The Life Cycle Model and Household Savings: Micro Evidence from Urban China*

Abhijit Banerjee† Xin Meng‡ Nancy Qian§

26th September 2010

Abstract

This study investigates the extent to which China's high household savings rates can be explained by the life cycle theory. First, we document that Chinese parents depend on their children for support when elderly and that sons provide more support than daughters. Second, we test the two predictions of a simple life-cycle model that account for these two facts: 1) parents will increase savings when they have fewer children; and 2) the reduction in fertility will increase savings more for parents who have only daughters. To establish causality, we exploit the plausibly exogenous decline in fertility in China caused by family planning programs which began in the early 1970s and the fact that there was no sex-selection for our sample of urban Chinese households that had children during the 1960s and 70s. Our results show that for parents who have a daughter as the eldest child, having one child less increases the savings rate by over 14,000 RMB, which is approximatley 27% of average income; for those who have a son as the eldest child, there is no effect. Finally, we apply our estimates to a simple life-cycle model of savings to predict the level and rates of savings, and thus assess the extent to which the life-cycle model can explain the data.

Keywords: Savings, China's Saving Puzzle, Life-cycle Hypothesis, Demographic Structure

*We are grateful to Paul Schultz, Chris Udry and Shangjin Wei for their insights; and David Y. Liu and Xiaoxue Zhao for excellent research assistance.

†banerjee@mit.edu

‡xin.meng@anu.edu.au

§nancy.qian@yale.edu.

1 Introduction

One of the most puzzling facts in economics in the past three decades is the extraodinary growth and high levels of Chinese household savings, which increased dramatically from a stable rate of approximately 5%, during 1949-78, to a high of almost 34% in 1994.[1] Interestingly, this pattern in savings ratios was paralleled by similar changes in demographic structure. Figure 1 shows the dramatic increase in savings rates and the ratio of the employed population and minors (E/M).

This paper studies the extent to which the high levels of household savings can be explained by a life-cycle model where elderly parents depend on adult children for support. There is very little social security provision for Chinese workers today. Recent retirement surveys show that the elderly depend largely on their adult children for support. Moreover, they receive much more support from sons than from daughters, both in terms of monetary transfers and cohabitation. In this context, a standard life-cycle model will make two clear predictions. First, a reduction in fertility will increase savings. Second, the increase will be larger in magnitude for parents who have daughters relative to those who have sons.

This paper conducts three exercises to investigate the importance of the life-cycle hypothesis in explaining China's high savings rates. First, we use recent survey data to document the importance of children to elderly parents, and the difference in support from daughters and sons. Over half of the retired elderly population in China live with at least one adult children. Sons provide twice as much transfers to parents as daughters both in terms of the number of transfers and the average amount of each transfer. These facts are consistent with the widely held beliefs that Chinese parents depend on their children in old age and that sons provide more support than daughters.

Second, we test the theoretical prediction of the life-cycle model we described earlier by examining the causal effect of the number of children on parents' savings, and how that effect differs for parents with daughters versus those with sons. The main empirical difficulty arises from the endogeneity of the fertility and savings decisions. For example, a negative correlation between fertility and savings could reflect the presence of a third omitted factor such as income growth. Households facing higher future income growth, a factor that the econometrician cannot observe, will have more children and save more. Therefore, the observed correlation will be confounded by this omitted variation and not reflect the true causal impact of fertility on savings. Similarly, one could be concerned that the sex of the children are endogenous to sex selection. Since the 1980s, the sex-ratio of children have become increasing biased towards boys.[2]

We address this difficulty by exploiting the plausibly exogenous variation in fertility caused by a shift in family planning policies. In 1972, family planning policies shifted from a pro-fertility agenda to one that focused on curbing fertility. This policy shift affected the entire nation. Since the initial policies promoted longer spacing between births and were soon replace by policies that curbed the total number of births, parents who had their first

[1] This was as high as the savings rates of Japan in the 1960s, despite the fact that GDP per capita in China during the increase in savings rate was many times lower (Horioka, 1990).

[2] Past studies have shown that the sex of children in China are responsive to relative female income and household income.

child during or after 1972 had fewer children on average than parents who had their first child prior to 1972. Furthermore, by focusing our study on urban Chinese households that had their first child five years before the policy shift, we study a context in which there was no sex-selection. Therefore, the sex of the first child can be interpreted as exogenous and we do not need to be concerned about selection biases in interpreting the estimated effects of having a son rather than having a daughter. Our empirical strategy estimates the effect of the total number of children a household has and its interaction with the sex of the eldest child. We instrument for the number of children with a dummy variable for whether the first child was born during or after 1972; and instrument for the interaction of the number of children and the sex of the eldest child with the interaction of whether the eldest child was born in or after 1972 and the sex of the eldest child. Conditional on a large set of baseline controls, the 2SLS estimate of the number of children establishes the causal effect of an additional child on savings for households that have a daughter as a first child.[3] The sum of this estimate and the estimate for the interaction term reveals the effect for households that have a son as the first child.

The empirical analysis uses a unique survey conducted in 2008 that reports both savings and the total fertility of a household. We focus only on urban households because both family planning policies and savings instruments vary greatly between rural and urban areas. Moreover, rural-urban differences in factors such as government enforcement against sex-selection and the economic advantage of having a son relative to having a daughter mean that while sex-selection features prominently in rural areas, it is less of a problem in urban areas. This is especially the case for our sample, which is restricted to households where the first child is born within five years of the policy shift, and therefore born many years before sex-selective abortion was made possible through the introduction of ultrasound B machines. The fact that the sex of a child is exogenously assigned in our sample is evidenced by the fact that the children of the households in our sample, who are on average in their early thirties when the survey was conducted in 2008, have approximately 49.8% males, which is similar to the gender ratios in Western countries not known for boy-preference.

Our first stage estimates show that households that had their first child during or after 1972 had on average one less child than households that had their first child prior to 1972. This is consistent with the fact that average household size decreased from slightly above two to approximately one. The 2SLS estimates show that reducing the number of children by one causes parents to increase savings by 14,008 RMB, which is approximately 27% of average household income. This effect is almost entirely driven by households for whom the first child is a daughter, which in our context effectively means that the only child is a daughter. Our main empirical results are consistent with the life-cycle model and suggest that it plays an important role in explaining the high savings rates.

In addition to the main results, we find that the difference in savings behavior between parents that have daughters versus those who have sons is driven by the gender wage gap

[3]The baseline controls include the age of the household head (and its squared term), the education of the household head (and its squared term), a dummy for whether the household head is more than 55 years of age, the age of the youngest child, a dummy for whether the younges child is less than 22 years of age, and city fixed effects.

in their city of residence. In other words, in cities where women earn similar wages as men, a reduction in fertility has broadly similar effects on parents that have daughters and sons. However, in cities where women earn much less than men, a reduction in fertility will increase the savings of parents with daughters much more than of parents with sons. If parents use current labor market wages to infer their future financial support from children, then these results provide additional support for the hypothesis that parents view their children as a form of savings, and when their access to this savings instrument is restricted, they must save more in cash.

There are two important caveats to the interpretation of the main results. First, there is the concern that our estimates are driven by the fact that parents that had their first child after family planning policies were introduced in 1972 are younger on average and in a different point in their life cycle savings behavior when we observe them in our survey. This could cause them to save more than parents who had their first child before 1972 and who are also older. To address this possibility, we show results both controlling and not controlling for the age of the household head and find that this change in controls has little effect on the estimates. In addition, we conduct and alternative experiment and compare the savings rates of those who are over sixty years of age with those who are below, irrespective of when their first child was born. While the older cohorts have more children, and therefore should save less according to our theory, we find no effect of age, suggesting that the pure age effect is, if anything, positive, and our estimates may understate the magnitude of the true effect of fertility on savings. Base on these findings, we conclude that our main results are not confounded by the selection of the age of the parent.

Second, we may be concerned that parents who have children earlier in life may differ from those who have children when they are younger – they may be more ambitious or have higher unobserved human capital (observed human capital is controlled for). This concern is especially true after we control for age of the parents. To investigate this, we directly examine the effect of age at first birth on savings. We find no correlation between the age at first birth and savings. Therefore, we conclude that our main results are robust to these factors.

Our final exercise assesses the extent to which a simple life-cycle model can explain observed Chinese households savings. We use our empirical estimates together with a simple life-cycle model of savings in the presence of transfers from children to parents and credit constraints to predict the level and rates of savings under realistic assumptions about income growth and interest rates. This calibration exercise shows that we cannot easily match both the observed magnitudes of savings levels and the responsiveness of savings to an additional child or an additional male child. In particular, models that generate high enough levels of savings tend to rely on children having a low propensity to make transfers to parents, which makes it difficult to explain the large effects that we find for the feritlity and gender and of children. These results suggest that in order to fit the data, the standard life cycle model should be augmented to include other forces that increase savings. On obvious possibility is a perception of high risk, which is realistic in light of the rapid changes and historical tumult that China has experienced in recent decades.

This paper makes significant progress towards understanding the determinants of China's high savings rates and the importance and limitations of the life-cycle model more generally. The evidence on the extent to which the LCH can explain China's high savings rates has been mixed. On the one hand, Modigliani and Cao (2004) uses correlational evidence and time series data to show that changes in the demographic structure can explain a substantial amount of the time-series variation in savings rates. Chamon and Presad (2008) attempts to quantify the contributions of life cycle motives and other motives and find that the former are an important factor. Blanchard and Giavazzi (2005) suggest that life cycle motives are especially important when there is high uncertainty about future income and when state provision of welfare and pension is poor (Diamond, 2004). On the other hand, studies such as Horioka and Wan (2006), which apply a LCH model to Chinese panel data find little support for it.

To the best of our knowledge, our study is novel in exploiting changes in the demographic structure induced by family planning policies to identify the causal effects of changes in the demographic structure on savings. A recent study by Wei and Zhang (2009) also relates family planning policies to savings in China. They argue that the increase in savings rate is driven by the increase in sex selection caused by the One Child Policy (OCP) and marriage market competition.[4] Our study differs from theirs in that there is little sex selection in our context. More importantly, we are mainly interested in the role that family size plays in determining savings. Note that in controlling for city fixed effects, our empirical strategy always controls for regional differences in sex-ratios.

This paper is organized as follows. Section 2 discusses the background on savings and family planning in China. Section 3 documents the support that children provide to elderly parents. Section 4 presents the empirical strategy. Section 5 describes the data. Section 6 presents the results. Section **??** uses our empirical estimates to assess the extent to which the life-cycle model can explain observed savings in urban China. Section 8 offers concluding remarks.

2 Background

2.1 Savings

During the first three decades of the People's Republic of China (1949-78), household savings were approximately 5% of total household income. It began a rapid and constant rise during the reform era beginning in 1978.

Chinese urban households had few instruments for savings. Money could be deposited in Urban Credit Cooperatives (UCC), which later became city-level "Commercial" banks; or it could be held as cash. Banks (and UCCs) provide very low interest rates for savings deposits. During the 1980s, annual real interest rates for savings deposit ranged from 0.7 to 1%. In the late 1990s, housing became a savings vehicle with the privatization of the urban housing stock. And more recently, reforms of financial markets have allowed a

[4]They argue that the One Child Policy together with son preference increased the fraction of males and therefore increased compeititon for brides on the marriage market. Hence, in anticipation of paying increased bride prices, parents of boys must increase their savings.

small number of urban households to invest in stocks. However, this is unlikely to affect many households. Almost all household savings in urban areas are still deposited in banks (He and Cao, 2007).[5]

The most common way to save in China, as in any developing country, is arguable by investing in children. Children will provide care for elderly parents. We document this with recent survey data in Section 3.

2.2 Family Planning

The history of family planning policy is public information and documented (in Chinese) by the China Population Information Network (POPIN), a branch of the China Population Development and Research Center (CPDRC or CPIRC).[6] In this section, we summarize the policies that are most relevant for our study. Family planning has been featured in the discussions of Chinese policy makers since the founding of the People's Republic in 1949. It can be characterized as two phases. The first phase promoted fertility. In the beginning of the Communist regime, the government actually advocated large families as it regarded high population growth as a sign of superiority of socialism. In 1952, the government published a regulation to restrict sterilization or abortions. For example, women were banned from sterilization operations or abortions unless they are over 35 or have six or more children, one of whom is over ten years old, or if giving birth may severely impair their health. By 1953, the Chinese population had reached 600 million and the annual population growth rate was at 23%. There was a shortage in housing and spaces in schools for the growing number of children, leading some to advocate for shifting family planning policies from promoting fertility to reducing fertility.

The discussion on family planning policies shifted away from pro-fertility to fertility-control during the 1960s.[7] Attempts were made to curb population growth. However, most of these were half-hearted and led to no concrete actions or declines in birth rate. Part of the difficulty in implementing these new policies arose from the fact that many of the advocates were victims of the Anti-rightist Movement (1957) and that attention was taken away from family planning by other policies such as the Great Leap Forward (1955-60). In the early 1960s, after the Great Famine (1959-61) killed up to 30 million individuals, family

[5] According to the 2002 round of the China Household Income Project (CHIP), average urban households hold approximately 10% of their total savings in stocks and bonds.

[6] See http://www.cpirc.org.cn/yjwx/yjwx_detail.asp?id=308. CPIRC also produces following publications: China Population Today, China Family Planning Yearbook, Population and Family Planning, Population Abstracts and China Population Data Sheet.

[7] Deng Yingchao first proposed relaxing the ban on contraceptive/sterilizing measures to the then vice Premier, Deng Xiaoping, in 1953. In December 1953, Liu Shaoqi hosted a party meeting to clarify that the party is not against birth controlling measures in face to the rising population. In 1956, the central government published "1956-57 National Agricultural Development Outline" to advocate family planning in all areas, except where the ethnic minorities reside. In the same year, Zhou Enlai included a family planning agenda in the 2nd Five-Year Plan report and pointed out that the Ministry of public health should promote family planning measures. Mao Zedong also advocated family planning at the 11th meeting of the Supreme State Conference in 1957. Scholars and other politically active persons, most noticeably Ma Yinchu and Shao Lizi, also wrote articles to advocate governmental control on fertility.

planning became a renewed agenda for some members of the government.[8] However, the Cultural Revolution, which started in 1966, again disrupted the government's efforts at birth control. Many family planning offices were either canceled by the revolution committee" or simply became dysfunctional. By 1969, the Chinese population had reached 800 million, prompting Zhou Enlai to reemphasize the importance of family planning in 1970. In 1971, Mao Zedong followed suit and announced that "population must be controlled", which signaled a turning point in family planning policy practice in China.

The second phase of family planning policies is characterized by a shift of focus towards curbing fertility. The plans were laid out in 1971.[9] In practice, efforts began in earnest in 1972. On January 17, 1972, provincial leaders attended a meeting organized by the Ministry of Public Health where the central government demanded that local governments publicize and enforce Mao's instructions on family planning, and instructed all levels of government to establish or reinforce their bureaucracies for organizing or implementing family planning related tasks. In May that year, the Ministry of Public Health organized a national workshop on family planning measures where all provinces had to participate. These measures stated and clarified the shift in family planning policy and effectively re-activated the bureaucracy. In another mandatory meeting for all provincial leaders on November 1, 1972, it was agreed that within that year, 23 provinces had established the necessary bureaucracies for implementing family planning related policies.

The empirical strategy of this paper is to exploit this shift in family planning policy in 1972 for exogenous variation in total fertility rates. The initial birth control policies emphasized birth spacing after the first child rather than a strict cap on the total number of children. However, since women are biologically less likely to have children at older ages and because the family planning policy eventually tightened such that it capped the number of children parents could have, households that already have their first child after the policy shift will on average have fewer children in total. Therefore, our empirical strategy assumes that parents who had not yet had their first child by 1972 are likely to have fewer children than parents who had already had their first child.

There are several important facts to keep in mind for our empirical analysis. First, the policies for population control gradually tightened over time. For example, one of the first programs that was implemented was the "late, thin and few" policy. Couples were encouraged to have children later in life, have only one child (or at most two children), and have at least three or four years in between births. A more stringent version of family planning, known as the One Child Policy (OCP), was introduced in 1979/80. This

[8] In December 1962, the central government issued a notification that acknowledged that future agenda should include discussions of family planning policies. In October 1963, the State Council determined that family planning committees should be established at all levels of government to organize family planning activities.

[9] On Feb. 15th, 1971, Zhou Enlai re-emphasized the importance of family planning when meeting with the provincial representatives at the National Planning Conference in Beijing: "It's important to control population growth. Government should advocate late marriages and birth control, and vigorously publicize these policies from now on. On July 8th, the State Council published "the Report on Doing Well in Family Planning". The written instruction by the State Council on the document pointed out that "Family planning is an important issue that Chairman Mao has advocated for years. All levels of officially must treat the issue seriously."

punished households that had more than one child with fines, job loss, and the loss of access to public goods, and rewarded those with only one child with bonuses. Family planning polices also became better defined over time. For example, in 1978, the state defined details on things such as what counted as late marriages and the bonuses and subsidies for workers and farmers if they go through sterilizing operations, etc.[10] This means that the effect of family planning policies on total fertility is not uniform across households that have their first child after 1972; the later they have their first child, the fewer children they will have. This does not affect the validity of our strategy. But for interpreting our first stage estimates, it is important to keep in mind that they capture the average effect across all households that had their child after the policy shift.

Second, there is much regional variation in family planning policies. The greatest differences have existed between rural and urban areas, and between Han ethnic households (who make up approximately 92% of China's population) and ethnic minorities. For the purposes of our study, the key is that family planning policy is relatively uniform across urban areas (e.g., Ebenstein, 2010; Qian, 2009) and that there are relatively few ethnic minority households in most Chinese cities. Variation across cities does not affect the validity of our empirical strategy, which estimates the average change after 1972. However, in interpreting the results, one should keep in mind that we are capturing the average effects.

Finally, it is important to keep in mind that one of the outcomes of family planning was a rise in boy-biased sex ratios. This is mostly a rural phenomenon as it was typically easier to evade government enforcement in the countryside. Rural households that wanted a son often hid female infants and attempted to try to have a son by having a second child. In extreme cases, they could also commit female infanticide more easily than urban households. Many of the relaxations in family planning policies during the 1980s were the government's response to these phenomena as it attempted to curb sex-selection. We largely avoid potential selection issues caused by sex-selection by focusing on an urban sample. However, the introduction of reliable Ultrasound B machines used for pre-natal sex detection during the 1980s decreased the cost of sex-selection for all parents. And although the government banned its use to reveal the sex of a fetus, anecdotal accounts and the rapidly rising boy-biased sex ratios at birth suggest that this was not well-enforced. The fact that our sample only includes households that had their first child during the late 1960s and 1970s, before sex-selective abortion was available, largely avoids these potentially confounding effects. It is important to note that in our sample, there is no evidence of sex selection.

3 Dependence on Children

In this section, we document two facts to support the widely held beliefs that Chinese parents view their children as a form of saving for when they are elderly, and that sons provide more support to parents than daughters. Since we need to document the dependence on children of retired individuals, we use data from the *China Household Retirement*

[10] See "The Report on the State Councils Family Planning Groups First Meeting"(1978).

Longitudinal Survey (CHRLS). This is a survey at the household level that only samples the elderly. Due to the small sample size, we cannot break the sample into rural and urban. However, to make it broadly comparable to the data used in the main analysis, we restrict the sample to those who had their first child during the 1970s and early 1980s. Table 1 shows that, as expected, more than half of the surveyed elderly live with their children. Panel B shows that parents are more likely to live with their children if one of the parents is in poor health, which is consistent with the belief that children take care of their parents when the latter are elderly. The advantage of having many children is consistent with the observation that the fraction of parents living with at least one adult child increases from 56% for parents with only one child to over 70% for those with seven or eight children. Interestingly, Table 1 shows that parents are twice as likely to live with an adult child if the eldest child is a son rather than a daughter. Table 2 shows transfer income. It shows that approximately 40% of all transfers received in the past year are from the eldest child. However, in terms of regular transfers (those that parents expect to receive on a regular basis), parents receive almost three times as many from sons than daughters, and the average amount of each transfer from sons is 40% higher than those from daughters.

4 Empirical Strategy

A standard life-cycle model where parents depend on children for old age support and are credit constrained predicts that a reduction in fertility will increase savings. If sons are able to provide more support than daughters, than the reduction will be larger if parents have only daughters. To examine this empirically, we can estimate the following baseline second stage equation.

$$savings_{ijt} = \beta_1 kids_{ijt} + \beta_2 kids_{ijt} \times 1stSon_{ijt} + \Gamma X_{ijt} + \gamma_j + \varepsilon_{ijt} \quad (1)$$

Savings for household i which lives in region j and had their first child in year t is a function of: the total number of living children ever born to the household, $kids_{ijt}$; the interaction of the total number of kids and a dummy variable indicating whether the first child is a son, $kids_{ijt} \times 1stSon_{ijt}$; a vector of household level controls that includes the main effect of whether the first child is a son, X_{ijt}; region fixed effects, γ_i; and a household-specific error term, ε_{ijt}.

In this equation, the effect of an additional child for households that have a daughter as the first child is reflected by β_1 and the effect for households that have a son as the first child is reflected by $\beta_1 + \beta_2$. β_2 reflects the difference in the effect of having an additional child between parents that have a daughter as a first child to those that have a son. When we present the results, we will present both the individual coefficients and the sum to show the effect for the latter group. If a reduction in fertility increases savings for parents whose first child is a daughter, then $\hat{\beta}_1 > 0$. Similarly, if it increases savings for parents whose first child is a son, then

$_1 + \hat{\beta}_2 > 0$. If the reduction in fertility increases savings more for parents that have a daughter rather than a son, then $\hat{\beta}_2 < 0$.

The vector X_i includes many household specific controls. First, since savings rates differ over the life cycle and because mandatory retirement age for state employees is 55 for female workers and 60 for male workers, our main specification includes a dummy for if the household head is over 55 and controls linearly for the age of the household head, though we show results without it as well. In practice, the vast majority of urban workers in our data are in the labor force until at least 65 years of age. However, it is presumably more difficult to gain employment at these higher ages and those who do find employment may be paid lower wages. Our control variables addresses these potential differences. We also show a specification which does not control for the age of the household head in order to see whether the potential selection of households by whether they had a child earlier or later in life (once we control for age of household head, this is what we are comparing), makes a difference. Since one comparison is closer to a pure cohort comparison while the other involves comparing those within the same cohort who had children early and late, we would worry less about the selection effect mentioned above if the two specifications yield the same effects. As this is an important issue, additional strategies for dealing with it are discussed later in the paper in the section on robustness.

Second, we control for the years of education of the household head and its squared term to address the potential endogeneity of both savings and the age at first birth with respect to education.

Third, savings may also be affected by the age of the youngest child. In particular, parents will be able to save less if any one of their children still needs financial support from them. Therefore, we control for a linear measure of the age of the youngest child, its squared term, as well as a dummy variable for if she is less than 22 years of age, and therefore still possibly in school. Instead of using a linear measure for the age of the youngest child and a dummy for whether she is less than 22 years of age, we can alternatively use dummy variables to control for the age fully flexibly. This does not affect the results. Therefore, for the sake of brevity, we do not present these alternative specifications in the paper.[11]

Fourth, we control for city fixed effects, which will control for all birth-cohort invariant differences across cities. Therefore, factors such as regional sex ratios, which a recent study by Wei and Zhang (2009) argue affects savings rates, are implicitly controlled for in our baseline estimation.

Finally, we control for the sex of the first child to control for the fact that sons provide more support than daughters. We only control for the sex of the first child because the sex of later children are more likely to be endogenous, especially for those who are having additional children against the law. This is not likely to be a serious issue for our study since there is no evidence of sex selection in our sample. But we take this precautionary measure to be extra cautious.

Recall from the introduction that the main difficulty for interpreting the OLS equation is that the correlation of an additional children confounds the effect of having child with other factors such as income growth, and that there may be a reverse causal relationship

[11] They are available upon request.

between savings and fertility. For example, wealthier households or household that expect higher future income may also choose to have more children. To address this, we exploit the exogenous variation in the number of children born caused by the shift of family planning policies towards reducing fertility. Households who had their first child after family planning policies were introduced in 1972 would have fewer children on average.

There are two endogenous regressors in equation (1), for which we will have two instrumental variables. The first stage equation for equation (3). The two first stage equations can be written as the following:

$$kids_{ijt} = \alpha_1 Post72_{ijt} + \alpha_2 Post72_{ijt} \times Son_{ijt} + \Gamma X_{ijt} + \gamma_j + \eta_{ijt} \tag{2}$$

and

$$kids_{ijt} \times 1stSon_{ijt} = \alpha_1 Post72_{ijt} + \alpha_2 Post72_{ijt} \times Son_{ijt} + \Gamma X_{ijt} + \gamma_j + \eta_{ijt} \tag{3}$$

The number of children in household i living in region j with a first child who was born in year t is a function of: a dummy variable for whether the first child was during or after 1972, $Post72_{ijt}$; the interaction term between $Post72_{ijt}$ and a dummy if the first child is a son, Son_{ijt}; and the same vector of controls as in equation (1). α_1 reflects the effect of having a first child in 1972 or afterwards for households that have a daughter as a first child and $\alpha_1 + \alpha_2$ reflects the effect for households that have a son for the first child. If having a the first child during or after 1972 decreased total fertility, then $\hat{\alpha}_1, \hat{\alpha}_1 + \hat{\alpha}_2 < 0$.

The strategy relies on the assumption that, conditional on our baseline controls, households who had their first child after 1972 did not differ from households that had their first child before in any way other than total fertility. These policies were introduced nationwide (at least with respect to urban areas) and in urban areas, they were introduced independent of changes in aggregate economic circumstances.[12] Therefore, they are most likely exogenous to household level factors. Because our strategy relies on the average change after 1972, the identification strategy is not confounded by differences in enforcement of family planning policies across regions. In other words, if actual policies differed across regions, our estimates captures the average effect. Our numerous baseline controls address most of the concerns regarding endogeneity.

There are two important caveats to our strategy. First, households in the treatment group (e.g., those that have their first child in or after 1972) and those in the control group (e.g., those that have their first child prior to 1972) differ in that parents of the latter are also likely to be those who chose to wait to have children. These parents may be more ambitious or have higher unobserved human capital. If age at first birth affects savings patterns (e.g. parents with higher earnings delay birth and save more), then our strategy will be confounded. To address this, we can directly examine the correlation between age at first birth and savings. We can also exclude the controls for the age of

[12] In rural areas, there are instances of exemptions for family planning policies in the case of economic hardship. However, this was rare – almost unheard of – in urban areas, where formal social security is typically better provided such that the poor can access other subsidies rather than be exempt from family planning policies.

the household head (and age of the household head squared). If we find that age at first birth is not correlated with savings and that controlling for age of household head does not change the estimated coefficients, then we can infer that our main strategy is not confounded by selection effects. Second, our main results could be confounded by the effects of having older parents who may have different saving patterns because they are at a different point of their life cycle savings pattern. To investigate this possibility, we will conduct a placebo "experiment" and compare savings of households where the household head is 55-60 years of age to households where the head is 61-65 years of age.

Finally, note that we present Newey-West corrected standard errors. Our sample contains 18 cities and 11 birth cohorts (of the eldest child). Therefore, the analysis is unsuitable for large sample correction methods such as clustering, which depend on asymptotic properties. However, we will show in the section on robustness that the significance of our estimates is not sensitive to the error correction method we choose.

5 Data

This study uses the Urban Household Survey (UHS) portion of the larger 2008 *Rural-Urban Migration in China and Indonesia* (RUMiCI) survey for China, a survey collected by one of the authors. The main advantage of our survey over other data is that it allows us to capture both the total number of children ever born and savings rate for a sufficient number of households for empirical analysis. This is not possible with other existing data.[13] In this paper, we only use urban data because family planning policies and access to savings instruments were relatively uniform in urban areas. Also, in there is little sex-selection in urban areas.

The sample frame used in the UHS is the same as the one used in the National Bureau of Statistics (NBS) *Annual Urban Household Income and Expenditure Survey* (UHIES). Sample selection is based on several stratifications at the provincial, city, county, township, and neighborhood community levels. Households are randomly selected within each chosen neighborhood community. The UHIES covers all 31 provinces, whereas the UHS sample households were drawn from nineteen cities in nine of the provinces of the UHIES sample.[14] This sampling frame typically miss migrant laborers. For our study, this is an advantage in that we can assume that urban households we observe in 2008 also had urban status when they had their first child.

[13] The UHIES surveys only asks about children currently residing at home. The China Health and Nutritional Surveys (CHNS) do not have good data on savings and have a very small urban sample. The China Household Income Project (CHIP) have incomplete fertility data and a very small urban sample. Note that our survey asks about children ever born. However, respondents may mistake the question and report only living children. Alternatively, they may be unwilling to recall children that have died. In this case, the birth year of the first child will be measured with error. On average, the reported birth year will be more recent than the actual birth year of the first child ever born. The extent of measurement error will depend on the extent of infant mortality during the early 1970s, which as very low (Banister and Hill, 2004). Therefore, it should not be a big problem for our estimates.

[14] The provinces included in the RUMiCI urban survey are: Shanghai, Guangdong, Jiangsu, Zhejiang, Henan, Anhui, Hubei, Sichuan, and Chongqing. The detailed list of cities can be found at http://rumici.anu.edu.au

The survey was conducted in March and April, 2008. In addition to general information (including fertility) for household members, the questionnaire also included the demographic characteristics, education, and employment situation of other family members who are not coresident with the household head and spouse, including parents, children, and siblings.[15] This allows us to know the total fertility history and characteristics of adult children such as sex, age and marital status. In our study, total fertility is synonomous with the total number of living children.

The information on household income and expenditure from the RUBMiCI in China are directly recorded from the UHIES survey, where the income and expenditure are collected using a diary record. Specifically, households are required to record each item (disaggregated for hundreds of product categories) purchased and income received for each day for a full year (in our case is for the year 2007). Enumerators visit sample households once or twice each month to review the records, assist the household with questions, and to take away the household records for data entry and the aggregation of the annual data at the local Statistical Bureau Office.

Our income and expenditure data and the sampling frame will share several of the problems that exist in the UHIES, which has been thoroughly discussed in by past studies such as Han, Wailes, and Cramer (1995), Ravallion and Chen (1999) and Gibson, Huang, and Rozelle (2003). According to these studies, the quality of the household surveys are in general good and most of the problems are confined to rural surveys. However, there are problems in the urban surveys that could affect studies of savings. First, the indicators used for consumption and expenditure lack consistency over time (e.g. the categories for durable consumption changed quite dramatically during two decades of rapid economic growth). Second, the urban surveys do not fully account for food consumption because they do not account for meals consumed away from home, although this is accounted for in expenditures for food. Finally, the onerous task of recording a daily diary of income, consumption and expenditure makes it difficult to recruit certain households. The first problem should not affect our study as we only use one cross-section and focus on urban residents. The third problem could cause us to underestimate consumption. We address this by using data on expenditures, which have been shown by the studies we cite above to be more accurate for urban household surveys. There is little we can do to directly address the last problem except to keep it in mind when considering the external validity of our results. According to interviews with NBS statisticians and a detailed examination of income and expenditure distributions conducted by researchers in study of the income distribution and income taxation using the UHIES data, researchers concluded that the households that refuse to participate are typically the poorest and the richest households (Piketty and Qian, 2009). This makes it difficult to use the UHIES to study the extreme tails of the income distribution, but should not affect our study, which focuses on the mean household.

Another important fact to keep in mind when assessing the external validity of our estimates is that China is the only country in the world that uses such comprehensive twelve month expenditure records.[16] Gibson, Huang and Rozelle (2003) found that ex-

[15] The questionnaires are available from http://rumici.anu.edu.au

[16] Surveys in many other countries observe households for a week, a fortnight, or a month, and estimates

trapolating annual totals from expenditures using some months of the year caused sharp decreases in expenditure measures.[17] This means that measures of household savings in China – the difference between income and expenditure – are not directly comparable to measures of household savings from other countries. (Unlike expenditures, income data is collected in a similar fashion as many other countries). In other words, if the same statistical methods employed in most of the world were also employed in China, then Chinese savings rates will be higher than what they are in our data (or any savings data that is based off of the UHIES). This error in measurement of what will be the dependent variable in our analysis should not affect our estimates. However, it needs to be taken into account when comparing mean savings rates in China with other countries. Specifically, one would need to know the correlation between household's expenditures with different months.

In our data, total household income is the sum of wage income (e.g. wages and subsidies from other labor income), operational income, property income, transfer income (e.g. pension and retirement allowances, social welfare benefits). Total expenditure is the sum of consumption expenditure (e.g. food; clothing; housing; family equipment; service; health; transpirations and communication; education; cultural and entertainment; other commodity and services), operational expenditure, property expenditure, transfer expenditure and social security expenditure (e.g. individually paid pension fund, individually paid public housing fund individually paid health care fund, individually paid unemployment fund, and other social security).[18]

Our main outcome measure is savings, which we measure in the standard manner as the difference between total income and total expenditure, $S = Y - E$. We also examine savings ratio, which is simply savings divided by total income, S/Y.

The data is organized to be household level birth cohort panel according to the birth year of the first child. The empirical analysis focuses on households that had their first child five years before or after the policy shift in 1972, i.e., 1967-77. This restriction excludes households that have no children. This makes little difference to our data as almost all couples in the comparable age range are married and have at least one child. The length of the window is arbitrarily chosen. Our sample will also exclude households headed by individuals over 65 years of age because they are likely to be on a different point of the life-cycle. Specifically, they may have already begun to dissave and therefore will not be comparable to younger households. The final sample contains 489 households

of income and consumption from these periods are annualised by multiplying by 52, 26, or 12. The length of the recall period typically depends on the category of consumption, with long reference periods used for costly and/or infrequently consumed items and short reference periods for frequently consumed and minor items that would be easily forgotten (ILO, 1994).

[17] They also found that such extrapolations sharply increased measures of inequality. This may be due to the fact that by using data from only a few months, random shocks to expenditures are given too much weight. Also, see Deaton (1997) for a detailed discussion of the statistical tradeoffs of different data collection methods.

[18] Food expenditure is the sum of expenditure on the following categories: grain, wheat, and rice coarse grains; dried vegetables pork, beef, and mutton; edible vegetable oil, fresh vegetables, dried vegetables, poultry, meat, eggs, fish; sugar, cigarettes, liquor, fruit, wine, beer, fresh melons and fruits cake; and milk.

in eighteen cities.

Table 3 shows the descriptive statistics. Households in our sample on average have total incomes of 52,067 RMB and expenditures of 32,681 RMB. Savings are on average 19,386 RMB, which is 31% of income. The average household has approximately one child, 49.8% of which are male, an important fact to keep in mind because it supports our claim that there is no sex-selective abortion in our sample. On average, parents had their first child in 1973 and their youngest child in 1976. This means that when the survey was conducted in 2008, households in our sample on average had children age 32-35 year of age. Our sample contains households headed by individuals 51-65 years of age. On average, household heads are approximately 61 years of age and have approximately ten years of education (e.g., one year of high school education) and approximately 42% of our sample is headed by women. In the case of our survey, this does not necessarily mean that there was no male spouse for the female household head. Sometimes, it simply indicates that the survey respondent was the adult female of the household. To be cautious and to avoid the potentially confounding effects from having a female household head, we will control for this in our regressions.

To observe the change in fertility over time, we estimate the correlation between the year of birth of the eldest child and the total number of children born to a household by regressing the latter on the dummy variables for the former (with no additional controls). 1967 is the omitted reference group. The coefficients and standard errors are shown in column (1) of Appendix Table A1. They show that fertility declines for households that have their first child in 1972 or afterwards and that the birth year dummies for those years are jointly significantly different from zero. The pattern is illustrated in Figure 2, which plots the coefficients and their 95% confidence intervals. It shows that there is a trend break in the coefficients beginning in 1972. This is consistent with our interpreting 1972 as the effective date of when population control policies began in earnest. We can repeat this exercise for savings. The coefficients are shown in column (2) of Appendix Table 2. They and their 95% confidence intervals are plotted in Figure 3. We observe an increase in savings levels for households that had their first child after 1972. This descriptive evidence suggests that parents that had their first child after the shift in family planning policies had fewer children and higher savings. In the next section, we present the causal evidence.

6 Empirical Results

6.1 Main Results

Our empirical analysis examines three outcomes: the level of savings, savings rate (savings as a fraction of income) and income. In a developed economy, if parents with fewer children and have higher savings rates for many years will, over time, earn returns from investing their savings, which means that they should have higher income in the long run. In our context, we expect the effect of saving more on income to be less than most other contexts because Chinese urban households have historically had very limited access to investment opportunities. However, the sign of the effect of fertility on income can still serve as a

consistency check on the mechanisms driving the results.

Table 4 shows the OLS estimates of the correlation between household size and savings from equation (1). The coefficient for the number of children reflects the correlation for households that have a daughter as the eldest child. The sum of this coefficient and the coefficient of the interaction term of the number of children and whether the first is a son reflects the correlation for households that have a son as the first child. This sum and its p-value is presented towards the bottom of the table. Panel A columns (1)-(8) shows that the number of children a household has is uncorrelated with savings or savings rates when the eldest child is a daughter. The estimates for the coefficient of the number of children are small in magnitude and statistically insignificant. Columns (3)-(5) of panel A show that the number of children and savings levels are positively correlated when the first child is a son. The joint estimate is statistically significant at the 5% and 1% levels. The estimates in panel C show that the correlation between household size and income is not statistically significant for households that have a daughter as the eldest child. But for households that have a son as the eldest, income and the number of children are positively correlated and statistically significant at the 1% level.

Next, we estimate the first stage effects of the effect of having the first child born after 1972 on family size from equation (2). The estimates are shown in Table 5. Column (1) shows the correlation while only controlling for city fixed effects. Columns (2)-(4) add controls for household characteristics. Columns (1)-(4) show that on average, parents who had their first child after 1972 have approximately one less child than those who had their first child before. In column (5), we estimate the baseline equations where we add controls for whether the first child is a son and the interaction of that term with whether the first child was born in 1972 or after. The coefficient for whether a child was born in 1972 or afterwards reflects the effect on households that have daughters for a first child. The sum of this coefficient and the interaction of whether the first child is a son reflects the effect on households that have a son as a first child. This sum and its p-value is shown at the bottom of the table. The baseline estimates show that parents who have their first child after 1972 will have approximately one less child than those who had their first before; and parents who first have son as the first child are slightly less likely to have a second child. This is consistent with the observation that parents with son preference use a stopping rule in that they are more likely to stop having children if they have a son. These estimate for whether the first child is born in 1972 or after is statistically significant at the 1% level in all specifications. The estimate for whether the first child is a male and its interaction with whether a first child is born in 1972 or after is statistically significant at the 1% levels. The estimates of the second first stage equation, equation (3), are presented in column (6). In both columns (5)-(6), we see that the F-Statistic for the joint significance of the two instruments are large. The Kleibergen-Papp F statistic for the first stage is similarly large and well above the Stock-Yogo (2005) critical value for weak instruments. Therefore, we conclude that our instruments are strong and significant.

Recall that to avoid concerns related to sex-selection of children, we focus our sample to urban households that have their first child before sex selective abortion was available and approximately 49.8% of the children in this sample are male, which supports our claim. However, one could still be concerned that the sex-ratio became more boy biased

after fertility control policies began in 1972. To address this, we directly examine the effect of having a first child in 1972 or after on the fraction of sons born conditional on all of the baseline controls. In column (7), we see that there is no effect. The magnitude of the estimate is near zero and it is statistically insignificant. Moreover, the sign is negative.

To estimate the reduced form effect of having the first child born after 1972, we repeat the first stage estimation with the second stage outcome variables as dependent variable. The estimates are shown in Table 6. Column (3) of panel A shows the estimates from the baseline specification. It shows that for households that have a daughter as a first child, having the first child in 1972 or after increases savings by 12,397 RMB. The estimate is statistically significant at the 1% level. The estimate for the interaction term shows that the effect is significantly smaller for households that have a son for a first child. It is also statistically significant at the 1% level. The sum of the coefficients at the bottom of panel A show that having a first child after 1972 does not have any significant effects on the savings of households that have a son as the first child. The estimates on savings rates in panel B show a similar pattern. The baseline estimation in column (7) shows that having the first child after the policy shift causes savings rates to be 10 percentage-points higher. However, the effect is greatly reduced in magnitude if the eldest child is a son rather than a daughter. Both the estimates for the dummy for having the eldest child in 1972 or after and its interaction with having a son as the first child are statistically significant at the 1% levels.

In columns (5) and (8), we show the estimated effects on savings and savings rates where we drop the controls for the age of the household head and its squared term. The estimated effects are very similar to the baseline estimates, which suggests that our baseline results are not driven by the selection of the age of the household head.

In panel C, we present the reduced form estimates on income levels. The signs are similar to the signs of the estimate effects on savings. However, the estimates are not statistically significant different from zero.

Finally, we present the 2SLS in Table 7. The baseline specification in column (4) shows that having one more child decreases savings by 14,234 RMB if the first child is a daughter. The estimated interaction term shows that the magnitude of the reduction in saving is 12,234 RMB less for those that have a son as the first child. The coefficients for the number of children and its interaction with a dummy variable of whether the first child is a son is statistically significant at the 1% level. Interestingly, these results also say that all of the effects of fertility on savings are driven by households who have daughters. This is consistent with the descriptive statistics that we presented in section 2, which show that sons provide much more old age support to parents than daughters. In panel B, the estimates on savings rates produce a similar pattern. The baseline estimate in column (7) show that having one more child reduces savings rates by 11.2 percentage-points. This estimate is statistically significant at the 1% level. The interaction terms show that if the first child is a son rather than a daughter, this effect will be much reduced in magnitude. However, the interaction term is not significant.

As with the reduced form estimates, the results in columns (4) and (8) in panel A from when we drop the controls for the age of the household head and its squared term, are very similar to the baseline results. This suggests that our estimates are not driven

by the selection of the age of household heads.

In panel C, we present the 2SLS estimates on the effect of household size on income. The estimate for the number of children is negative, which is consistent with the fact that savings generate positive returns over time, which is reflected in the income measure. However, it is not significantly different from zero. As we mentioned earlier, this is most likely due to the fact that Chinese households had very few investment options historically.

Besides the main results, there are two interesting points to note. First, a comparison of the OLS estimates in Table 4 and the 2SLS estimates in Table 7 show that the latter are significantly larger in magnitude and have the opposite sign. This is consistent with the belief that fertility is an endogenous decision such that it is the richer households that can afford to have more children and these households will also have higher levels of savings and savings rates as long as there are diminishing returns to consumption. Second, note that the coefficient for whether a first child is a son is very large, negative and statistically significant. This suggests that parents who have sons save much less than those that have daughters. Together with the fact that sons seem to provide twice as much support for retired parents as daughters, this is consistent with our hypothesis that parents view their children as a form of savings.

6.2 Robustness

This section tests the sensitivity of our results. First, we check that our empirical strategy is not confounded by the fact that households that had their first child after 1972 will also have younger households heads and younger children by conducting two placebo experiments. To see that selection in the age of the household heads is unlikely to play a major role in driving our results, note that the main 2SLS estimate change little when we drop the control for the age of the households head (Table 7 panel A columns (5) and (10)). To investigate whether the results are driven by the age of the children, we conduct a placebo experiment where we compare households that are headed by individuals who are younger than sixty years of age to those that are headed by individuals older than sixty years of age. For this exercise, we restrict our sample to households where the household head is age 55-65. If our main estimates are driven by the fact that households that had their first child after 1972 are younger rather than by the reduction in fertility caused by family planning policies, then we should observe similar results from this comparison. The first stage estimates are shown in Table 8 Column (1). It shows that the first stage estimate is smaller in magnitude but still statistically significant. Households with younger household heads typically have 0.22 fewer children on average. The estimate is statistically significant at the 5% level. However, Columns (3) and (5) shows that the number of children have no effect on savings and saving rate when it is instrumented with these alternative instruments. Therefore we conclude that our main estimates are not driven by the age differences of parents or children between the control and treatment groups.

Next, we investigate the possibility that our results are confounded by the relationship between the age at first birth and savings. For example, if, controlling for age, parents with higher unobserved human capital have children later in life and save more, then our main strategy, which exploits variation from the birth year of the first born child will be

confounded. In the main analysis, we have already addressed this by examining whether the results differ when we do not control for the age of the parent. We found that dropping these controls made little difference, which suggests that are estimates are not biased by selection. Here, we address this potential problem by directly estimating the effect of the age at first birth for the household head and household savings. In an estimation where we control for the same controls as in the main estimates, we find that the age at first birth is uncorrelated with savings. The correlation coefficients with savings and savings rates are 478 (se 567) and 0.0043 (se 0.0055). Therefore, we conclude that our estimates are unlikely to be driven by selection bias.

Finally, we check the robustness of our standard errors to different correction methods. The main results presents Newey-West robust standard errors. Alternative ways to correct for standard errors are to cluster them at the city or birth year level (for the first child) to correct for autocorrelated shocks on savings within cities or correlated shocks across cities within birth years. However, since we only have eighteen cities and eleven birth years, such correction may induce small sample bias. Table 9 shows that this is indeed the case when we cluster at the birth year level, which produce standard errors that are very similar to the uncorrected ones, and that are smaller than the Newey-West robust standard errors and those clustered at the city level. Table 9 also shows that the Newey-West robust standard errors we present in the main results are slightly smaller in magnitude than those clustered at the city level, but that the latter is still statistically significant at the 1% level. Therefore, we conclude that our estimates are robust to different error correction methods.

In addition to these results, we also control for city-specific time trends to address the possibility that there are differential trends across cities which could confound our results. Adding this rigorous set of controls increases the magnitudes of our estimates. But they are not precisely estimated. Therefore, for the sake of brevity, we do not present them in the paper.[19]

6.3 Additional Results

The main results show that savings levels vary with the sex of the first child; the effect of reducing fertility on increased savings is mainly due to those who have a daughter as a first child. Here, we investigate the extent for which the difference in savings behavior between those who have sons versus daughters is due to the fact that parents expect daughters to provide less support when they are elderly. One way to explore this is to the divide the sample according to cities where adult female wages as a fraction of adult male wages (as reported in our survey) is above or below the sample median (0.83). If parents forecast the future earnings of daughters relative to sons based on the current relative female wages of their local labor market, then the difference in savings behavior between those that have daughters versus sons will be larger in cities where the current gender wage gap is large. In Table 10, columns (1) and (2), we show the estimate for the sample of cities that have below and above median gender wage gaps. In each column, the estimated effect of having an additional child and its interaction with whether the first

[19] They are available upon request.

is a son have the same signs as the main results, and they are all statistically significant at the 5% and 1% levels. A comparison of the estimated interaction effects show that the difference in savings behavior between parents who have daughters versus those that have sons is indeed much larger in cities with high gender wage gaps. This suggests that the gender wage gap play an important role in determining the child-sex-specific savings behavior. It is, more generally, additional support for the theory that parents savings behavior is partly determined by the expected returns from having children.

Next, we investigate if the effect of fertility on savings varies according to the marital status of children. Marital status can matter for many reasons. For example, parents with unmarried children may save more for the purposes of providing a dowry or bride price. In urban China, there is no tradition of providing dowries of bride prices when a child marries. However, one may think that parents may still plan to provide financial support for children more informally (e.g., wedding, housing). In columns (3) and (4), we compare the estimated effects for parents whose eldest child is single versus married. The estimates are not always statistically significant due to the small sample size. However, they suggest that there is little difference in the relationship between fertility and savings rates between these two groups. In columns (5) and (6), we examine the effect of fertility on savings for households that have at least one single child and household for which all children are married. As with the earlier estimates, they show that the marital status of children does not affect the relationship between fertility and savings rates.

7 A Simple Life-Cycle Model of Savings

We imagine a very simple life-cycle: People live for T years. There are five stages of life. At age S a couple gets married, starts working, has children and begins to earn incomes and support their parents. At $2S$, their children, now S years old, also start working and supporting their parents. At age $T - S > 2S$, their parents die. At age R such that $T > R > T - S$, they retire. At age T, they die.

Assume that every male child gives a fraction μ^m of his income to his parent once he starts working, and every female child gives a fraction μ^f of her income to her parents. And suppose wages of a man who started working at time s at time t is $\widetilde{w}(s,t)$, and that of a woman is $\lambda\widetilde{w}(t,s)$. Since both members of a couple are the same age, their total transfer to their parents is $(\mu^m + \lambda\mu^f)\widetilde{w}(t,s)$. It is convenient to express everything in terms of total family income so we define the average transfer rate to parents as $\mu_p = \frac{(\mu^m+\lambda\mu^f)\widetilde{w}(t,s)}{w(t,s)}$, where $w(t,s) = (1+\lambda)\widetilde{w}(t,s)$ is total family income. The amount the couple gets from their children depends on their number and gender composition: one daughter born in year s at time t will give an amount $\mu_c(G)w(t,s) = \lambda\mu^f\widetilde{w}(t,s)$; one son will give $\mu_c(B)w(t,s) = \mu^m\widetilde{w}(t,s)$; two daughters will give $\mu_c(GG)w(t,s) = 2\lambda\mu^f\widetilde{w}(t,s)$; two sons $\mu_c(BB)w(t,s) = 2\mu^m\widetilde{w}(t,s)$; one son and one daughter $\mu_c(GB)w(t,s) = (\mu^m+\lambda\mu^f)\widetilde{w}(t,s)$. Note that this means that we can never separate λ and μ^f. Therefore we set $\lambda = 1$ and hence $w(t,s) = 2\widetilde{w}(t,s)$. It will be convenient to also suppress the arguments of $\mu_c(\cdot,\cdot)$ unless specifically needed.

Asset accumulation by the couple in each period (except when they are a child) can

be characterized as the following (here time is being measured from when this generation was born).

$$a_{t+1} = a_t(1+r) + y_t - c_t$$

where

$$\begin{aligned} y_t &= 2w(S,t)(1-\mu_p) \qquad 2S > t \geq S \\ &= 2w(S,t)(1-\mu_p) + \mu_c w(2S,t).....T - S > t \geq 2S \\ &= w(S,t) + \mu_c w(2S,t).....R > t \geq T - S \\ &= \mu_c w(2S,t).....R > t \geq T - S \end{aligned}$$

We will assume that wages are subject to both an experience effect and a cohort effect:

$$w(s,t) = w_0(1+\beta)^s(1+\alpha)^{t-s} \tag{4}$$

where w_0 is the starting wage of the cohort that entered the labor market at time zero, $w_0(1+\beta)^s$ is the starting wage of the cohort that entered the labor market at time s, and $(1+\alpha)^{t-s}$ is the experience premium for having been in the labor market for $t-s$ years.

Assume that at time the couple maximizes a utility function of the form

$$\sum_{t=S}^{T} \delta^t u(c_t)$$

subject to the above asset accumulation constraints, plus a terminal condition $a_T \geq 0$. Pending discussion of this condition, assume that $\delta(r+1) = 1$. Then the optimal consumption path without credit constraints is to consume a constant amount. However assume that there are credit constraints i.e. $a_t \geq 0$. Then we have the following result.

Result: There always exists at least one t_1 $(R > t_1 \geq S)$, and c pair such that

$$c\sum_{t \geq t_1}^{T} \frac{1}{(1+r)^t} = \sum_{t \geq t_1}^{T} \frac{y_t}{(1+r)^t}.$$

Among such (t_1, c) pairs let t_1^*, c^* be the one among with the lowest value of t_1. If it satisfies the property

$$\mu_c w_\beta (1+\beta)^{T-2S} \leq c^*, \tag{5}$$

then the optimal consumption path is

$$\begin{aligned} c_t &= y_t \text{ for all } S \leq t < t_1^* \\ c_t &= c^* \text{ for all } T > t \geq t_1^*. \end{aligned}$$

The intuition for this result is simple: The consumer is trying to minimize variation in consumption. Since income goes up first and then (for the most part) comes down,

the best the consumer can do is to wait until their income is high enough and then start saving for the future when their income is lower. The condition 5 is there to make sure that the transfers from children don't grow so fast that income actually goes up at some point after retirement. To explain the observed savings rates, we will focus mainly on the case where the propensities to transfer (μ_c's) are relatively close to zero. Therefore we can usually satisfy condition 5.

7.1 Discussion of model

The model we introduced above is very special along a number of dimensions. In particular children do not have any direct costs associated with them, which is obviously unrealistic. However this typically pushes towards depressing savings rates since couples have to spend on children relatively early in life, and as we remarked in the introduction our challenge is to get the model to generate enough savings. We also assumed that $\delta(1+r) = 1$, which, given that real interest rates in China are around 1% makes the agents very patient indeed. However patience also goes in the direction of making people more inclined to save. A more problematic assumption is that there is no uncertainty, and hence no precautionary motive to save. We will try to address this concern in subsequent versions of this paper.

7.2 Assumptions about parameters

We have already said the average real interest rate on bank deposits in China over the period is about 1%. T the life-span is set to be 80 years, which is higher than the national life expectancy (73 years).[20] However what we are interested in is the life expectancy when they start earning, which is obviously higher, and our data is for urban Chinese who live longer.

In terms of work life we assume that people join the labor force at twenty-five and retire at sixty. Both numbers are probably too conservative–people probably join earlier and while official retirement age for men is sixty and for women is fifty-five, many people continue to work until at least sixty-five. However a longer work life relative to the period of retirement will tend to depress savings.

Finally to estimate cohort and experience effects on earnings, we use household level data from UHIES from 1987 till 2005 to estimate the earnings model given above in 4. In other words we run a regression of the form

$$\log w(s,t) = \log w_0 + s\log(1+\beta) + (t-s)\log(1+\alpha) + \varepsilon_{st}.$$

We estimate this separately for the 1987-1992 period, the 1993-1996 period and the 1997-2005 period, because of the enormous variation in the inflation rates across these sub-periods (the average inflation rate goes from 9% in the first sub-period, to 16% in the second to 1% in the third). The results show large variation in both the per year experience and the cohort effects:

[20] World Bank (2008).

	1987-92	1993-96	1997-2005	Average
Experience effect	6.5%	12%	8%	8.5%
Cohort effect	3.5%	9%	6.5%	6%

In our main results we will use the average for the entire period but we will also look at what happens if focus on the rates from specific sub-periods.

In terms of demographics about 25% of our population has more than two children the present computations of the model assumes that people either have one or two children. As we argue above, for this population it is reasonable to assume that there is no sex selective abortions, so that the probability of a male child in each birth is 50%.

Finally, our descriptive statistics, reported above, suggests that boys give about twice as much as girls to their parents. We will make use of this fact when we select values for μ^m and μ^f.

7.3 Results

The model generates a predicted value of the average savings rate. Our empirical analysis of savings is based on people who had their first child between 1967 and 1977. They are all between 51 and 65 years of age. Though a lot of them above the retirement age a lot of them are still working. Assuming that they have retired will mechanically depress the savings rate. We therefore compute saving rates for the oldest 15 year cohort that would be working in our model–those between 45 and 60–who also have the highest average savings rate. Our savings rate is therefore define to be

$$\int_{t_1^*}^{54} \frac{y_\tau(1-\mu_p)-c^*}{15y_\tau}d\tau + \int_{55}^{60} \frac{y_\tau - c^*}{15y_\tau}d\tau$$

where t_1^* is the year when the couple starts to save and c^* is how much they consume every year once they start to save (see the above result for formal definitions).

The table below shows the implications of a sample alternative parametric assumptions. The first two columns represent the values of the wage growth rates, α and β. The next two give alternative propensities to transfer for male (μ^m) and female (μ^f) children. The next five columns are our results: The first(B) gives the savings rate of a couple that has one boy; the next(G) one with one girl; the next(BB) one with two boys; and so on. The tenth column shows the average savings rate generated by the model. The last three columns give the predicted values for our regression coefficients–respectively the effect of the number of children (#), the effect of having a male first child (son), and the interaction of the two (#*son).

α	β	μ^m	μ^f	B	G	BB	BG	GG	Av	#	son	#*son
0.085	0.06	0.133	0.066	0.29	0.34	0.21	0.25	0.29	27%	-0.07	-0.06	0.01
0.085	0.06	0.2	0.1	0.24	0.3	0.15	0.19	0.24	21%	-.085	-0.075	0.015
0.085	0.06	0.2	0	0.25	0.38	.019	0.25	0.38	28%	-.065	-0.165	.035
0.065	0.035	0.133	0.066	0.39	0.41	0.32	0.35.	0.39	36%	-0.04	-0.015	-.005
0.065	0.035	0.2	0.1	0.29	0.36	0.21	0.26	0.29	27%	-0.85	-0.1	0.03
0.065	0.035	0.2	0	0.3	0.43	0.23	0.3	0.43	32%	-0.65	-0.16	0.03

The first three rows assume that the savers are making their plans taking as given the actual numbers on the experience effect and the cohort effect ($\alpha = 0.85, \beta = 0.06$). What varies across them is our assumptions about the propensity to transfer. It is evident that if parents expect their children to transfer more on average, they will save less. Thus a couple that has only daughters in a world where daughters do not transfer at all will save 38% of their income (row four column six, nine)), whereas a couple with two boys where each boy will transfer 20% of his income should only save 19% (row four, column seven).

The second three rows assume that ($\alpha = 0.65, \beta = 0.035$). These are the numbers corresponding to the 1987-1992 period. This is when most of sample were just entering the labor market, and the assumption is that this experience frames their view of the future.

The cases considered in the table are designed to illustrate the key features of our model. First, even with relatively low propensities to transfer (the maximum is 0.2), the models with the actual numbers for wage growth generate less savings than we observe in the data: the average savings rate in the data is 31% while in the model the maximum is 28%. This is reached when $\mu^m = 0.2$ and $\mu^f = 0$, which assumes that contrary to the data, girls give nothing to their parents (or are assumed to give nothing by their parents). If we make the assumption, more consistent with the data that girls give roughly half of what boys give, we can generate a savings rate of 27% on average by assigning very low values to the propensity to transfer (0.133 for boys and 0.066 for girls), which is perhaps not so different from 31%. However now we encounter a different problem: when transfers are relatively unimportant, the effect of an extra child or an extra boy on the savings rate tend to be small. The predicted values from our model for this case (μ^m =0.133 and μ^f =0.066 for girls) of the coefficient on an extra child is negative 7 percentage points, that of the first being male is negative 6 percentage points and the interaction of the two is positive 1 percentage point. In the data the corresponding effects are negative 11 percentage points, negative 16% percentage points and positive 7.7 percentage points. In other words the predicted effects from the model are dwarfed by what we observe in the data. We get significantly closer to observed coefficients when we assume that $\mu^f = 0$, and $\mu^m = 0.2$, but the predicted coefficients on the number of children is still quite a bit smaller.

Moving to a lower expected wage growth rate ($\alpha = 0.065, \beta = 0.035$) helps match the observed savings rate better, but the basic trade-off between a lower propensity to transfer which yields higher savings rates and a higher propensity to transfer which permits the coefficients to be larger, remains. None of the cases we have computed so far get very close to matching both the savings rates and the coefficients on the number and gender of children. We come closest when we set $\mu^f = 0$, but this seems counterfactual.

8 Conclusion

The importance of children in savings decisions has long been recognized by economists. China, through its family planning policies, changed demographics in a way that very few countries have been able to do. This provides a unique laboratory to study the effect of changes in the demographic structure on savings and wealth. In this study, we document

that children provide a substantial amount of support for elderly parents and that sons provide more support than daughters. We then show empirical estimates that support the predictions of a simple life-cycle model that take these facts into account and assumes that parents are credit constrained. We find that the exogenous reduction in fertility due to family planning policy causes a significant increase in household savings, and that all of the increase is driven by parents that have a daughter as a first child. Since during this time, family planning policy reduced the number of children from approximately two to one, this means that all of the effects of the fertility reduction are driven by parents that have a daughter as their only child. These results provide strong support that the life-cycle model and changes in demographic structure play an important role in explaining household savings rates in China.

In comparing our results to the macro-evidence, it is important to note that those studies typically capture the savings behavior of the entire population where as our study focuses on a sample of urban individuals that range from 51 to 65 years of age, and almost all of whom are still working.

It is also important to note that while the results of this study suggest that the LCH and the change in demographic structure are important contributors in explaining household savings in China, they do not rule out or compete with the contribution of other factors. In fact, the calibration exercise using the empirical estimates illustrates the limitations of a simple life-cycle model in showing that it cannot perfectly fit the data without additional assumptions. This is an interesting avenue for future research.

References

[1] Banister, Judith and Hill, Kenneth "Mortality in China 1964-2000" Population Studies, Vol 58 (1), pp.55-75.

[2] Blanchard, Olivier J. and Francesco Giavazzi, 2005. "Rebalancing Growth in China: A Three-Handed Approach." MIT Department of Economics Working Paper No. 05-32.

[3] Cai, Fang, Giles, John and Meng, Xin (2006) "How well do children insure parnets against low retirement income? An analysis using survey data from urban China" Michigan State University Working Paper.

[4] Chamon, Marcos, and Eswar Prasad, 2008, "Why Are Saving Rates of Urban Households in China Rising?" American Economic Journal – Macroeconomics, forthcoming.

[5] Deaton, Angus. (1997) The Analysis of Household Surveys: A Microeconometric Approach to Development Policy Johns Hopkins, Baltimore.

[6] Diamond, Peter, (2004) "Report on social security reform in China", China Economic Research and Advisory Programme, November.

[7] Fang, Cheng, Eric Wailes and Gail Cramer (1995) "China's Rural and Urban Household Survey Data: Collection, Availability, and Problems" Center for Agriculture and Rural Development Working Paper 98-WP202.

[8] Gibson, John, Jikun Huang and Scott Rozelle (2003) "Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey," Review of Income and Wealth, Blackwell Publishing, vol. 49(1), pages 53-68, 03.

[9] He, Xinhua and Yongfu Cao, 2007. "Understanding High Saving Rate in China," China & World Economy 15 (1): 1 – 13.

[10] Holland, Tom. 2002. "China's Pension Puzzle," Far East. Econ. Rev. 165:21, pp. 42–43.

[11] Horioka, C.Y. 1990. "Why Is Japan's Private Saving Rate So High?" in Developments in Japanese Economics. R. Sato and T. Negishi, eds. Tokyo: Academic Press.

[12] Horioka, Charles Y., Junmin Wan, "The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data," NBER Working Paper No. 12723, November 2006.

[13] ILO (1994), "Household Income and Expenditure Surveys," Year Book of Labour Statistics, Sources and Methods, Volume 6, International Labour Office, Geneva.

[14] Kuijs, Louis. 2006, "How Would China's Saving and Investment Evolve?" World Bank Policy Research Working Paper 3958. Washington: World Bank.

[15] Modigliani, Franco. 1970. "The Life Cycle Hypothesis of Saving and Intercountry Differences in the Saving Ratio," in Induction, Growth and Trade, Essays in Honor of Sir Roy Harrod. W.A. Elits, M.F. Scott, and J.N. Wolfe, eds. Oxford.

[16] ———. 1989. "Life Cycle, Individual Thrift and the Wealth of Nations," in The Collected Papers of Franco Modigliani. Simon Johnson, ed. Cambridge, MA and London: MIT Press.

[17] ———. 1990. "Recent Declines in the Saving Rates: A Life Cycle Perspective," Revista di Politica Econ., Dec., pp. 5–42.

[18] ———. 1996. "The Life Cycle Hypothesis of Saving, the Demand for Wealth and the Supply of Capital," Social Research.

[19] Modigliani, Franco, and Shi Larry Cao, 2004, "The Chinese Saving Puzzle and the Life Cycle Hypothesis," Journal of Economic Literature, 42: 145-170.

[20] Qi, Tianxiang, 2000, "A study of Chinese household saving in a transition period: Some discussion on the relationship of uncertainty and household saving," Jingji yanjiu (Economic Research Journal), Vol.35, No. 9, pp. 25–33.(In Chinese)

[21] Qian, Nancy (2009) "Quantity-Quality: The One Child Disadvantage on School Enrollment" Yale Working Paper.

[22] Qian, Yingyi, 1998, "Urban and Rural Household Saving in China," International Monetary Fund Staff Papers, 35(4): 592-627.

[23] Ravallion, M., and Chen, S. (1999), "When Economic Reform is Faster Than Statistical Reform: Measuring and Explaining Income Inequality in Rural China," Oxford Bulletin of Economics and Statistics 61, 75-102.

[24] Shi, Jianhuai, Haiting Zhu, 2004, "Precautionary saving of Chinese urban households and its intensity: 1999–2003," China Center for Economic Research Working Paper No. C2004013, Beijing: Beijing University.

[25] Wei, Shangjin and Zhang, Xiaobo "The competitive saving motive: Evidence from rising sex ratios and savings rates in China" NBER Workin gPaper 15093

Table 1: Cohabitation of Elderly Parents and Adult Children

	The Fraction of Parents Living with Adult Children							
	Number of Living Children							
Eldest Child	1	2	3	4	5	6	7	8
	A. All Household							
All	0.56	0.57	0.54	0.63	0.63	0.65	0.71	0.71
Son	0.64	0.65	0.64	0.74	0.74	0.93	1.00	0.67
Daughter	0.33	0.41	0.37	0.26	0.48	0.13	0.20	1.00
	B. One or more Parent in Poor Health							
All	0.65	0.56	0.46	0.60	0.57	0.82	0.75	0.75
Son	0.72	0.65	0.58	0.65	0.58	1.00	1.00	0.67
Daughter	0.40	0.33	0.28	0.33	0.55	0.33	0.00	1.00

Source: CHRLS. Sample is restricted to households for which at least one child was born during 1972-84. Sample is not restricted based on residence or hukou. There are 939 household level observations.

Notes: 31% of the sample report that at least one parent is in poor health. 16.3% report that at least one parent is in very poor health.

Table 2: Transfers from Children to Parents

	Eldest Child	
Transfers from the Past Year	Son	Daughter
Total Number of Transfers Received	590	299
# of Transfers from Eldest Child	235	117
Fraction of Transfers that were from Eldest Child	0.40	0.39
Number of "Regular" Transfers from Eldest Child	43	18
Fraction of Transfers from Eldest Child that were "Regular" Transfers	0.18	0.15
Average Amount of "Regular" Transfer (RMB)	4930.23	2894.12

Source: CHRLS. Sample is restricted to households for which at least one child was born during 1972-84. Sample is not restricted based on residence or hukou. There are 939 household level observations.

Table 3: Means

Variable	Obs	Mean	Std. Dev.
Income	489	52066.61	37070.81
Expenditure	489	32681.09	26117.09
Savings (Income-Expenditure)	489	19385.52	25361.96
Savings Rate (Savings/Income)	489	0.31	0.29
# Kids	489	2.04	0.84
Fraction male	489	0.498	0.38
Year of Birth of First Child	489	1973.02	2.94
Year of Birth of the Last Child	489	1976.36	4.23
Age of Household Head	489	60.66	3.03
Years of Education for the HH Head	489	9.73	1.49
Fraction of Female HH Heads	489	0.42	0.49

Table 4: OLS Estimates of the Correlation between Fertility and Savings

	Dependent Variables									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
				Baseline					Baseline	
	A. Ln Savings					B. Savings/Income				
# Kids	1,855.623	1,233.519	-2,037.420	-2,215.916	-1,893.550	0.016	0.003	-0.012	-0.012	-0.008
	(1,513.150)	(1,411.208)	(1,767.015)	(1,675.062)	(1,598.970)	(0.010)	(0.015)	(0.021)	(0.018)	(0.016)
# Kids x 1st is a Son			7,701.003	7,698.483	7,682.631			0.034	0.034	0.033
			(2,593.695)	(2,652.013)	(2,665.556)			(0.033)	(0.026)	(0.026)
1st is a Son		2,981.131	-12,611.191	-12,554.431	-12,493.437		0.001	-0.067	-0.067	-0.066
		(2,210.635)	(4,948.349)	(5,781.084)	(5,782.556)		(0.026)	(0.069)	(0.060)	(0.060)
Controls										
HH Head Age	N	Y	Y	Y	N	N	Y	Y	Y	N
HH Head Age Squared	N	Y	Y	Y	N	N	Y	Y	Y	N
HH Head Years of Edu	N	Y	Y	Y	Y	N	Y	Y	Y	Y
HH Head Years of Edu Squares	N	Y	Y	Y	Y	N	Y	Y	Y	Y
HH Head Age >55	N	N	Y	Y	Y	N	N	Y	Y	Y
Age of Youngest Child	N	Y	Y	Y	Y	N	Y	Y	Y	Y
Youngest Child Age < 22	N	N	Y	Y	Y	N	N	Y	Y	Y
Mother is HH Head	N	N	N	Y	Y	N	N	N	Y	Y
Observations	489	489	489	489	489	489	489	489	489	489
Joint F: # Kids + # Kids x 1st is a Son			5664	5483	5789			0.0219	0.0220	0.0251
p-value			0.0255	0.0103	0.00696			0.322	0.314	0.240
	C. Income									
# Kids	3,930.063	4,744.146	-256.598	-498.847	-670.027					
	(2,462.045)	(1,966.182)	(2,112.127)	(2,418.754)	(2,331.442)					
# Kids x 1st is a Son			11,773.606	11,770.186	11,723.533					
			(2,875.462)	(3,387.986)	(3,370.910)					
1st is a Son		4,892.051	-18,946.125	-18,869.092	-18,754.929					
		(2,842.397)	(5,529.868)	(7,332.716)	(7,295.897)					
Controls										
HH Head Age	N	Y	Y	Y	N					
HH Head Age Squared	N	Y	Y	Y	N					
HH Head Years of Edu	N	Y	Y	Y	Y					
HH Head Years of Edu Squares	N	Y	Y	Y	Y					
HH Head Age >55	N	N	Y	Y	Y					
Age of Youngest Child	N	Y	Y	Y	Y					
Youngest Child Age < 22	N	N	Y	Y	Y					
Mother is HH Head	N	N	N	Y	Y					
Observations	489	489	489	489	489					
Joint F: # Kids + # Kids x 1st is a Son			11517	11271	11054					
p-value			0.00140	1.48e-05	2.15e-05					

All estimates control for city fixed effects. Robust standard errors are presented in parentheses. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger.

Table 5: The First Stage Estimates of the Effect of Having a First Child during or after 1972 on Total Fertility

	Dependent Variables						
	# Kids	# Kids	# Kids	# Kids	# Kids	# Kids x 1st is a Son	1st is a Son
	(1)	(2)	(3)	(4)	(5) Baseline	(6) Baseline	(7)
1st Born 1972+	-0.619	-0.712	-0.773	-0.751	-0.922	-0.043	-0.041
	(0.071)	(0.084)	(0.086)	(0.087)	(0.110)	(0.041)	(0.057)
1st Born 1972+ x 1st is a son					0.319	-0.527	
					(0.129)	(0.086)	
1st is a Son					-0.269	2.279	
					(0.102)	(0.065)	
Controls							
HH Head Age	N	Y	Y	Y	Y	Y	N
HH Head Age Squared	N	Y	Y	Y	Y	Y	N
HH Head Years of Edu	N	Y	Y	Y	Y	Y	Y
HH Head Years of Edu Squares	N	Y	Y	Y	Y	Y	Y
HH Head Age >55	N	N	Y	Y	Y	Y	Y
Age of Youngest Child	N	Y	Y	Y	Y	Y	Y
Youngest Child Age < 22	N	N	Y	Y	Y	Y	Y
Mother is HH Head	N	N	N	Y	Y	Y	Y
Observations	489	489	489	489	489	489	489
Joint: 1st Born 1972+ 1st Born 1972 x 1st is a Son					-0.603	-0.570	
p-value					1.50e-08	3.83e-10	
F-Statistic for Joint Significance					41.49	20.91	
Kleibergen-Paap F statistic					21.12		

All estimates control for city fixed effects. Robust standard errors are presented in parentheses. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger. The Stock-Yogo (2005) statistic for 10% maximal IV size is 7.03.

Table 7: The Reduced Form Estimates of the Effect of Having the First Child during or after 1972 on Savings

	Dependent Variables									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
				Baseline					Baseline	
	A. Savings					B. Savings/Income				
1st Born 1972+	7,527.889	5,713.623	11,506.016	12,397.174	10,689.857	0.057	0.057	0.097	0.100	0.086
	(1,826.697)	(2,612.811)	(3,665.662)	(3,780.687)	(3,519.273)	(0.029)	(0.030)	(0.043)	(0.044)	(0.040)
1st Born 1972+ x 1st is a son			-10,979.239	-10,917.129	-11,100.555			-0.077	-0.076	-0.078
			(4,405.831)	(4,398.967)	(4,411.587)			(0.054)	(0.054)	(0.054)
1st is a Son		2,997.797	9,099.395	9,205.125	9,238.178		0.002	0.045	0.045	0.046
		(2,186.879)	(3,336.418)	(3,341.292)	(3,339.285)		(0.027)	(0.042)	(0.042)	(0.042)
Controls										
HH Head Age	N	Y	Y	Y	N	N	Y	Y	Y	N
HH Head Age Squared	N	Y	Y	Y	N	N	Y	Y	Y	N
HH Head Years of Edu	N	Y	Y	Y	Y	N	Y	Y	Y	Y
HH Head Years of Edu Squares	N	Y	Y	Y	Y	N	Y	Y	Y	Y
HH Head Age >55	N	N	Y	Y	Y	N	N	Y	Y	Y
Age of Youngest Child	N	Y	Y	Y	Y	N	Y	Y	Y	Y
Youngest Child Age < 22	N	N	Y	Y	Y	N	N	Y	Y	Y
Mother is HH Head	N	N	N	Y	Y	N	N	N	Y	Y
Observations	489	489	489	489	489	489	489	489	489	489
Joint: 1st Born 1972+ 1st Born 1972 x 1st is a Son			526.8	1480	-410.7			0.0205	0.0238	0.00782
p-value			0.867	0.641	0.890			0.582	0.539	0.833
	C. Income									
1st Born 1972+	7,179.705	205.359	3,793.111	4,782.209	4,137.786					
	(1,666.174)	(3,378.096)	(4,485.835)	(4,512.921)	(4,341.533)					
1st Born 1972+ x 1st is a son			-6,800.435	-6,731.499	-6,832.448					
			(5,536.200)	(5,524.668)	(5,526.351)					
1st is a Son		4,519.150	8,298.422	8,415.771	8,460.013					
		(2,830.048)	(4,121.150)	(4,119.841)	(4,108.443)					
Controls										
HH Head Age	N	Y	Y	Y	N					
HH Head Age Squared	N	Y	Y	Y	N					
HH Head Years of Edu	N	Y	Y	Y	Y					
HH Head Years of Edu Squares	N	Y	Y	Y	Y					
HH Head Age >55	N	N	Y	Y	Y					
Age of Youngest Child	N	Y	Y	Y	Y					
Youngest Child Age < 22	N	N	Y	Y	Y					
Mother is HH Head	N	N	N	Y	Y					
Observations	489	489	489	489	489					
Joint: 1st Born 1972+ 1st Born 1972 x 1st is a Son			-3007	-1949	-2695					
p-value			0.479	0.641	0.508					

All estimates control for city fixed effects. Robust standard errors are presented in parentheses. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger.

Table 7: 2SLS Estimates of the Effect of Fertility on Savings

	Dependent Variables									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
				Baseline					Baseline	
	A. Savings					B. Savings/Income				
# Kids	-11,531.187	-7,367.422	-12,906.470	-14,008.412	-11,632.678	-0.087	-0.073	-0.107	-0.112	-0.092
	(4,524.971)	(3,582.007)	(4,471.447)	(4,667.909)	(4,104.585)	(0.057)	(0.040)	(0.051)	(0.052)	(0.046)
# Kids x 1st is a Son			13,039.994	12,234.026	14,120.616			0.081	0.077	0.094
			(6,947.008)	(7,007.983)	(6,834.650)			(0.084)	(0.085)	(0.085)
1st is a Son		2,297.605	-24,104.676	-22,440.060	-26,063.176		-0.005	-0.168	-0.161	-0.193
		(2,235.728)	(14,144.161)	(14,280.470)	(13,952.113)		(0.027)	(0.171)	(0.173)	(0.172)
Controls										
HH Head Age	N	Y	Y	Y	N	N	Y	Y	Y	N
HH Head Age Squared	N	Y	Y	Y	N	N	Y	Y	Y	N
HH Head Years of Edu	N	Y	Y	Y	Y	N	Y	Y	Y	Y
HH Head Years of Edu Squares	N	Y	Y	Y	Y	N	Y	Y	Y	Y
HH Head Age >55	N	N	Y	Y	Y	N	N	Y	Y	Y
Age of Youngest Child	N	Y	Y	Y	Y	N	Y	Y	Y	Y
Youngest Child Age < 22	N	N	Y	Y	Y	N	N	Y	Y	Y
Mother is HH Head	N	N	N	Y	Y	N	N	N	Y	Y
Observations	489	489	489	489	489	489	489	489	489	489
Joint F: # Kids + # Kids x 1st is a Son			133.5	-1774	2488			-0.0268	-0.0352	0.00178
p-value			0.981	0.762	0.613			0.686	0.616	0.977
	C. Income									
# Kids	-10,997.841	-264.799	-4,570.870	-5,619.951	-4,890.406					
	(3,942.525)	(4,360.297)	(4,968.290)	(5,061.521)	(4,651.841)					
# Kids x 1st is a Son			10,137.327	9,370.021	10,016.151					
			(8,430.953)	(8,443.108)	(8,337.241)					
1st is a Son		4,493.984	-16,031.221	-14,446.458	-15,676.232					
		(2,883.859)	(17,292.606)	(17,347.403)	(17,134.242)					
Controls										
HH Head Age	N	Y	Y	Y	N					
HH Head Age Squared	N	Y	Y	Y	N					
HH Head Years of Edu	N	Y	Y	Y	Y					
HH Head Years of Edu Squares	N	Y	Y	Y	Y					
HH Head Age >55	N	N	Y	Y	Y					
Age of Youngest Child	N	Y	Y	Y	Y					
Youngest Child Age < 22	N	N	Y	Y	Y					
Mother is HH Head	N	N	N	Y	Y					
Observations	489	489	489	489	489					
Joint F: # Kids + # Kids x 1st is a Son			5566	3750	5126					
p-value			0.434	0.601	0.427					

All estimates control for city fixed effects. Robust standard errors are presented in parentheses. # Kids is instrumented with a dummy variable for if the eldest child is born in 1972 or afterwards; # Kids x 1st is a Son is instrumented by a dummy variable for whether the first child is born in 1972 or afterwards x a dummy variable for whether the first child is a son. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger.

Table 8: The Effect of Fertility on Savings -- Placebo Experiment

	Dependent Variables				
	# Kids	Savings		Savings/Income	
	(1)	(2)	(3)	(4)	(5)
	1st	OLS	2SLS	OLS	2SLS
# of Kids			-1,498.017		0.606
			(35,004.647)		(0.433)
#Kids x 1st is Son			6,937.523		-0.050
			(8,263.569)		(0.122)
1st is Son	0.010	2,397.747	-11,809.002	-0.011	0.085
	(0.092)	(2,743.275)	(15,157.435)	(0.031)	(0.208)
HH Head Age<60	-0.218	2,129.418		-0.145	
	(0.114)	(8,967.291)		(0.069)	
HH Head Age<60 x 1st is Son	0.063	-4,356.380		0.069	
	(0.104)	(6,544.603)		(0.044)	
Observations	991	991	991	991	991
R-squared	0.310	0.040	0.042	0.076	

All regressions control for age of HH head, age of HH head squared, edu years of HH head, edu years of HH head squared, age of the youngest child, a dummy for if the youngest is less than 22 years of age, a dummy for if the mother is the HH head, and city fixed effects. Robust standard errors are presented in parentheses. The instruments are HH Head Age<60 and HH Head Age<60 x 1st is son. Notes: Sample uses data from the UHS where households are headed by individuals age 55-65.

Table 9: The Effect of Fertility on Savings -- Robustness of Standard Errors

	Dependent Variables	
	(1)	(2)
	Savings	Savings/Income
# Kids	-11632.67764	-0.09199
Robust (Main)	(4,104.58463)	(0.04561)
No Correction	(3,794.14369)	(0.04487)
Cluster at City Level	(5,188.19736)	(0.05592)
Cluster at YOB level	(3,684.35997)	(0.04040)
Observations	489	489
Joint	2488	0.00178
p-value	0.613	0.977

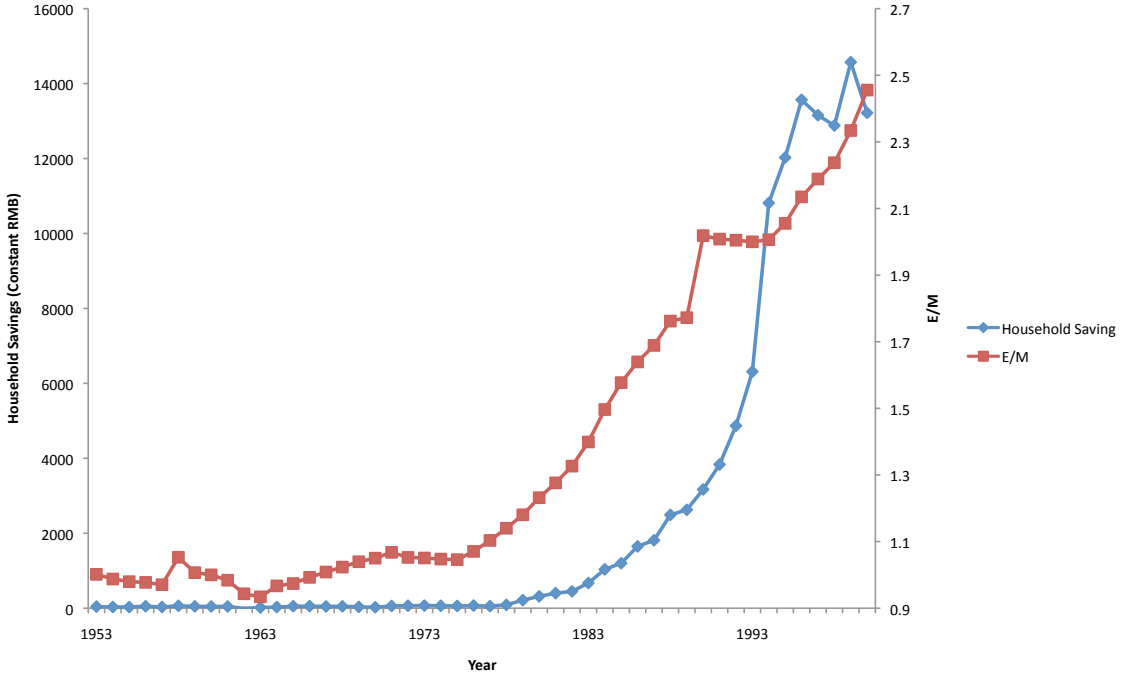
All regressions control for t# Kids x 1st is a Son, a dummy for if the 1st child is a son, the age of HH head, age of HH head squared, edu years of HH head, edu years of HH head squared, age of the youngest child, a dummy for if the youngest is less than 22 years of age, a dummy for if the mother is the HH head, and city fixed effects. # Kids is instrumented with a dummy variable for if the eldest child is born in 1972 or afterwards; # Kids x 1st is a Son is instrumented by a dummy variable for whether the first child is born in 1972 or afterwards x a dummy variable for whether the first child is a son. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger.

Table 10: The Effect of Fertility on Savings -- Heterogeneous Effects

	Dependent Variable: Savings					
	City Gender Wage Gap		Marital Status of Children			
	Fem/Male Wage Above Median	Fem/Male Wage Below Median	1st Single	1st Married	Have Single Child	All Children Married
	(1)	(2)	(3)	(4)	(5)	(6)
# Kids	-9,697.162	-18,424.923	-13,145.608	-14,018.948	-12,547.512	-11,361.633
	(5,665.906)	(7,585.530)	(26,490.617)	(5,029.993)	(10,410.100)	(5,099.858)
#Kids x 1st is Son	9,707.532	19,292.452	26,613.805	10,089.528	35,584.614	6,944.167
	(8,397.894)	(12,399.799)	(30,955.858)	(8,735.538)	(27,747.755)	(7,540.032)
1st is Son	-20,284.143	-34,521.303	-45,781.689	-16,944.791	-63,098.034	-12,124.469
	(17,820.140)	(24,582.455)	(51,365.740)	(18,555.650)	(54,032.414)	(15,835.847)
Observations	210	279	87	402	148	341
Joint F: # Kids + # Kids x 1st is a Son	10.37	867.5	13468	-3929	23037	-4417
p-value	0.999	0.929	0.665	0.597	0.370	0.475

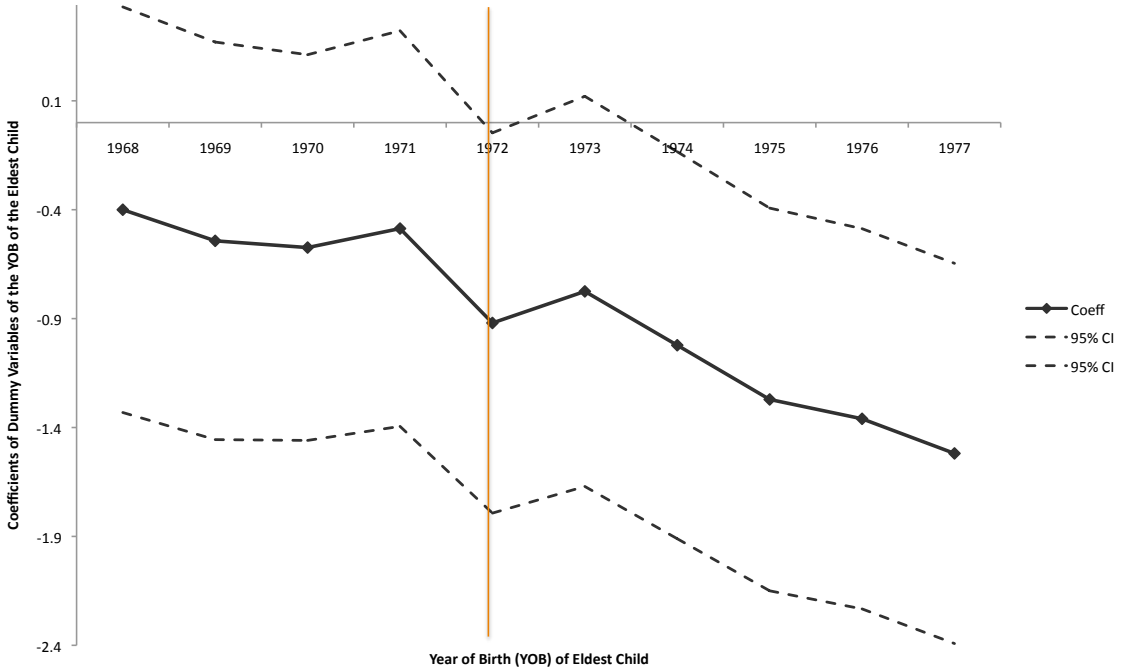
All regressions control for age of HH head, age of HH head squared,a dummy for if the HH head is older than 55, edu years of HH head, edu years of HH head squared, age of the youngest child, a dummy for if the youngest is less than 22 years of age, a dummy for if the mother is the HH head, and city fixed effects. # Kids is instrumented for by a dummy variable for the eldest child being born Post1972. # Kids x 1st Son is instrumented for by Post1972 x 1st is Son. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger.

Figure 1: Chinese Savings Rates and the Demographic Structure over Time



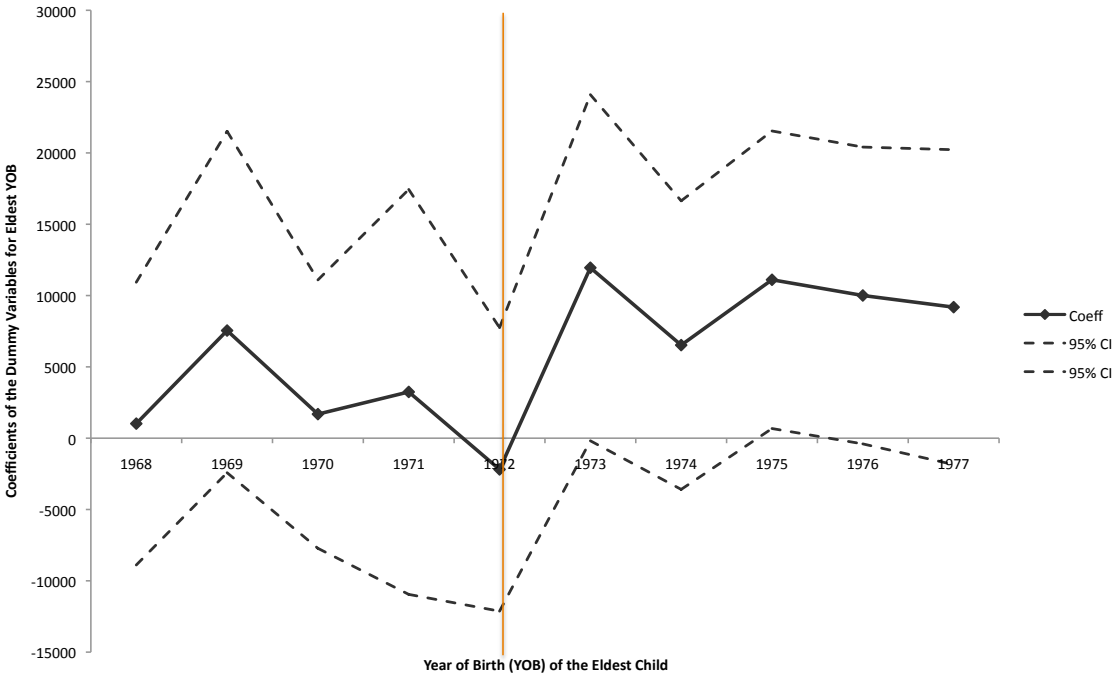
Source: Modigliani and Cao (2003) (Original source: China Statistical Yearbooks)

Figure 2: The Correlation between the Year of Birth of the Eldest Child and Total Household Size
-- The coefficients of the dummy variables for the YOB of the first child



Source: Authors' computations.

Figure 3: The Correlation between the Year of Birth of the Eldest Child and Household Savings
-- The coefficients of the dummy variables for the YOB of the first child



Source: Authors' computations.

Appendix Table A1: The Correlation between the Birth Year of the First Child and Total Fertility and Savings

	Dependent Variables	
	(1)	(2)
	# Kids	Savings
Born 1968	-0.400	1021
	(0.474)	(5045)
Born 1969	-0.543	9547
	(0.465)	(6091)
Born 1970	-0.574	1686
	(0.451)	(4788)
Born 1971	-0.486	3243
	(0.462)	(7224)
Born 1972	-0.920	-2190
	(0.444)	(5058)
Born 1973	-0.775	11951
	(0.456)	(6175)
Born 1974	-1.022	6525
	(0.452)	(5148)
Born 1975	-1.271	11107
	(0.447)	(5309)
Born 1976	-1.360	10005
	(0.444)	(5295)
Born 1977	-1.519	9186
	(0.444)	(5618)
Observations	489	489
R-squared	0.226	0.034
Joint F	9.831	2.898

Robust standard errors are presented in parentheses. Notes: Sample uses data from the UHS where households have their first child during 1967-77, and the age of household is 65 or younger.