

2.2 Sample Construction

Since 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 the transactions activity at a quarterly level to alleviate potential seasonality within quarters. We drop the first quarter of a new customer and the last quarter of a customer closing all his/her accounts to avoid large abnormal deposits or withdrawals around opening or closing of an account.

As is true for much research employing data from a single large bank or financial institution (eg. Ganong and Noel (2019) and Agarwal and Qian (2014)), it is possible that individuals in our sample may have accounts in other banks. In such a case, we would 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. Since paycheck income constitutes the greatest source of total personal income in China, we first restrict our sample to individuals who receive continuous paycheck income during our sample period.[6] 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. Secondly, 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 saving account in our sample. We exclude individuals with only one checking account and no saving account who keep a low account balance most of the time with infrequent transactions.[7]

[6] The amount of paycheck income should also be at a reasonable level. If the annual paycheck income is less than the 10% percentile of the entire sample (roughly 1500 yuan), it would not be considered the primary paycheck income. The corresponding individuals are then excluded from our analysis.

[7] Specifically, we regard an individual as having a low account balance if the sum of balances in all his/her 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 3 (roughly once every month), or

Finally, 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 between 2010 and 2017 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).[8] For consistency with the aggregation approach used by 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 nation-wide 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.[9]

2.3 Income, Consumption and Saving Rates

Income and spending. – Quarterly income (expenditure) of a customer is defined as the sum of all inflows (outflows) in all of the 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% of total income, on average. Cash deposits constitute around 13.8% of total income. Other income sources include transfers

the average time is over 100 days but there are less than one outflow transaction every 50 days, the individual would be considered as inactive and excluded from our analysis sample.

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

[9] Specifically, according to the 2012 Hohhot Statistical Yearbook that 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% (= 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.

from other accounts (4.1%) and miscellaneous income (0.3%) such as interest income, subsidies, and tax refunds.

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. In addition, some transfers out of accounts may go towards investment (eg. real estate or real or financial assets) or to intra-household 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 cards or mobile payments on real estate, financial assets, investment goods (for example, gold, silver, antique, etc.), and commercial insurance are regarded as investments rather than consumption. Finally, outflows categorized as lawsuit fees and fines are also subtracted from spending.

Saving rate. – Quarterly saving rates of individuals are defined as one minus the ratio of consumption to income, while the average annual saving rate of the entire population ("aggregate saving rate") is defined as one minus the ratio of average annual consumption per capita divided by average annual income per capita.

Figure 1 depicts the trend of aggregate household saving rates from several data sources. The dark blue solid line shows the saving rates calculated with our bank sample. As Figure 1 demonstrates, the average household saving rate increased from around 22% in 2010 to nearly 32% in 2017. We compare the household saving rate in our sample with other, publicly available

aggregate saving rates. The orange, yellow and green dashed line in Figure 1 show the aggregate saving rates from NBS, China Family Panel Studies (CFPS), and CHFS, respectively. Data from the NBS are annual while the two surveys, CFPS and CHFS, are conducted every two years. CFPS has completed four waves in even years (2010, 2012, 2014, and 2016) during our sample period and CHFS has completed four waves in odd years (2011, 2013, 2015, and 2017). These surveys contain information from interviews of a random sample of households that are surveyed about their financials in the previous year. In order to compare similar time periods, we plot the saving rates that correspond to the previous year. Given the relatively small sample sizes in the surveys, we include in our samples all urban households in the ten Middle and West China provinces which are similar to Inner Mongolia in terms of both distance from the heavily industrialized coastal regions and economic development.

Figure 1 shows that the magnitudes of estimated aggregate household saving rates are quite similar across these different data sources. This observation alleviates some concerns about measurement errors in consumption in our data because of frequent cash payments. The estimates of saving rates are much higher than those in other major countries, ranging from 20% to 35%. For example according to the OECD, in 2013, the saving rate was 0.7% in Japan, 3.1% in the U.K., 5.2% in South Korea, and 6.6% in the U.S.[10] In addition, three out of the four measures (except the CHFS) show a similar upward trend during the sample period – suggesting that our data is reflective of the overall saving rate in China.

2.4 Loan Application and Approval Data

In some of the empirical tests in the paper we assess customers' access to credit by exploiting data on loan applications and approvals in the same bank. The dataset contains all loan applications submitted to the bank during our sample period, which constitute a sample of more than 25,000 separate applications. The bank provides various types of retail loans, including housing 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 bank seeing an increasing share of consumer term loans in the following years. Since the credit card industry is less

[10] Please see https://data.oecd.org/hha/household-savings.htm, last visit on April 3, 2021.

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 include gender, age, ethnic group, marital status, educational level, job position, zip code, and their employer's name and industry. The financial information include applicants' annual income, annual household income, collateral assets, and their housing status; that is, whether they own a house or an apartment, rent one, or share one with other family members. The data also include the proposed terms of the loan, including the type of loan, the amount of loan, the term of loan, interest rate, overdue rate, repayment method, and the 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%, which is consistent with CHFS survey data.

2.5 Other Key Variables

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

Income variability. – Following Carroll (1992) and Choi et al. (2017) we calculate individual's 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 (1)) using loan applications data from the bank.

$$\begin{aligned} Approved = {} & \beta_1 MaleDummy + \beta_2 Age + \beta_3 AgeSquared + \beta_4 \log(AnnualIncome) \\ & + \beta_5 BankEmployeeDummy + IndustryDummies + QuarterDummies \\ & + \gamma X + \epsilon \end{aligned} \tag{1}$$

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 value 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 dataset but 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. Table A2 in the appendix reports the regression results for this specification.

After estimating equation (1), we calculate credit accessibility as the fitted odds ratio for every individual-quarter in our transaction data. Specifically, we use equation (2) where subscript *i* indicates individuals and *t* indicates calendar quarters. The hatted betas are the corresponding estimated coefficients from equation (1). The greater the value, the higher the probability that the individual will receive the loan conditional on application and can be utilized as a measure of financial constraints.

$$\begin{aligned} & Credit\ Accessibility_{i,t} \\ & = \exp\{\hat{\beta}_1 MaleDummy_i + \hat{\beta}_2 Age_{i,t} + \hat{\beta}_3 AgeSquared_{i,t} \\ & + \hat{\beta}_4\, log\big(YearIncome_{i,t}\big) + \hat{\beta}_5 BankEmployeeDummy_i \\ & + \widehat{IndustryDummies}_i + \widehat{QuarterDummies}_t\} \end{aligned} \tag{2}$$

Family Status. – We use three variables to describe an individual's family status, namely, the marital status of his/her own, the marital status of his/her children, and whether he/she has a second child. First of all, $GettingMarried_{i,t}$ is a dummy which equals one if individual *i* has been married in quarter *t*, and equals zero otherwise.

Second, we also proxy their children'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.[11]

[11] The threshold of 30 is motivated by the results from the following Cox hazard regression on marriage as shown in Figure A1.

Specifically, we estimate the following Cox hazard model of marital decision using our administrative data.

$$\lambda(t) = \lambda_0(t)\exp\{\beta_1 \log(Income) + \beta_2 SavingRate + \beta_3 X\} \quad (3)$$

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 A2, 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 the age of 27, and the portion of single individuals has decreased quickly to nearly 50% by the age of 30. Therefore, 30 is a reasonable threshold for a proxy of children's marital status.

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

2.6 Summary Statistics

Panel A of Table 1 provides summary statistics on the financial and demographic variables in our sample. As we explained in Section 3.3, financial variables are aggregated at the individual-quarter level. The average quarterly income in the entire sample is 22,873.18 yuan and the average consumption expenditure is 16,465.60 yuan, which leads to an aggregated saving rate of 28.01%. In contrast, the average household quarterly saving rate is 19.06% because of the low (or even negative) quarterly saving rates of some individuals. Income Variability differs greatly across individuals and quarters, with an interquartile range of 0.0435 to 0.3167. The average Credit Accessibility measure is 15.52, indicating a 94% probability of receiving a bank loan based on the applicant's observed characteristics in our transactional dataset.[12] The first (third) quartile is 10.17 (19.08), which is equivalent to a probability of 91% (95%). The mean age of individuals in our

[12] Using an odds ratio of 15.52, the corresponding probability of receiving a bank loan is calculated as 15.52 / (1 + 15.52) = 94%. Probabilities in the following sentence are calculated in the same way.

sample in 2014 is 42, and 57% are males. We summarize family status in the table based on the latest value for each individual. 20.23% are married or got married during our sample period, 2.49% have a child older than 30, and 5.42% have a second child.

In order 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 expenses in our sample and two household finance surveys in China. We further aggregate our quarterly data into annual levels so that they are comparable to the survey data. Also, we include in the survey sample all urban households in ten Middle and West China provinces that are similar to Inner Mongolia to maintain a large sample size.

Table 1 Panel B demonstrates that our data is similar to the two surveys in terms of annual income, consumption expenses, and household saving rates. Moreover, our data have two important advantages over the survey data. Our sample size is much larger, our data is more frequent, and the view of consumption and income are likely more accurate. We have a continuous panel consisting of 160,151 individual-year observations, while the surveys are conducted every two years, and in each survey cohort there are fewer than ten thousand at each cross-section.

3. Factors Affecting Chinese Household Saving Rates

In this section, we analyze the determinants of the saving rate in China focusing on factors that have been shown to be important determinants of saving in the literature. We begin our analysis by focusing on income which has been shown to be one of the most important factors affecting saving. We then examine the impacts of several other factors such as including income variability, credit accessibility, marital status, number of children and childbirth in a unified framework. Importantly, by utilizing the implementation of the "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 saving rates of treated individuals, suggesting that the number of children leads to higher saving rates in China.

3.1 The Relationship between Consumption and Income

It has been widely documented that the income elasticity of consumption is lower than one and is decreasing in income, which leads to an increasing saving rate as income grows.[13] We verify this relationship 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:

$$\Delta Log\big(Consumption_{i,t}\big) = \beta_1 \Delta Log\big(Income_{i,t}\big) + \alpha_i + \gamma_t + \epsilon_{i,t} \tag{4}$$

where $Income_{i,t}$ and $Consumption_{i,t}$ are quarterly income and consumption of individual i in quarter t as defined in Section 3.4. α_i and γ_t are individual fixed effects and calendar-quarter fixed effects, respectively. By definition, β_1 in the equation is the average income elasticity of consumption.

To examine the cross-sectional differences in saving 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 (4) as an additional explanatory variable. Equation (5) illustrates the specification where the coefficient of the interaction term, β_2, captures the relationship between the elasticity of consumption and income quintiles.

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

The first two columns of Table 2 report the regression results from estimating equations (4) and (5), respectively. Column 1 shows that the average income elasticity of consumption is around 0.448, which indicates that a 1% increase in income will lead to a 0.448% increase in consumption. The magnitude of the elasticity is higher than that in the U.S, implying that Chinese households are more sensitive to short term income fluctuation than U.S. households.[14] Column 2

[13] See, for instance, Dynan, Skinner and Zeldes (2004) and Fagereng, Holm, and Natvik (2021) which demonstrate reductions in MPC as income increases across households.
[14] Baker (2018) gives an estimation of 0.295 on the income elasticity of consumption.

demonstrates that there is a negative relationship between elasticity and income; a one-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 wage account for 81.8% 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 the 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 relationship 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.[15] 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 the log-difference of the Chinese Coal Price Index, $\Delta Log(CoalPriceIndex)$.

We use a two-stage least squared regression (2SLS) to estimate the income elasticity of consumption with instrumental variables. Specifically, we estimate the following panel fixed-effects model,

$$\Delta Log\left(Consumption_{i,t}\right) = \beta_1 \, \Delta Log\left(\widetilde{Income}_{i,t}\right) + \alpha_i + \gamma_t + \epsilon_{i,t} \qquad (6)$$

[15] 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 which 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 following first-stage equation,

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

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 fixed effects and calendar-quarter fixed effects, respectively. Appendix Table 2 reports the results from the first-stage regressions.

For the specification with income quintiles (equation (5)), we obtain the 2SLS estimators in a similar way using the same instruments and their interactions with the income quintiles. Specifically, we estimate the following panel fixed-effects model,

$$\begin{aligned}&\Delta Log\left(Consumption_{i,t}\right) \\ &\quad = \beta_1\, \Delta Log\left(\widetilde{Income}_{i,t}\right) + \beta_2 \Delta Log\left(\widetilde{Income}_{i,t}\right) \times IncQuintile_i + \alpha_i + \gamma_t + \epsilon_{i,t}\end{aligned} \tag{8}$$

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,

$$\begin{aligned}\Delta Log\left(Income_{i,t}\right) = {} & \beta_1 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t) \\ & + \beta_2 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) \\ & + \beta_3 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i \\ & + \beta_4 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i + \alpha_i + \gamma_t + \epsilon_{i,t}\end{aligned} \tag{9}$$

$$\begin{aligned}\Delta Log\left(Income_{i,t}\right) \times IncQuintile_i = {} & \beta_1 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t) \\ & + \beta_2 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) \\ & + \beta_3 CoalPos_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i \\ & + \beta_4 CoalNeg_{i,t} \times \Delta Log(CoalPriceIndex_t) \times IncQuintile_i \\ & + \alpha_i + \gamma_t + \epsilon_{i,t}\end{aligned} \tag{10}$$

Columns 5 and 6 in Table 2 present the results from estimating equations (6) and (8), respectively.

We find that the income elasticity of consumption becomes greater when using instrumental variables and isolating income fluctuations driven by coal price changes as compared to the first two columns. 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 as well. Individuals with high income are likely able to smooth consumption better faced with unexpected income shocks than those with low income. Columns 7 and 8 replicate the results in columns 5 and 6 by replacing quarterly income with quarterly wages. The results are similar in signs and magnitude. *F*-statistics of the four specifications are all greater than the critical value, mitigating the weak instrument concern.

3.2 Financial Constraints, Financial Volatility, and Savings

Next, we further examine the relationship between saving 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 Section 3.5. We test the effects of income variability and financial constraints by estimating the following panel data model with demographic control variables,

$$SavingRate_{i,t} = \beta_1 \, Log\big(Income_{i,t}\big) + \beta_2 IncomeVariability + \beta_3 CreditAccessibility + \beta_4 \boldsymbol{X}_i + \gamma_t + \epsilon_{i,t} \quad (11)$$

where subscript *i* and *t* represent individual and quarter, respectively. The dependent variable of the model is the quarterly saving rate, $SavingRate_{i,t}$. 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 to 4 report the results from estimating regression (11). Column 1 shows the baseline result in which the saving rate is highly positively correlated with income. Generally speaking, a 1% increase in quarterly income is associated with a 0.119 percentage points increase

in the quarterly saving rate. This result is consistent with our findings and discussions in Section 4.1.

In addition, saving rates also vary across different demographic groups. First, the average saving rate among males is around 5.95 percentage points lower than that among females. Second, there is a strong positive correlation between age and saving rate after controlling for income. One thing to note here is that this positive coefficient does not necessarily imply a universally upward-sloping age-saving profile. Since the influence of income level is dominating, the shape of the age-saving profile is mainly determined by the age-income profile. In fact, without controlling for income, we get an inverted U-shaped age-saving profile, which is consistent with the discussions in Coeurdacier, Guibaud, and Jin (2015).

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

The negative coefficient on credit accessibility suggests that individuals with easier access to bank credit have lower saving rates. Since 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 savings rate of approximately seven percentage point.

Column 4 includes both income variability and credit accessibility in the same regression. The magnitudes and significance of both coefficients remain almost unchanged, indicating that the two variables capture different elements of an individuals' financial status. In addition, the directions of demographic controls are similar to those in the baseline regression after adding the financial status variables.

Next, we replace the personal characteristics controls $\boldsymbol{X}_i$ in columns 1 to 4 with individual fixed-effects α_i which absorb any time-invariant individual-specified characteristics. Columns 5 to 8 show the corresponding results. We find that the coefficients remain significant although their magnitudes change. The impacts of income and income variability become stronger while the impact of credit accessibility slightly decreases. Overall, these results suggest that our findings are

quite robust and are not subject to major unobserved heterogeneity concerns. Since fixed-effects models can better account for unobservable characteristics, we use them in the remainder of the analysis.

3.3 Marriage and Saving Rate

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

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

$$Y_{i,t} = \sum_{\tau=-8}^{32} \beta_\tau \times DMarry_{i,t,\tau} + \sum_{\tau=-8}^{32} \delta_\tau \times DChild_{i,t,\tau} + \mu + \alpha_i + \gamma_t + \epsilon_{i,t} \quad (12)$$

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. $DMarry_{i,t,\tau} = 1$ if quarter *t* is τ quarters away from the quarter of marriage for individual *i*; and zero otherwise. The β_τ's capture the changes in $Y_{i,t}$ around marriage. $DChild_{i,t,\tau}$ is a dummy variable for the timing of childbirth. $DChild_{i,t,\tau} = 1$ if quarter *t* is τ quarters away from the quarter of childbirth for individual *i*; and zero otherwise. The δ_τ's 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 saving rate (that is, the coefficients of the β_τ's) in Panel 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. By dividing the sample by gender, we find that this income pattern is mainly driven 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 saving rate in the quarter of marriage, as shown in Panel C. But the saving rate recovers 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 competitive marriage market saving motive.

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 "*GettingMarried*" dummy 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. This result is consistent with our previous discussions that getting married does not permanently affect individuals' own saving rate levels.

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 already gotten married. More details about this proxy variable are discussed in Section 3.5. As column 10 shows, children's marriage has a weak negative effect on the saving rate. As their children age, parents' saving rates decrease by 2.43 percentage points (t = 1.84) on average. This finding suggests that parents, other than their children, exhibit stronger competitive saving motives. Parents in China save extra money for their children in order to build up greater advantages in the marital market and reduce savings rates after their children get married.

3.4 Heterogeneity Tests

In the previous sections, we have shown how financial constraints and family status affect saving rates. In order 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 estimate fixed-effects models separately, 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. But for group "Age≤30" and "30<Age≤40", dummy "I(*ChildAge*>30)" is omitted as none of the individuals in these two groups has a child older than 30.

As the first four columns of Table 4 demonstrate, the effect of Log(*Income*) and income variability both increase with age up to 50. As individuals get older, their quarterly saving rates are more sensitive to income variability up the age of 50. Beyond the age of 50, individuals in our sample tend to become less sensitive to income fluctuations as they have accumulated substantial

savings for the previous several decades. In addition, the impact of credit accessibility also increases in age, suggesting that credit constraints become more severe among the older population. Lastly, marital status has no significant effect across different age groups.

Columns 5-8 separately examine individuals in different income quartiles; the results for the lowest income quartile are presented in column 5, while the results for the highest income quartile are displayed in column 8. The results show that our factors of interest are more influential among individuals of middle income quartiles than those of extreme groups. Individuals in the first quartile hardly have any money to save and maintain a low saving rate in general, so the impacts of financial and family status, if any, are relatively small. In contrast, the savings of high-income individuals are less responsive as the factors are less binding for them. They have enough savings, good credits, and many other alternative financial facilities against unfavorable fluctuations in income.

In columns 9 and 10, we estimate the model for females and males, respectively. The signs of the estimated coefficients are mostly consistent with the full sample results, but the magnitudes of males' coefficients are generally greater. One thing to note is that the dummy of "getting married" is significant at the 10% level among females, indicating a mild decline in saving rates after they are married. 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 only the employees of the bank that provided us the data. 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 results are reported in column 11 of Table 4. 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. The magnitudes are somewhat smaller than the baseline results in Table 3 column 10, which might be due to specific characteristics of bank employees, such as higher income, better financial literacy, and fewer financial constraints.

4. The Relationship between the "One-Child Policy" and Saving Rate

Choukhmane, Coeurdacier, and Jin (2019) argue that the "One-Child Policy" induce parents to save more due to the expected lower support from only one child, and this can explain at least

30% of high saving rate in China. In this section, we analyze the relationship between "One-Child Policy" and saving rate by utilizing a policy change implemented in China in 2014 which completely removed the "One-Child Policy". Using this shift in policy as an exogenous shock to the number of children a household can have, we are able to test the interaction between family structure and saving rate using a difference-in-differences (DID) design.

4.1 The Two-Child Policy in China

Almost thirty years in place, the one-child policy in China began to be gradually relaxed. This adjustment was aimed at increasing the fertility rate in China which had fallen to unprecedentedly low levels following many years of the one-child policy being in effect. The implementation of the two-child policy consisted of three phases. The first phase started in 2008 when the population policy authority realized that high population growth in China was no longer a concern. Instead, the decreasing fertility rates in recent years could lead to serious socioeconomic problems, including the fading demographic dividends, an aging society, and sexual imbalance. As a result, the government made the first attempt to encourage fertility by initiating the two-child policy. Since the regulator worried that the relief of population control might trigger a boom in the number of newborns, only couples who were both the only children of their parents were allowed to have a second child during this first phase.

However, the first policy adjustment was too conservative and only about 4 million couples (out of 118 million couples with one child in 2014, 3.4%) in China were eligible (Zhai, Li, and Chen, 2016) and as a result total fertility rate remained lower than desired by national policymakers. To this end, the Chinese central government initiated the second phase of the two-child policy in early 2014 and 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% of the couples with one child (Zhai, Li, and Chen, 2016; Zhang and Wang, 2014).

Despite the implementation of this second phase, worries about low fertility rates were accumulating, and the regulator finally decided to remove the one-child policy completely in late 2015. Beginning in January 2016, all couples were allowed to have two children. Population experts in China estimate that an additional 91 million couples can benefit from this policy, and 17.2 million additional 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 exogenous shock to fertility constraints. The exact policy implementation date varies across provinces. In Inner Mongolia, the first news report about the second-phase policy was published on January 3rd, 2014, by the provincial government saying that the policy will come out before mid-2014.[16] The actual policy implementation date was on March 31st, 2014, three months after the announcement.[17] Therefore, we 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 birth 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% to almost 50% after the implementation the policy. Our analysis sample shows a similar pattern in panel B, suggesting a significant impact of the policy after 2014.

We also examine the effectiveness of the policy by age since the elder couples generally have less willingness and ability to have more children. Figure 4 presents urban women's second-child fertility rates in different ages ranging from 22 to 45 around 2014. For each age group, the fertility rate is calculated as the number of second-child born 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, 2014.

The overall pattern in Figure 4 is similar to that in Figure 3. More importantly, the rise in fertility rates seem to be more significant in women younger than 41. The fertility rates of women aged 42 to 45 are quite low in both the pre-treatment and treatment periods.

To formalize this observation, we further run a series of regressions of fertility rate responses to the policy change. Specifically, we estimate the following regression:

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

where subscripts *a* and *t* represent age and year, respectively. We use the fertility rate in logarithm as the dependent variable since the treatment effect is more likely to be proportional rather than additive. For each integer *k* between 23 and 44 (inclusive), we estimate the response difference

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

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 AIC's (Akaike Information Criterion) of these models as displayed in Appendix Figure A2. The graph shows that AIC reaches the minimum at $k = 41$, indicating that 41 is indeed the optimal age break to separate the responsive populations.[18]

4.2 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 mainly comes from females who probably reduce their labor supply temporarily during pregnancy. On the other hand, 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 saving rates. Once again, these changes are more prominent among females. The impacts of childbirth on males' financial status are mild.

4.3 The Impacts of the Policy on Saving Rate

As mentioned above, we first exclude individuals older than 41 as 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 for having an additional child. The grouping dummy, *ZeroOrOneChild*, equals one if the individuals belong to the treatment group; otherwise, it equals zero.

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

[18] Evaluating the models using 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 several financial variables, including income, spending, account balance, income variability, credit accessibility, as well as demographic variables including age, sex, marital status, minority ethnics, and industry dummies. For each individual in the treatment group, we then match him/her with the nearest neighbor in the control group based on computed propensity score. This helps to more flexibly control for some of the larger differences in observable characteristics between groups.[19] Individuals in the control group are weighted by their frequency of matches as they can be matched several times with different individuals in the treatment group.

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

$$\begin{aligned} SavingRate_{i,t} = {} & \beta_1 Policy_t \times ZeroOrOneChild_i + \beta_2 ZeroOrOneChild_i \times Log(CoalPriceIndex_t) \\ & + \beta_3 \boldsymbol{X}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t} \end{aligned} \tag{13}$$

where subscripts 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 saving 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 occurring in the same period. 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 quarterly income in logarithm, income variability, credit accessibility, and the *GettingMarried* dummy. Individual fixed-effects, α_i, and calendar-quarter fixed-effects, γ_t, are also included to account for unobserved individual-specified characteristics and common trend in saving rates, respectively.

In another specification, we further split the treated population into two groups, one with no child and another one with one child, and estimate the following model:

$$\begin{aligned} SavingRate_{i,t} = {} & \beta_1 Policy_t \times ZeroChild_i + \beta_2 Policy_t \times OneChild_i \\ & + \beta_3 ZeroChild_i \times Log(CoalPriceIndex_t) \\ & + \beta_4 OneChild_i \times Log(CoalPriceIndex_t) + \beta_5 \boldsymbol{X}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t} \end{aligned} \tag{14}$$

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

where the indicator *ZeroChild* (*OneChild*) equals one if the individuals have no (one) child; and zero otherwise. This specification 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 saving rate of the treatment group (where the individuals have less 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). This finding suggests that, instead of considering the additional child as providing insurance against their 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.

Column 2 refines the treatment group assignment and shows a 1.68 percentage points (t = 1.76) difference between the saving rate responses of the zero-child and one-child groups. There are at least two reasons for the fact that the two-child policy is more influential for the saving rate of individuals with no child. First, the two-child policy offers them a large decision set (i.e., 0, 1, or 2 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 as they have already had one child that requires them to spend a larger share of their income in the present.

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

To further address the potential differences between the treatment and the control groups, we explicitly test whether the parallel trend assumption holds in our specification by replacing the policy dummy with a set of annual dummies the distance between the observation year and the policy year. Specifically, we estimate the following specification,

$$SavingRate_{i,t} = \sum_{\substack{\tau=-4 \\ \tau\neq-1}}^{3} \delta_\tau I(year_t - 2014 = \tau) \times ZeroOrOneChild_t + \beta_2 ZeroOrOneChild_i \times Log(CoalPriceIndex_t) + \beta_3 \boldsymbol{X}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t} \quad (15)$$

where τ is the distance (in years) between the year of quarter t and 2014 – the policy year. Year 2013 ($\tau = -1$) is regarded as the base year and the corresponding coefficient, δ_{-1}, is set to zero. Other control variables and fixed effects are also included as before.

Appendix Figure A3 plots the estimated δ_τ's along with the 95% confidence levels. The results confirm that the parallel trend assumption is valid in our specification since the estimated δ_τ's with negative τ do not significantly deviate from zero in all the three graphs.

4.4 Increasing Saving Rate and Effects on Fertility

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

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

$$\begin{aligned} HasNewChild = {} & \beta_1 \Delta SavingRate_{2014} + \beta_2 \Delta SavingRate \times OneChild \\ & + \beta_3 \Delta SavingRate \times I(31 \le Age_{2014} \le 41) \\ & + \beta_4 \Delta SavingRate \times I(Age_{2014} \ge 42) \\ & + \beta_5 \Delta SavingRate \times OneChild \times I(31 \le Age_{2014} \le 41) \\ & + \beta_6 \Delta SavingRate \times OneChild \times I(Age_{2014} \ge 42) \\ & + OneChild + Log(AvgIncome) + MaleDummy + MinorityPercentage \\ & + IndustryDummies + \epsilon \end{aligned} \tag{16}$$

where subscript i that indicates individuals in the subsample is omitted for conciseness. The dependent variable, *HasNewChild*, is a dummy indicating if the individual has a new child between 2015 and 2017 (inclusive) – one to three years after implementing the two-child policy. 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 individual's saving rate between the end of 2014 and 2013.

It measures the change in saving rate in the policy year. We thoroughly interact $\Delta SavingRate$ with the *OneChild* dummy and two age group dummies, $I(Age_{2014} \geq 42)$ and $I(Age_{2014} \geq 42)$, in order to capture the explanatory power of saving rate changes in different demographic groups. We use the age of 41 as the threshold as it is the upper bound for the policy to be significantly effective, and then we focus on a 10-year age group ranging from 31 to 41. Individuals younger than 31 are regarded as the base group. This grouping strategy also ensures the three categories are of similar sizes. The results are also robust to other different age group specifications.

Other control variables are the average log income in 2013-2014, a male dummy, and a set of industry dummies. We also include the percentage of minority ethnic population in the individual's residential city. It is a proxy for the individual being minority ethnic that needs to be controlled since minorities in China generally face a less restrictive population policy and tend to have higher fertility rates.

Table 6 reports the results from fitting the model in equation (16). In columns 1 to 3, we estimate the model with logit regressions as they are easy to interpret. The results indicate that a saving rate increase can predict the probability of having an additional child in the future among the individuals aged between 31 and 41 with one child. Specifically, the estimated coefficients of the three-way interaction term, $\Delta SavingRate \times$OneChild$\times$ I($31 \leq Age_{2014} \leq 41$), are around 1.5, which correspond to a 0.34 basis points marginal increase in the probability of fertility on average if the saving rate increases by one percentage point. The three-way interaction for the age of 42 or older is also positive with a marginal effect of 1.9 basis points at the mean after a one percentage point increase in saving rate, but are not significant in the logit regressions.

Note that the effect of individuals' fertility ends in 2017 due to the limited length of our sample period. In order to account for the time censoring problem, in columns 4 to 6 of Table 6, equation (16) is estimated again with Cox regressions where the timing variable is the time that elapsed between the first quarter of 2014 and the quarter of having a new child in the following years. From columns 4 to 6, we find that Cox regressions improve the significance of coefficients, especially for the individuals having one child, in both age groups.

Overall, these results confirm that increased saving 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.

5. Saving Behavior Among Linked Households

Most literature on Chinese saving rates, including Choukhmane, Coeurdacier, and Jin (2019) and Wei and Zhang (2011), focuses on household rather than individual behavior. While this is partially due to the nature of the survey data employed by many such papers, savings rate decisions are often made at a household level. This paper has focused on how savings react to income, income volatility and credit constraints at an individual level, but we find similar impacts when restricting to accounts for which we can link individuals to household units.

5.1 Household Sample

The household sample is derived from two sources of data discussed above in Section 2: the transaction data from our sample bank, and the marriage and children birth data from administrative system. For the household transaction data, similar practices aiming at reducing the risk of having samples with un-observable income and consumption are taken as in Section 2.2. However, when linking individuals to household units, we relax the requirement that every individual must have at least eight quarters of transaction records. 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 cases like such, we link individual records up to one year prior to marriage, enabling us to check the effect of marriage on saving. Limiting to one year of pre-marriage linkage also reduces the risk of over-estimating the intensity of relationship. After matching, the household sample includes 2,049 families and 11,874 records.

5.2 Variables

We construct four kinds of variables on the household sample:

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 in the same way as in Section 2.3. The quarterly paycheck income and consumption of a family are

defined in a similar way, also using the measurement in Section 2.3. The quarterly saving 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 2.5. The only 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 the 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 constructed in Section 2.5. 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 2.5, 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 any child in the family, is at least thirty years old.

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

5.3 Results

In order to test the effect of income variability and financial constraints on household saving, we adopt the same framework as in Section 3.2. The regression design is no different from equation (11) except the subscript now stands for household rather than individual.

To make our results more representative, households are weighted by the individual sample weight of husband which is mentioned in Section 2.2. To choose the husband's sample weight as mirroring the representativeness of the family is based on the assumption 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 individual sample weights of wives, or not using sample weights at all.

The regression results are reported in Table 8. We find strong significance for the income, income variability and credit constraints variable on their effects on saving, consistent with our previous findings in Section 3.2. Moreover, the magnitude of response in household sample is very close to that in individual sample, though they do exhibit some differences. Specifically, 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 results utilizing individual income and savings rates were somewhat attenuated by measurement error.

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

In this paper, we have tested several factors that might influence saving rates across individuals in China. In this section, we try to quantify how these factors have influenced the evolution of savings rates over time in our sample period. Namely, we choose to examine the variables which have we have observed co-varying strongly with savings rates in Equation (11): income, income variability, credit accessibility, marriage status and having children.

Equation (11) identifies how saving rates respond to changes in these five factors at an 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. 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:

$$ctr_i = chg_i * coef_i \tag{17}$$

Further, we define the proportion of a factor's contribution to the predicted change in savings rates as its relative explanatory power:

$$rep_i = \frac{ctr_i}{\sum_j ctr_j} \tag{18}$$

Table 9 gives the result of this exercise, contrasting 2010 and 2017, the end-points of our sample. The results are similar when also including individual 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 savings rates are responsive to financial attributes like income volatility and credit accessibility and also marriage and demographic policy in the cross section. However, in the time series, these variables cannot explain the increase in savings rate across our sample window. For instance, while income volatility may induce increases in savings rates for an individual, aggregate income volatility did not increase in tandem with savings rates across all households.

7. Conclusion

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

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 detailed transaction data to other financial information regarding credit access as well as administrative records on demographic factors such as children, marriage, and household formation. We first show that households in China tend to respond similarly to financial shocks when compared to Western households. We then demonstrate that financial factors such 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.

A second potential driver of high savings rates are demographic or political factors. 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 differences in differences strategy, we find little to support this view, at least in the short, run. In fact, the relaxation of the policy tended to significantly *increase* savings rates in the following years.

References

Bai, Chong-En, and Binzhen Wu, 2014, Health insurance and consumption: Evidence from china's new cooperative medical scheme, *Journal of Comparative Economics* 42, 450-469.

Baker, Scott R. 2018. Debt and the response to household income shocks: validation and application of linked financial account data. *Journal of Political Economy*

Baker, Scott R and Kueng, Lorenz. 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, *NBER working paper* No. 20050.

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, February 2011

Bussière, Matthieu, Yannick Kalantzis, Romain Lafarguette, and Terry Sicular, 2013, Understanding household savings in China: the role of the housing market and borrowing constraints, *MPRA working paper* No. 44611.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas, 2008, "An equilibrium model of 'global imbalances' and low interest rates," *American Economic Review*, Vol. 98, No. 1, pp. 358–393.

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, *Journal of Development Economics* 105, 164-177.

Chamon, Marcos D., and Eswar S. Prasad, 2010, Why are saving rates of urban households in china rising?, *American Economic Journal: Macroeconomics* 2, 93-130.

Chen H, Qian W, Wen Q. 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, *Journal of Money, Credit and Banking* 49, 635-661.

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, *American Economic Review* 105, 2838-2881.

Curtis, Chadwick C., Steven Lugauer, and Nelson C. Mark, 2015, Demographic patterns and household saving in China, *American Economic Journal: Macroeconomics* 7, 58-94.

Du, Qingyuan, and Shang-Jin Wei, 2013, A theory of the competitive saving motive, *Journal of International Economics* 91, 275-289.

Feng, Jin, Lixin He, and Hiroshi Sato, 2011, Public pension and household saving: Evidence from urban China, *Journal of Comparative Economics* 39, 470-485.

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, *Journal of Monetary Economics* 94, 94-113.

İmrohoroğlu, Ayşe, and Kai Zhao, 2018, The Chinese saving rate: Long-term care risks, family insurance, and demographics, *Journal of Monetary Economics* 96, 33-52.

Wei, Shang-Jin, and Xiaobo Zhang, 2011, The competitive saving motive: Evidence from rising sex ratios and savings rates in China, *Journal of Political Economy* 119, 511-564.

Wen, Yi, 2010, Saving and growth under borrowing constraints: Explaining the "high saving rate" puzzle, *Fed working paper*.

Zhai, Zhenwu, Long Li, and Jiaju Chen, 2016, Target population and extra births under the universal two-child policy, *Population Research* 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, *Journal of Development Economics* 111, 34-47.

Figure 1. Comparison of Aggregate Saving Rates in The Bank Sample, Official Statistics and Other Surveys

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

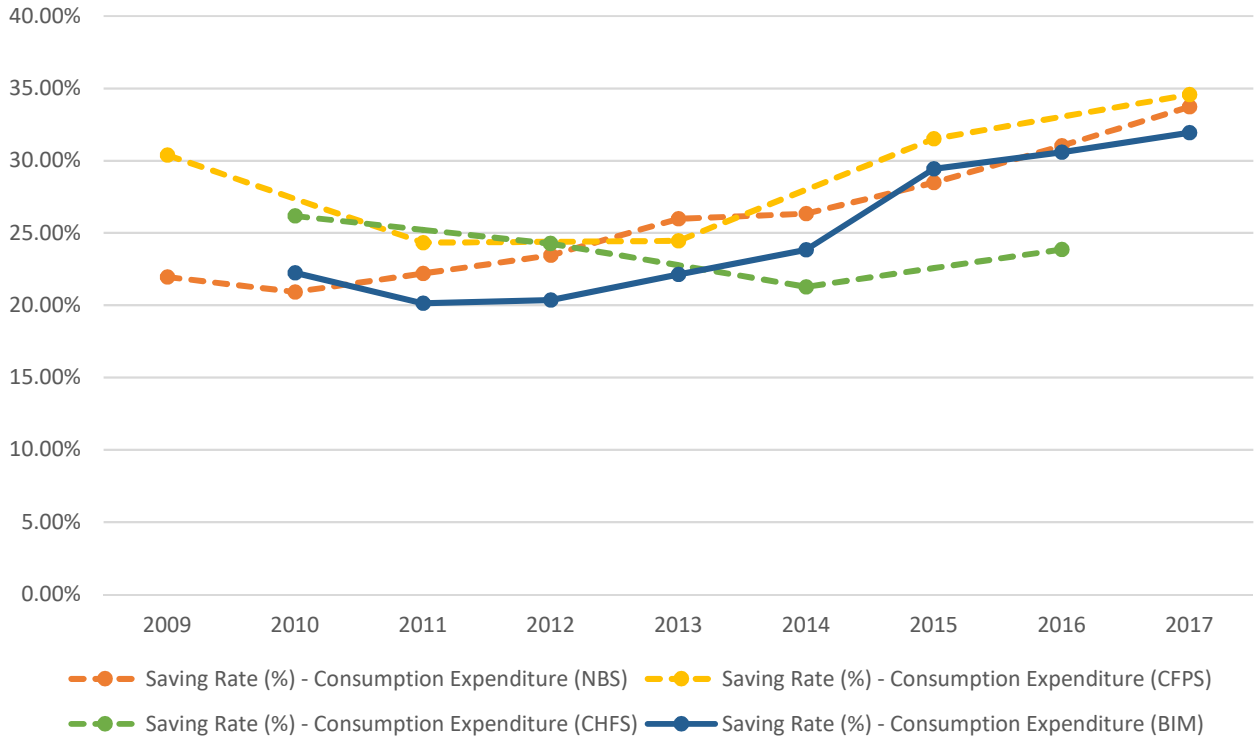


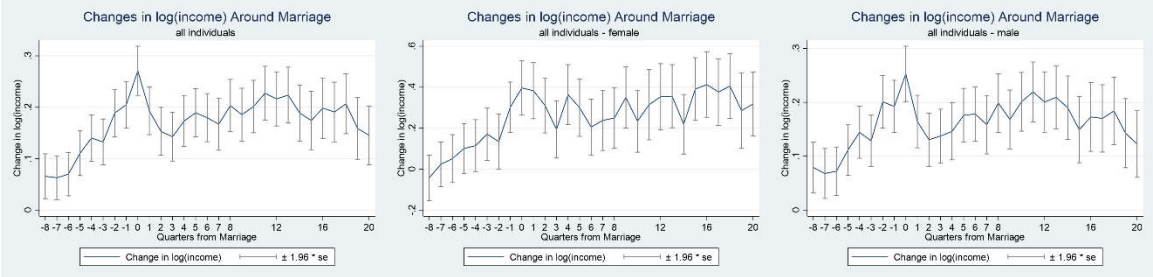
Figure 2. Changes of Income and Consumption around Marriage

Changes of income around marriage and child birth are estimated by equation:

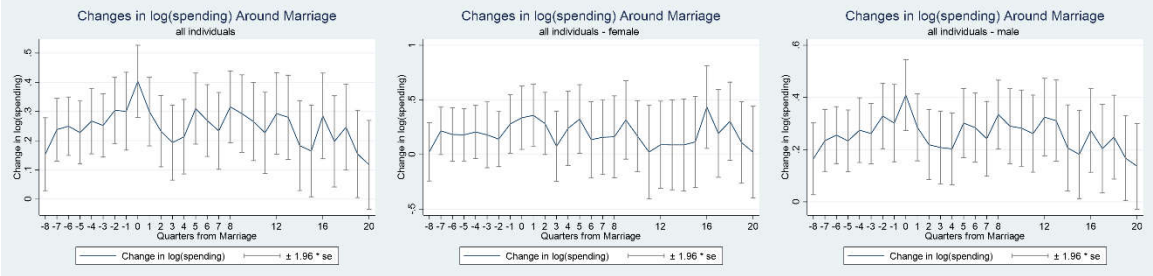
$$\log(income)_{i,t} = \alpha_i + \gamma_t + \mu + \sum_{\tau=-8}^{32} \beta_\tau \times DMarry_{i,t,\tau} + \sum_{\tau=-8}^{32} \delta_\tau \times DChild_{i,t,\tau} + \varepsilon_{i,t}$$

$DMarry_{i,t,\tau}$ is a dummy variable for marriage. $DMarry_{i,t,\tau} = 1$ if for individual i quarter t is τ quarters away from the quarter of marriage; otherwise, it equals 0. $DChild_{i,t,\tau}$ is a dummy variable for child birth. $DChild_{i,t,\tau} = 1$ if for individual i quarter t is τ quarters away from the quarter of child birth. 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, for females subsample, and for male subsample, respectively. Panel B shows the estimations of β_τ about consumption by replacing $\log(income)$ with $\log(consumption)$. Panel C shows the estimations of β_τ about consumption by replacing $\log(income)$ with saving rate. Some control variables in Table 3 including income variability and financial constraint proxy are also included in the saving rate regression.

Panel A. Changes of Income around Marriage



Panel B. Changes of Consumption around Marriage



Panel C. Changes of Saving Rate around Marriage

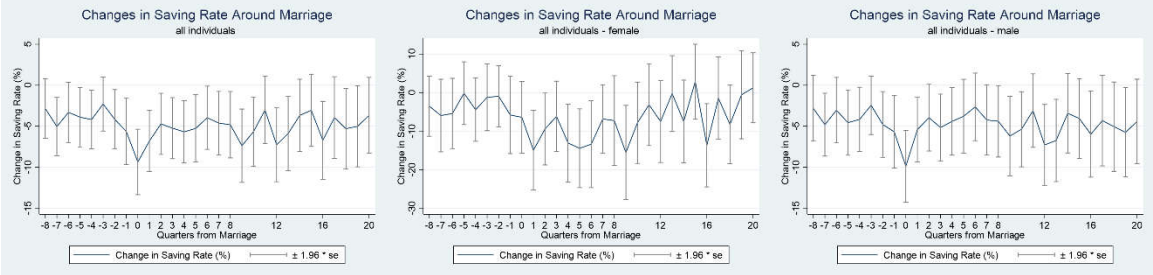
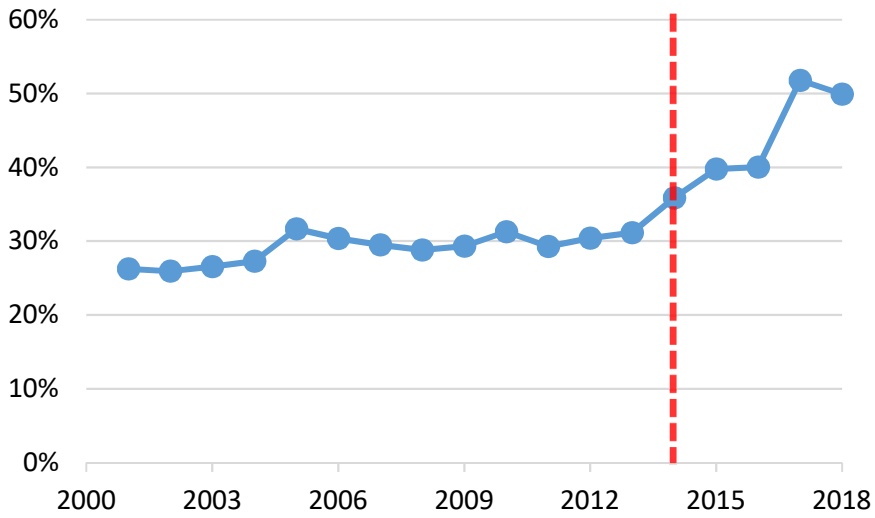


Figure 3. The Effectiveness of the Two-Child Policy

This figure shows the percentage of the number of family's second child born in the year among the total children born in the year. Panel A shows the national-wide statistics from the National Bureau of Statistics and Panel B shows the percentages calculated from our matched Bank sample.

Panel A. Percentage of Second Child in Total Birth: National Statistics



Panel B. Percentage of Second Child in Total Birth: Bank Sample

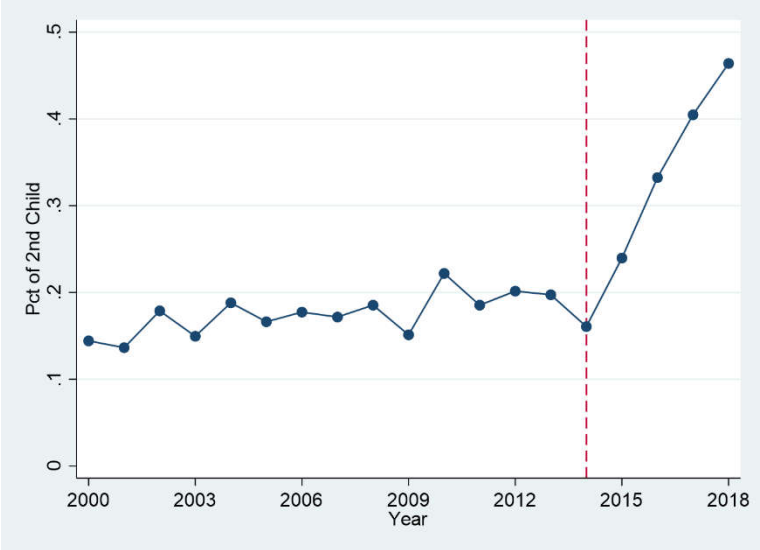
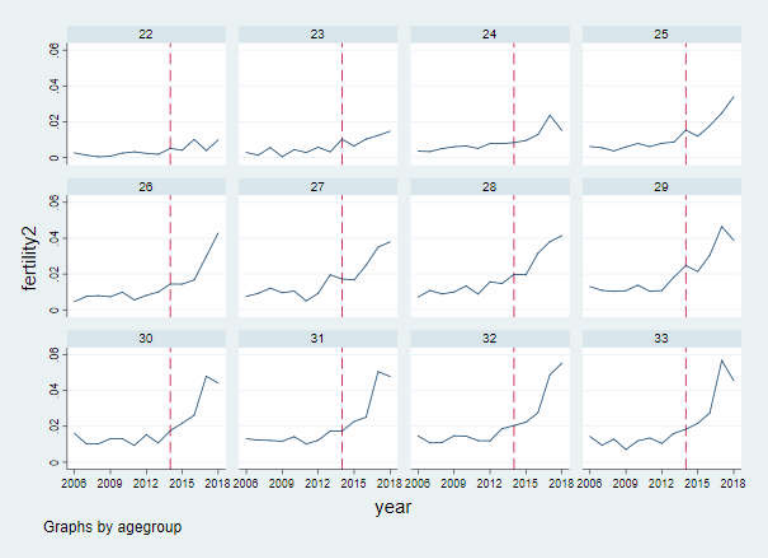


Figure 4. Urban Women Second-Child Fertility Rates in Different Age Groups Around the Two-Child Policy

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. Age Group 22~33



Panel B. Age Group 34~45

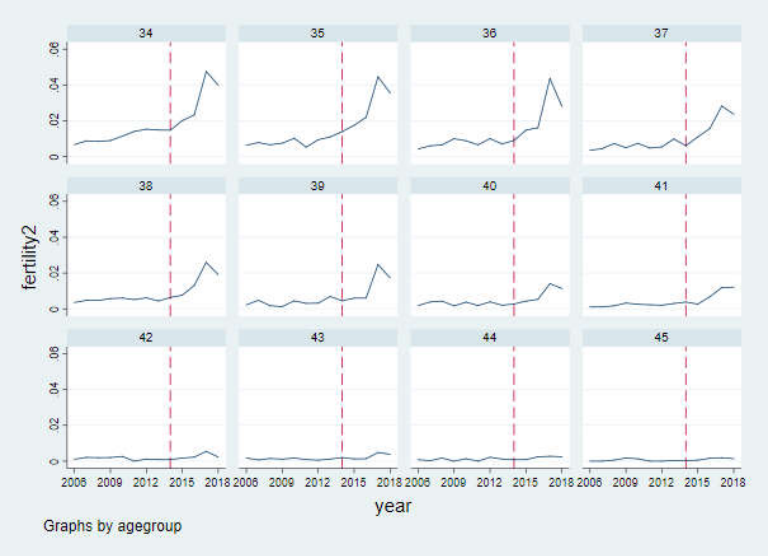


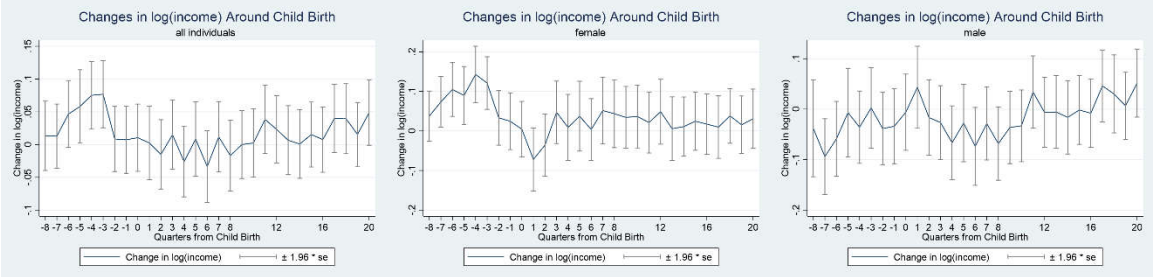
Figure 5. Changes of Income and Consumption around Child Birth

Changes of income around marriage and child birth are estimated by equation:

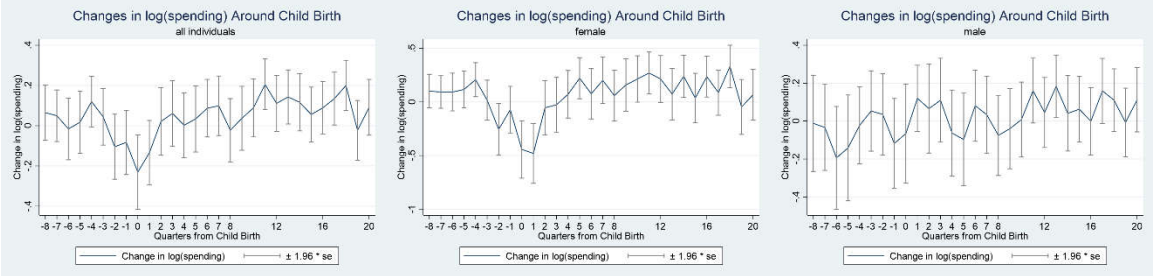
$$\log(income)_{i,t} = \alpha_i + \gamma_t + \mu + \sum_{\tau=-8}^{32} \beta_\tau \times DMarry_{i,t,\tau} + \sum_{\tau=-8}^{32} \delta_\tau \times DChild_{i,t,\tau} + \varepsilon_{i,t}$$

$DMarry_{i,t,\tau}$ is a dummy variable for marriage. $DMarry_{i,t,\tau} = 1$ if for individual i quarter t is τ quarters away from the quarter of marriage; otherwise, it equals 0. $DChild_{i,t,\tau}$ is a dummy variable for child birth. $DChild_{i,t,\tau} = 1$ if for individual i quarter t is τ quarters away from the quarter of child birth. 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 females subsample, and for 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 saving rate. Some control variables in Table 3 including income variability and financial constraint proxy are also included in the saving rate regression.

Panel A. Changes of Income around Child Birth



Panel B. Changes of Consumption around Child Birth



Panel C. Changes of Saving Rate around Child Birth

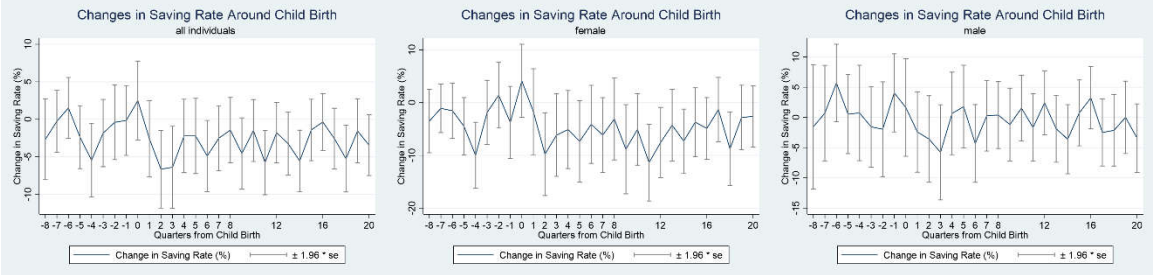


Table 1. Summary Statistics

This table shows the summary statistics of our sample from the Bank and from 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. Saving rate equals one minus the ratio of consumption to income. "Income Variability" is calculated using wage according to Carroll (1992) and Choi et al. (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 and equals 1 if the individual is married in the quarter. "I(ChildAge>30)" is a dummy and equals 1 if the individual has a child older than 30 in the quarter; otherwise. "Has2ndChild" is a dummy and equals 1 if the individual has the second child in the quarter, including the time of pregnancy. The aggregate saving rate is one minus the ratio of average consumption to average income. See Section 1.4 and 1.5 for additional details. For demographics and family status dummies, the latest value of each individual are summarized. Panel B shows the distributions of income, consumption and saving rate, along with the aggregate saving rate in the Bank sample and some other commonly used survey data in China, including CFPS (China Family Panel Survey) and CHFS (China Household Finance Survey). The quarterly Bank sample has been aggregated to annual frequency so that they have the same horizon as the surveys. In the survey data, the sample consists of urban households in Middle and West China provinces which are similar to Inner Mongolia in terms of geographic location and economic development.

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
Saving 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 Saving 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
Saving Rate	13.39%	64.25%	-0.46%	9.80%	36.94%	160,151
Aggregate Saving 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
Saving Rate	13.80%	39.84%	-11.51%	20.02%	42.00%	1,004
Aggregate Saving 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
Saving Rate	10.28%	50.98%	-14.27%	20.69%	46.16%	887
Aggregate Saving 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
Saving Rate	13.57%	55.94%	-10.17%	28.19%	53.25%	4,990
Aggregate Saving 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
Saving Rate	10.99%	57.60%	-14.10%	25.97%	51.67%	7,518
Aggregate Saving 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
Saving Rate	12.32%	54.51%	-10.62%	26.51%	50.68%	7,785
Aggregate Saving Rate	23.87%					

Table 2. The Income Elasticity of Consumption and Its Relationship with Income Quantiles

IncQuintile is the income quintile ranging from 1 (minimum) to 5 (maximum).

	Dependent Variable: ΔLog(Consumption)							
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) IV	(6) IV	(7) IV	(8) 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

Cluster robust standard errors clustered by individual are reported in parentheses.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3. The Impact of Income Level, Income Variability and Financial Constraints on Saving Rate

"Income Variability", "Financial Constraints", "Married", "I(ChildAge>30)" are as defined in Table 1.

	Dependent Variable: Saving 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 Variability		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)						
Industry Dummies	Yes	Yes	Yes	Yes	No	No	No	No	No	No
Age×Industry Dummies	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

Clustered robust standard errors clustered by individual are reported in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4. The Impact of Income Level, Income Variability and Financial Constraints on Saving Rate: Heterogeneity Tests

"Income Variability", "Financial Constraints", "Married", "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.

	Dependent Variable: Saving 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

Clustered robust standard errors clustered by individual are reported in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5. The Impact of Second-Child Policy on Saving Rate: DID Analysis

$Policy_t = 1$ if quarter t is in year 2014 or later; otherwise, it equals to 0. $ZeroChild_i = 1$ if individual i had no child in 2014; otherwise, it equals to 0. $OneChild_i = 1$ if individual i had only one child in 2014; otherwise, it equals to 0. Other explanatory variables are defined in the same way as before.

	Dependent Variable: Saving 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

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6. The Relationship between Saving Rate Increase and the Tendency of Having a New Child after the Two-Child Policy

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 saving rate between the end of 2013 and 2014. "Minority Percentage" is the percentage of minority ethnic population in the individual's residential city.

	Dependent Variable: Dummy: Having a New Child between 2015 and 2017?					
	Logit Regression			Cox Regression		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta SavingRate_{2014}$	0.00855	0.0358	0.0567	0.0103	0.0376	0.0563
	(0.222)	(0.231)	(0.196)	(0.223)	(0.232)	(0.193)
$\Delta SavingRate \times$ OneChild	-0.835	-0.846	-0.765	-0.699*	-0.704*	-0.627
	(0.609)	(0.623)	(0.627)	(0.392)	(0.397)	(0.385)
$\Delta SavingRate \times$ I(31≤Age2014≤41)	-0.0926	-0.112	-0.0968	-0.0922	-0.112	-0.0928
	(0.232)	(0.241)	(0.205)	(0.231)	(0.241)	(0.202)
$\Delta SavingRate \times$ I(Age2014≥42)	0.0294	-0.000123	0.0194	0.0264	-0.00329	0.0179
	(0.223)	(0.232)	(0.197)	(0.223)	(0.233)	(0.194)
$\Delta SavingRate \times$OneChild× I(31≤Age2014≤41)	1.492**	1.538**	1.527**	1.300***	1.343***	1.345***
	(0.656)	(0.670)	(0.672)	(0.436)	(0.443)	(0.437)
$\Delta SavingRate \times$OneChild× I(Age2014≥42)	0.857	0.855	0.853	0.719*	0.711*	0.714*
	(0.613)	(0.628)	(0.635)	(0.398)	(0.403)	(0.397)
OneChild	0.255**	0.186*	0.140	0.249**	0.182*	0.128
	(0.104)	(0.106)	(0.112)	(0.102)	(0.104)	(0.109)
Log(Avg. Income in 2013~2014)		-0.279***	-0.281***		-0.273***	-0.269***
		(0.0453)	(0.0619)		(0.0444)	(0.0598)
Male Dummy			-1.034***			-1.006***
			(0.113)			(0.111)
Minority Percentage			0.610			0.613
			(0.394)			(0.386)
Observations	14,249	14,249	14,249	14,249	14,249	14,249
- Industry Dummies	No	No	Yes	No	No	Yes

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

Table 7. Summary Statistics for Sample of Linked Households

This table shows the summary statistics of the quarterly household sample. The data source includes transaction data from Bank of Inner Mongolia, and the marriage and child birth data from administrative system. Records no earlier than the year before the marriage year have been kept. Income and consumption are the sum of the husband's and wife's income and consumption, respectively. Saving rate equals one minus the ratio of consumption to income. "Income Variability" is calculated using wage by the same method as used in table 3, expect the detrended family paycheck income is regressed on husband's and wife's age, age square and the interaction of age and industry. "Credit Accessibility" is the sum of husband's and wife's individual credit accessibility constructed by the same method as table 3. "GettingMarried" is a dummy and equals 1 if the couple get married; otherwise, it equals 0. "Haschild30" is a dummy and equals 1 if the individual has a child older than 30 in the quarter; otherwise, it equals 0. It is a proxy of the saving incentive for unmarried child. "Duration" is the number of quarters in which the household of the records have observations. The aggregate saving rate is one minus the ratio of average consumption to average income.

	Mean	Standard Deviation	25% Percentile	Median	75% Percentile	Observations
Income(yuan)	41762.27	60620.23	16657.8	25707.09	42863.68	11874
Consumption(yuan)	30306.08	40002.59	11610	20348.85	34336.58	11874
Saving	0.126916	1.104786	-0.02279	0.185588	0.484499	11874
Income variability	0.125926	0.176038	0.021657	0.059164	0.1445	11064
Credit accessibility	33.09342	13.53368	23.0144	30.8162	40.67237	11874
Husband Age	36.43524	10.71995	28	33	42	11874
Wife Age	34.83359	10.26765	27	31	41	11874
GettingMarried	0.870136	0.336167	1	1	1	11874
Haschild30	0.006401	0.07975	0	0	0	11874
Duration(quarter)	10.73135	7.026948	5	9	15	11874
Aggregate Saving Rate	27.43%					

Table 8. The Impact of Income Level, Income Variability and Financial Constraints on Saving Rate

"Income Variability", "Credit Accessibility", "GettingMarried", "Haschd30", are defined as in Table 7.

VARIABLES	Dependent variable: Saving 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 hh	1580	1558	1580	1558	1558	1558

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 9. Relative explanatory power of factors

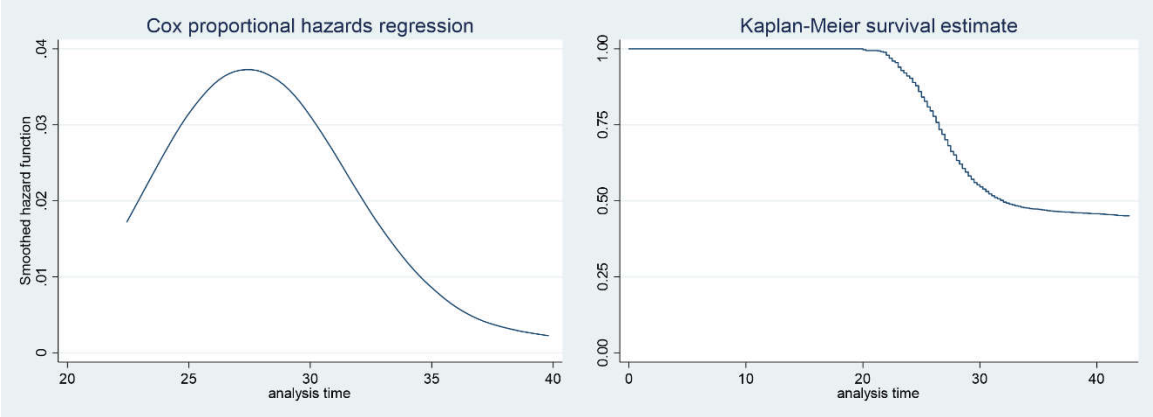
We decompose the predicted saving rate change between year 2010 and year 2017, the beginning and end year of our sample, into contributions from five explanatory variables based on regression equation (11). The table below presents the proportion of each explanatory variable's contribution to the predicted saving rate change, which are their relative explanatory power. The standard error is calculated using nlcom command in Stata.

	(1)	(2)
VARIABLES	Saving rate change	Saving 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

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Figure A1. The Timing of Marriage

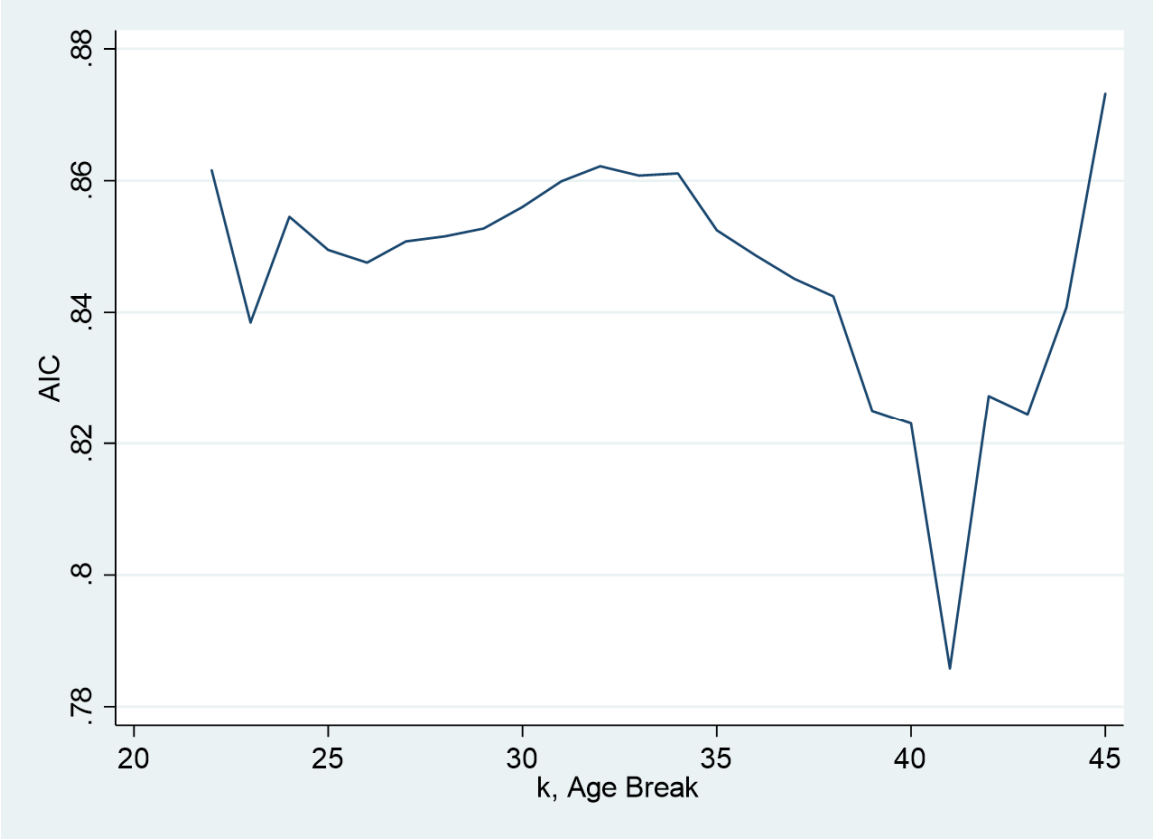
This figure shows the results from a Cox hazard regression about marital decision using the loan application dataset. The timing variable is age and the control variables include income level, income variability, sex, industry dummies and calendar-quarter dummies. Panel (a) shows the estimated smoothed 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.



(a) (b)

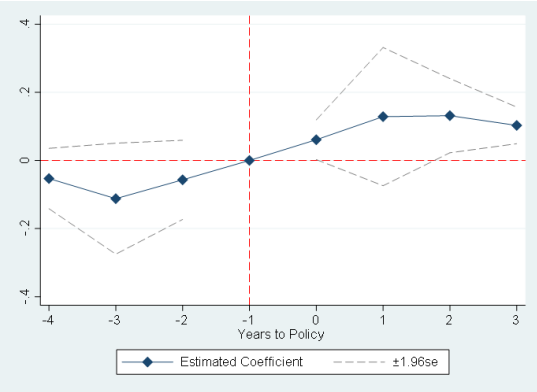
Appendix Figure A2. The Performance of the DID Models about Second-Child Fertility Rate with Different Age Breaks

This figure displays the performance of the DID models about second-child fertility rate as discussed in Section 5.1. Different age breaks (represented by symbol *k* in the equations) are used in the models and the corresponding AIC's are plotted.

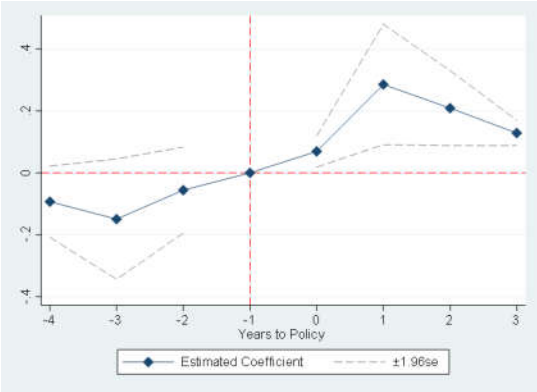


Appendix Figure A3. Tests of the Parallel Trend Assumption in DID Analysis

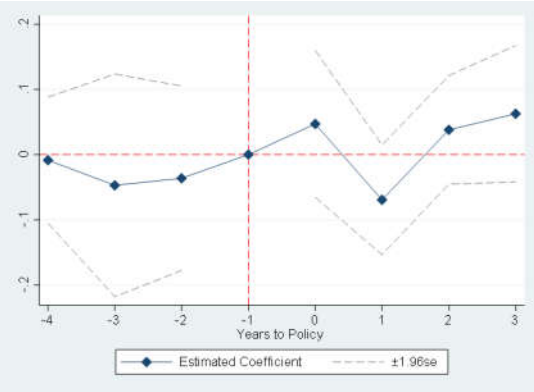
This figure shows the results from equation (16) which tests the validity of the parallel trend assumption in our DID specifications. The graphs plot the estimated δ_τ's along with the 95% confidence levels, which capture the saving rate impacts of the two-child policy in the τth year after the implement 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 zero. Figure (a) to (c) presents the results for the whole matched sample, the female matched subsample, and the male matched subsample, respectively.



(a)



(b)



(c)

Appendix Table A1. The Income Elasticity of Consumption: First-Stage Regressions

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 their relationship with the coal industry. CoalPos = 1 if the industry's profits are generally positively-correlated with coal prices; otherwise, CoalPos = 0. CoalNeg = 1 if the industry's profits are generally negatively-correlated with coal prices; otherwise, CoalNeg = 0. ΔLog(CoalPriceIndex) is the log-difference of China Coal Price Index. Quarter fixed-effects and individual fixed-effects are also included in the regressions. Cluster robust standard errors clustered by individual are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

	(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

Appendix Table A2. Estimations of Individuals' Credit Accessibility

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 equals 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 2010~2017. Robust standard errors are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

	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