2-16-2009

Sex Ratios and Savings Rates: Evidence from "Excess Men" in China

Shang-Jin Wei (Columbia University and NBER)

and

Xiaobo Zhang (IFPRI)

Abstract

High savings rates in many countries are a major factor in the global current account imbalances. The Chinese savings rate has been high and rising, contributing to one of the world's largest and fast-growing current account surpluses. The life cycle theory and precautionary savings motive, in spite of their popularity, are not compatible with recent time series patterns. The paper tests a new hypothesis based on a biological savings motive: a high and rising sex ratio imbalance, by increasing the competition by men in the marriage market, may have induced the Chinese to postpone consumption significantly in favor of wealth accumulation. We report both "macro" and "micro" evidence in support of the hypothesis. First, across provinces, local saving rates are found to be strongly positively associated with local sex ratio imbalances, after accounting for demographics, social safety net and other factors. The relationship continues to hold with an instrumental variable approach. Second, across households, those with a son tend to save more in regions with a more skewed sex ratio, holding constant other household features. The rise in sex ratio imbalance can potentially account for half of the actual increase in household savings in the recent decade.

Contact information: Shang-Jin Wei, Columbia Business School, 619 Uris Hall, 3022 Broadway, New York, NY 10027, USA. Shangjin.wei@columbia.edu; Xiaobo Zhang, International Food Policy Research Institute (IFPRI), 2033 K Street, NW, Washington, DC, 2006, USA. X.zhang@cgiar.org.

The authors would like to thank Marcos Chamon, Avi Ebernstein, Lena Edlund, Roger Gordon, Justin Lin, Olivier Jeanne, Simon Johnson, Junsheng Zhang, and seminar participants at the NBER and various universities for helpful discussions, and John Klopfer, Lisa Moorman and Jin Yang for able research assistance.

1. Motivation and Overview

Many countries have a high savings rate, which often manifests itself in a large current account surplus and rising foreign exchange reserves. A notable example that has attracted international attention recently is the People's Republic of China, whose national savings accounted for close to 50 percent of GDP in 2007. This savings rate is higher than in China's past: the national savings rate was around 30% of GDP in the early 1980s. It is higher than most other countries, including many East Asian countries known to have a high savings rate, and higher than China's already-high investment-to-GDP ratio. China's household savings are approximately half of its national savings.

Why do households in China save so much? There are three classes of explanations in the existing literature. The first is the life cycle theory (Modigliani, 1970), which predicts that the savings rate rises with the share of working age population in the total population. This explanation appears to work at the aggregate level as the youth dependence ratio declines and the share of working age population rises due to China's strict population control policy that was adopted in 1980 (Modigliani and Cao, 2004). However, this explanation doesn't appear to work well at the household level. According to Chamon and Prasad (2008), who study urban Chinese households, the savings rate as a function of the age of the head of household exhibited an inverse U shape only in 1995—as predicted by the life cycle theory—but the relationship flattened over time, resulting eventually in a U shape by 2005. This suggests that the life cycle theory would have difficulty with the recent rise in China's savings rate.

The second class of explanation has to do with a precautionary savings motive in combination of a rise in income uncertainty. This is the explanation favored by Blanchard and Giavazzi (2005) and Chamon and Prasad (2008). The logic of the explanation goes as follows: the economic reform of the last three decades has undermined the traditional social safety net under the state-owned, central-planning system and increased job and income uncertainty at the same time. Urban workers used to work for state-owned firms that provided job security, housing, health benefits and pensions. Now, seven out of every ten workers either work for a private firm or are self-employed with no job security.

This could motivate people to save more than they otherwise would have. While this explanation appears to be consistent with a comparison between 1979 and 2009, it is likely contradicted by a comparison between 2000 and 2009. By 2000, China was already firmly a private-sector driven, market-based economy. Since 2000, while there has been no substantial increase in job/income uncertainty, there has been a steady and visible improvement in social safety net. For example, a hybrid pension system that includes an individual account and government funding has been developed and put in place across all cities during this period. A health insurance system has also started to become a norm in the urban areas. A minimum-income scheme ("di bao") was rolled out progressively to cover all cities in the nation: under this scheme, urban households whose income is below a threshold would receive a transfer from the government regularly (Ravallion 2008). The *di bao* system was further expanded to rural areas in 2007. The logic of a precautionary savings motive should have predicted a decline in Chinese savings rate, yet it has increased by close to ten percentage points of GDP during this period.

The third possible explanation is cultural norm. But cultural norm tends to be persistent, and if it plays a role, it would also have difficulty in explaining the visible rise in the savings rate in the last two decades.

In this paper, we test a potentially new explanation based on a biological motive that the high and rising savings rate is, in a significant part, due to a high and rising sex ratio imbalance, or a widening gap between the numbers of men and women. We summarize our explanation and findings in three steps: (a) some fact - how skewed is the sex ratio? (b) The logic - why would a rising sex ratio imbalance lead to a rise in the savings rate? And (c) some evidence – do variations across regions and households support the hypothesis?

If it is left to the nature, the sex ratio at birth is supposed to be 105 boys per 100 girls (with human biology compensating for a slightly higher mortality rate for infant boys than girls). The sex ratio in China was close to that in 1980 (with 107 boys per 100 girls), but climbed steadily over time to over 120 boys/100 girls in 2005 (Li, 2007). This implies an increased likelihood that a man, in adulthood, will not to be able to find a wife. As a result, by 2005, there are about 40 million men who cannot mathematically get married due to a shortage of women.

How would this affect the savings behavior? It is safe to assume that most men have a strong desire to get married. As the likelihood of not being able to find a mate increases, the parents of a son (or the son himself) are willing to try harder to improve the son's prospects for marriage. Postponing consumption and raising household savings rate may be one such thing that could enhance the probability of attracting a partner.

The initial inspiration for the idea in the paper comes from (our visits to the National Zoo in Washington DC, and) the literature on evolutionary biology. In nature, male sea lions or walruses tend to be physically larger than corresponding females; in fact, this pattern holds for most species. A leading hypothesis is sexual dimorphism, which conjectures that because bigger and stronger males have an advantage in competing for and retaining females, this difference in size gets reinforced over time by natural selection (Weckerly 1998). It is likely that humans had to do the same a long time ago; that is why men, on average, are somewhat taller and larger than women. By now, the (sexual) returns to progressively larger males are much lower, if not negative. Men may discover that possessing a larger house, a bigger savings account, and more general wealth is a more effective mating strategy.

A relevant economics literature is the work on status goods, positional goods, and social norm (e.g., Cole, Mailath and Postlewaite, 1992; Hopkins and Kornienko, 2004). By allowing certain goods to offer utility beyond its direct consumption value (i.e., through "status,' which in turn could affect the prospect of finding a marriage partner), it is easy to show that the consumption and savings behavior would be altered. It is important to point out, however, that the effect of a rise of a sex ratio imbalance on savings rate is conceptually ambiguous. On one hand, when the competition in the marriage market increases, men (or their parents) may compete by increasing their conspicuous consumption, *if* conspicuous consumption could effectively signal their attractiveness; this would result in a decline in the savings rate. On the other hand, the probability of securing a marriage partner may depend more on being able to show a higher level of wealth than showing off a few flashy goods. Furthermore, a visible way to demonstrate a man's desirability is by owning a larger house (at least in a developing

country like China), which requires a larger savings before the purchase is made than otherwise. Both can induce households with a son to raise their savings rate[1].

When the savings effect dominates the conspicuous consumption effect, one may still ask whether women (or parents of a girl) can decrease their savings to completely offset extra savings by men (or their parents). As women become more scarce, it is tempting to think that they (or their parents) may increase their consumption (and hence reduce their saving) to take advantage of the change in their favor. However, there can be spillovers that go in the opposite direction: for example, the increased competition by men for potential mates may bid up local housing prices. As a consequence, even parents with a daughter in a region with a more skewed sex ratio might have to save more in order to afford housing (or at least might not reduce their saving by as much as they would otherwise do).

To summarize, while the economics literature suggests that savings behavior would be altered when people care about getting married and the savings/consumption patterns could affect the probability of getting married. The sign of the effect on savings behavior, however, can go either way. Under the conditions that a higher wealth level is a more effective mating strategy than more conspicuous consumption, and competition by men spills over to households with a daughter via a rise in the price of non-tradables, a rise in sex imbalance would result in a higher savings rate.

The discussion so far assumes that changes in sex ratio are exogenous: in particular, it assumes that the parental preference for boys does not respond to a change in the sex ratio, at least in the short run. One potential justification for our assumption of exogeneity is that such a preference is part of a culture, and as such, it changes only very slowly. Korea has experienced a sustained increase in its sex ratio imbalance for about 25 to 30 years, which has only recently started to decline; this evidence is consistent with

[1] Madonna's hit song, "Material Girl," contains the following lyrics: *"We are/ living in a material world/ and I am a material girl/ Some boys try and some boys lie but/ I don't let them play/ Only boys who save their pennies/ make my rainy day…*" To be clear, our hypothesis does not require women (in China or elsewhere) to be material girls in order to induce men to save more. Other factors can be important for the success in finding a mate. But other things being equal, as long as a wealthier man is preferred to a less wealthy one, men (or their parents) have an incentive to raise their savings rate in order to enhance their prospects for marriage.

our assumption (Guilmoto, 2007). In the empirical part, we will report instrumental variable regressions that allow for potential endogeneity of sex ratio.

We have not found any prior empirical work that links savings rates to sex ratio imbalances, either for a single country or across countries. Since a country's saving behavior is shaped by many factors including history, culture, and institutions—factors that may not be easily quantified—we regard a within-country study as a useful exercise, because factors that are common within a country are naturally controlled for.

Since 1980, both the sex ratio and the savings rate in China have been rising. In Figure 1, we present a time series plot of a standardized version of both variables,[2] with the sex ratio at birth lagged by twenty years. The sex ratio variable is lagged to account for the fact that the median age of first marriage for Chinese women is about 20. The two standardized variables, visually, are highly correlated (the actual correlation coefficient is 0.822). While this suggests that the sex ratio hypothesis of savings rates cannot be dismissed out of hand, this is not a rigorous proof of the hypothesis, either.

We provide two different types of evidence for our hypothesis. First, we run panel regressions on provincial-level savings rates from 1978-2006. In addition to the local sex ratio at birth, lagged by twenty years, we include as regressors local per capita income, the share of working age (25-60) population in local population, yearly fixed effects to account for changes in the national social security and health care systems, and provincial fixed effects to account for locally idiosyncratic factors. We run separate regressions for urban and rural areas, and show that local savings rates are systematically and positively correlated with local sex ratios (for the pre-marriage age cohort). Therefore, the positive association between these two variables holds across regions and over time. An instrumental variable approach suggests that the relationship likely reflects a causal effect of sex ratios on savings rates.

A second set of evidence that we provide examines household-level savings decisions, using data from household surveys in 122 rural counties and 70 cities. Since the cultural norm is such that men or parents of the groom are generally expected to cover the cost of weddings and of the newlyweds' first house or apartment, one might expect that households with a son would need and tend to save more than households with a

[2] standardized variable = (raw variable – mean)/standard deviation.

daughter. However, since new wives tend to live with the husbands' families, especially in rural areas, parents of a daughter may need to save more for their own financial security in old age, to the extent that they will be less able to rely on their daughters when they are old than parents with a son would be. This means that we cannot easily confirm or reject our hypothesis by comparing the average savings rates between households with a son or a daughter.

Our hypothesis has one distinct prediction for household data: a higher local sex ratio raises the degree of competition by men for potential mates, which may motivate parents with a son to save more than they would otherwise save (holding constant all other household characteristics). The effect of the local sex ratio on the savings rate of households with a daughter is a bit ambiguous. On one hand, if competition for husbands becomes less intense due to a higher male/female ratio, one may expect these households to save less. On the other hand, there could be general equilibrium effects that go in the other direction. For example, if a higher sex ratio raises the competition among men and bids up the local housing price (or that of other non-tradables), then even households with a daughter need to save more to afford housing, as a result.

What does the data say? In the rural areas, we find that the savings rate of three-person households with a son tends to increase with the local sex ratio, after accounting for the effect of household size, age and education level of the head of a household. In comparison, the savings rate of three-person households with a daughter or of four-person households with two daughters is not systematically related to the local sex ratio. In the urban areas, we find that the savings rate is a positive function of the local sex ratio both for households with a son and for those with a daughter, though the point estimate of the elasticity in this case is bigger for households with a son. These empirical patterns confirm the idea that savings rates are related to local sex ratio imbalances.

We also provide some additional, supplementary evidence. Working with data at the rural county level, we are able to examine two other variables of interest: house or apartment sizes and bank account balances. A bank account balance is literally a part of a household's savings, whereas the size of a house is partly a result of previous savings and can be used by a man (or his family) to enhance his attractiveness to a potential wife. In the data, we find both variables to be positively correlated with the local sex ratio lagged

by twenty years, after accounting for county fixed effects, average income, and the age profile of the county.

The rest of the paper is organized as follows. Section 2 describes the key variables and their sources. Section 3 reports the statistical results. Finally, Section 4 provides concluding remarks.

2. Data and Summary Statistics

2.1 Sex Ratios

Data on sex ratios at birth for cohorts born later than 1988, at the national level, are from Coale and Banister (1994, Table 3). Because the recorded sex ratio at birth in their table is by five-year birth cohorts, we use the third year of the five-year period as the corresponding year of birth. The sex ratios at birth in the period of 1988-1993 are from Gu and Roy (1995). The sex ratios for the birth cohorts of 1994-2000 are calculated directly from the *China Population Census 2000*.

Sex ratios at the provincial level for the age cohort 6-24 are inferred from the 2000 census. For example, the age cohort 6-24 in 2006 is identified as the 0-18 cohort in 2000; the 6-24 cohort in 1990 is identified as the 16-34 cohort in 2000. The sex ratio data at the county level analysis in 2000 are based on the birth cohort of ages 0-9 at county level, from *China Population Census 1990*.

2.2 Other data used in cross-regional panel regressions

Savings rate: There are several ways to define the savings rate. The first is the standard (Y-C-G)/Y, where Y is GDP, and C and G are household and government consumptions, respectively. Total GDP and the final consumption component of GDP are from *Comprehensive Statistical Data and Materials on 50 Years of New China* (CNBS), *Gross Domestic Product of China: 1996–2002* (CNBS).and *China Statistical Yearbook 2007* (CNBS). This variable is systematically available for a long period. However, it does not offer a rural-urban distinction.

The second savings variable is defined as the ratio of the difference between per capita disposable (net) income and living expenditure to per capita disposable (net) income in urban (rural) areas. The per capita disposable income and living expenditure in cities from 1985 to 1998 and the per capita rural net income and living expenditure for the period of 1978-1998 are from *Comprehensive Statistical Data and Materials on 50 Years of New China* (CNBS). The data in later years are from various issues of *China Statistical Yearbooks*.

Total residential deposits at the national level prior to 1999 are from *Comprehensive Statistical Data and Materials on 50 Years of New China* (CNBS). For later years, these data are available from various issues of *China Statistical Yearbooks*. The residential bank deposit and per capita GDP at the county level in 2002 are from *China County Social and Economic Statistical Yearbooks* (CNBS).

Living space. The living space per household at the county/district level is from *China Population Census 2000*.

2.3 Other data used in household level regressions

The rural and urban household survey data sets are obtained from the Chinese Household Income Project (2002), available from the website of the Inter-University Consortium for Political and Social Research (http://www.icpsr.umich.edu/cocoon/ICPSR/STUDY/21741.xml).

3. Statistical Evidence

We provide a series of complementary evidence. First, we examine "macro-level" evidence on the relationship between local sex ratios and local saving rates across Chinese provinces using panel fixed effects regressions. To go from correlation to causality, we also adopt an instrumental variable approach based on regional variations in the enforcement of the family planning policy. Second, we scrutinize micro, household level data. The primary goal is to check whether and how savings by households with a son (or a daughter) vary by the degree of the local sex ratio imbalance. Finally, we

provide supplementary evidence on regional variations in the size of houses/apartments, bank deposits, and time profile of savings in relation to a marriage event.

3.1 Panel regression evidence across Chinese provinces

We start by examining provincial-level data for any association between local savings rates and local sex ratios. It is important in this exercise to separate urban and rural areas as the two have different income and education levels and different social safety nets. Parental preference for a son is known to be stronger in rural than in urban areas. Perhaps most importantly, local marriage markets may work differently in each area. A majority of marriages in rural areas take place between a man and a woman from the same county[3]. Migrant workers often either get married before leaving their homes to work in a city, or return to their villages to get married. In comparison, a greater proportion of marriages in a city could involve a man or a woman (or both) from outside the city. An implication of this difference in the marriage market is that the same local sex ratio imbalance may exert a much smaller competitive pressure on men in a city than in a rural area. Due to these differences, we need to run separate regressions for rural and urban areas that allow for different sensitivities of local savings rates to local sex ratios.

We perform a panel fixed effects regression that links a location *j*'s savings rate in year *t* with the sex ratio for the appropriate age cohort in that same location and year, controlling for location fixed effects, year fixed effects, and other factors. To be precise, our specification is the following:

$$\text{Savings\_rate}_{k,j,t} = \beta\ \text{Sex\_ratio}_{k,j,t} + X_{k,j,t}\,\Gamma_{k,j,t} + \text{province fixed effects} + \text{year dummies} + e_{k,j,t}$$

where *k* = "rural" or "urban." Following Chamon and Prasad (2008), we define the local savings rate by log (net income/living expenditure) for rural residents and log (disposable income/living expenditure) for urban residents.[4]

[3] According to the *China Population Census 2000*, about 10% of people migrate to places outside their registered counties or districts in 2000. Among them, only 2.5% list marriage or family reunion as the major reason of migration.

[4] Income and expenditure data, which are aggregated based on nationwide rural and urban household surveys, are from various issues of *China Statistical Yearbooks*.

Ideally, we would like to have sex ratio for a fixed age cohort year by year. However, such data are not available as the population census has been carried out only once every few years (in 1982, 1990 and 2000). Moreover, only the 2000 census has data that separates rural and urban areas at the provincial level. Given these constraints on available data, we make the following short cut: we focus on the sex ratios for the age cohort 6-24 in all years, and derive the values from the 2000 population census. To be precise, for the year 2000, we know the exact sex ratio for this age cohort from the census. For the age cohort 6-24 in 2006, we infer the sex ratio with data for the age cohort 0-18 in the 2000 census, since the two groups should theoretically be the same. Similarly, for the age cohort 6-24 in 1990, we match it with the cohort 16-34 in the 2000 census; and so on.

A caveat with this method is that the actual sex ratio is likely to be different from the inferred one for all years other than 2000. In particular, because the mortality rates for boys and young men are generally slightly higher than those for girls and young women, we may under-estimate the true sex ratios for years before 2000 and over-estimate them for years after 2000. However, under the assumption that measurement errors are common across all regions in any given year (but may vary from year to year), we can eliminate the effect of measurement errors by including yearly fixed effects in regressions.

As control variables, we include both log income and squared log income. In addition, we account for local demographic features by including the proportions of local population in the age brackets of 0-24 and 25-60, respectively. Table 1a presents the summary statistics of the major variables used in the provincial level analysis.

In Column 1 of Table 2, we report the regression results for the rural areas. The effect of local income on local savings rates is essentially linear: a one percent increase in local income is associated with a higher local savings rate by 0.52 percentage points. The coefficients on the demographic variables are sensible and consistent with the life-cycle theory of saving: A greater share of working population (age 25-60) is associated with a higher local savings rate. On the other hand, a greater share of age cohort 0-24 is also associated with a modestly higher local savings rate, possibly because the latter cohort also includes young workers. The age group that is left out (i.e., 60 or older) is associated with a lower local savings rate. The coefficient on local sex ratio (for the age cohort 6-24) is 0.831 and statistically different from zero at the 1% level. An increase in the sex ratio

by 10 basis points (e.g., from 1.10 to 1.20) is associated with an increase in the local savings rate of 8.31 percentage points.

One might worry that income (earnings) inequality could affect local savings rates directly, and that a sex ratio imbalance is simply one source of earnings inequality. For a subset of provinces and years, we can measure local income inequality by the Gini coefficient (separate for rural and urban areas).[5] In Column 2 of Table 2, we add the local Gini coefficient as an additional control. This variable indeed has a positive coefficient (0.62). However, with a t-statistic of only 0.7, the coefficient is not statistically different from zero. With this much reduced sample, the coefficient on the local sex ratio is still positive and statistically significant. The point estimate of the coefficient on the local sex ratio jumps to 2.20.

We have also added local life expectancy as an additional regressor to account for the possibility that households raise savings rate when they expect to live longer. The coefficient on this variable is positive but statistically insignificant (not reported).

We perform a similar set of regressions for the urban areas across Chinese provinces. The results are reported in columns 3-4 of Table 2. With year dummies and province fixed effects, a higher local income is also associated with a higher local savings rate (Column 5). Local income inequality measured by the Gini coefficient is not statistically significant. Most important for our purpose, the coefficients on the local sex ratio are positive and statistically significant across both specifications. The point estimates are smaller than their counterparts for the rural regression. Using the point estimate in Column 3 as an illustration (0.585 on sex ratio), an increase in a local sex ratio by 10 basis points (e.g., from 1.10 to 1.20) is associated with a higher savings rate by 5.9 percentage points. (The share of local population that is in working age has a negative sign, which is surprising from the viewpoint of the life-cycle theory, but is consistent with the findings in Chamon and Prasad (2008)).

The existing literature has hypothesized that the deficiency in China's social safety net is what motivates the Chinese to save so much. In our regressions, since there is basically no social safety net to speak of in the rural areas, there is not much variation

[5] Data are available for 29 provinces in 1988, 1990 and 1993 and 28 provinces in 1996 and 1999 for the urban sample, and 27 provinces in 1988, and 28 provinces in all other years for the rural sample.

across regions in that dimension. However, in the urban areas, we can create some proxies for regional variation in the local social safety net. We use two proxies: the proportion of the local population that is enrolled in social security, and the proportion of the local population that is employed by state-owned enterprises (SOEs). Under the precautionary savings hypothesis, if a higher value in either variable signified a greater security of income, then the savings rate should decline. In Column 5 of Table 2, we include these variables. It turns out neither has a coefficient that is different from zero statistically. In fact, neither point estimate has a negative coefficient. This suggests that the precautionary saving motive, as proxied by these two variables, does not appear to be quantitatively important in explaining cross-regional differences in savings rates. We make this observation with one caveat: our evidence does not rule out the possibility that every region's saving rate has been made higher due to a poor national social safety net.

It is interesting to note that the estimated elasticity of local savings rates with respect to local sex ratios is greater for rural areas than for urban areas. Based on point estimates in Columns 1 and 3 of Table 2 and holding other factors constant, the actual increase in local sex ratios accounts for 68% and 18%, respectively, of the increase in local savings rates in the rural and urban areas. The difference in these estimated rural and urban elasticities is reasonable. First, the increase in sex ratio imbalances is more severe in the rural areas. Second, the marriage market tends to be more local in the rural areas. Rural migrant workers tend to return to their county of origin to get married. The 2000 population census suggests that 96% of marriages in the rural areas tend to take place between men and women from the same county. In comparison, a greater portion of marriages in a city may involve someone from outside the city.

While the absolute amount of per household savings is lower in the rural areas than in the urban areas due to differences in the income levels, rural households outnumber their urban counterparts by about 2 to 1. Therefore, approximately half of the increase in the household savings rate, nationally, can be attributed to the increase in China's sex ratio imbalance, holding other factors constant. This means that the effect of the sex ratio imbalance is economically significant.

3.2 Causes of sex ratio imbalance and instrumental variable regressions

The sex ratio imbalance is a result of a strict family planning policy established in the early 1980s to limit the number of children that a majority of Han Chinese families can have, in conjunction with a son-favored culture, and the (black market) availability of relatively inexpensive technologies to screen the sex of fetuses (primarily Ultrasound B) and to conduct abortions.[6] The degree of a local sex ratio imbalance is likely to be exogenous with respect to local savings decisions, especially since the relevant local sex ratios are determined by parental decisions – whether to undertake a sex selective abortion - taken many years prior to the corresponding savings variables in the regressions.

Nonetheless, one has to think hard about the possibility that sex ratios could be endogenous: For example, factors that affect sex-specific earnings in a region could simultaneously determine a preference for sons (Qian 2008) and household savings decisions. In addition, as we noted earlier, local sex ratios are measured with errors since they are inferred from the 2000 population census for years.

A solution to both the causality issue and measurement errors is to employ an instrumental variable approach, which we turn to now. We explore two determinants of local sex ratios which are unlikely to be affected by local savings rates, and for which we can get data. First, the national family policy exempts non-Han ethnic groups. That is, the 50-some ethnic minority groups do not face quotas on births. (The government allowed the exemption, possibly to avoid being criticized for using the family planning policy to marginalize minority groups). As a result, a region with a relatively high share of non-Han ethnic groups should likely have a relatively low sex ratio imbalance. To the extent that a Han Chinese man would be able to marry a non-Han woman, the competition among men in such a region would be lower than otherwise.

[6] China's family planning policy, commonly known as the "one-child policy," has many nuances. Since 1979, the central government has stipulated that Han families in the urban areas should generally have only child. Han is the main ethnic group and accounts for about 92% of the Chinese population. If both parents are single children, then they may have two children. Families in rural areas can generally have a second child if the first child is a daughter (this is referred to as the "1.5 children policy" by Eberstein, 2008). Ethnic minority (i.e., non-Han) groups are generally exempted from the limit on the number of children they may have. Non-Han groups account for a relatively significant share of local populations in Xinjiang, Yunnan, Ganshu, Guizhou, Inner Mongolia, and Tibet. Qian (2008) shows that regions with relatively high economic status due to high tea prices tend to exhibit a lower sex ratio imbalance.

Second, while the goal of family planning is national, the enforcement is local. Eberstein (2008) propose to use regional variations in the monetary penalties for violating the family planning policy as an instrument for the local sex ratio. Using data collected by Scharping (2003) and extending them to more recent years, Eberstein focus on two dimensions of penalties: (a) a monetary penalty for the violation of policy, expressed as a percent of annual income in the province, and (b) a dummy for the existence of extra penalty for having higher order unsanctioned births (e.g., for having a 3rd child in a 1-child zone, or for having a 4th child in a 2-child zone).[7]

Table 3 reports regressions that link these potential determinants to local sex ratios. Both types of variables appear to be related to the local sex ratios. First, the greater the fraction of the local population that is not subject to the family planning policy, the lower the local sex ratio imbalance. Second, greater financial penalties tend to be associated with a more skewed sex ratio. This effect is stronger in the rural areas than in the urban areas. (*86 % and 67% of the regional variation in the rural and urban sex ratios, respectively, are associated with the regional variations of these variables plus local income).* Of course, a proper first stage regression in a 2SLS needs to include other control variables in the main regression. In Columns 2 and 4, we include these controls.

In Table 4, we report the 2SLS estimation results for local savings rates where the local sex ratio is instrumented by the variables described in Table 3. The local sex ratio continues to have a positive coefficient that is statistically different from zero. The point estimates, 0.724 for the rural regression, and 0.501 for the urban regression, are somewhat smaller than the corresponding estimates without the instruments.

[We also perform 2SLS estimation using only data on monetary fines on violating birth quotas and obtain similar estimates.]

3.3 Household-level evidence

It is useful to go beyond regional aggregate comparisons and look at evidence at the household level. From data in the Chinese Household Income Project of 2002, which

[7] In principle, variations in the cost of sex screening technology especially the use of an Ultrasound B machine, and the economic status of women (such as that documented in Qian 2008) could also be candidates for instrumental variables. We, unfortunately, do not have the relevant data. Note, however, for the validity of the instrumental variable regressions, we do not have to have a complete list of the determinants of local sex ratio in the first stage.

covers 122 rural counties and 70 cities, we construct a sub-sample of households with one or two children older than five and a household head younger than 40.[8] Since the cultural norm is such that unmarried young people live with their parents, the survey does not contain many observations of households with a single young man or woman as the household head. Therefore, we are not able to analyze such households directly.

One may be tempted to compare savings rates for households with sons or daughters. But this comparison is not particularly informative with regard to our hypothesis. On one hand, if saving is in part motivated by competition for mates, one might expect those families with a son to save more. On the other hand, because the cultural norm is such that old people can more reliably expect to get help from their sons than daughters, families without a son may need to save more to prepare for their old-age days. These two opposing effects make it difficult to say whether, on net, families with a son should save more or less than families without a son. Table 5 reports average saving rates for households with children of different genders. In both rural and urban areas, households with a son (or two sons) have a moderately higher savings rate than those with a daughter (or two daughters). However, none of the difference in savings rates is statistically significant; the standard deviation of the savings rate with any given type of household easily overwhelms the difference in the savings rate between any two types of household. In any case, we cannot confirm or reject our hypothesis by comparing the savings rate across household types in this way.

Our hypothesis, however, implies a particular regional variation in saving rates: households with a son should save more in a region with a more unbalanced sex ratio, holding constant family income and other household characteristics. Moreover, this pattern is not predicted by either the life-cycle theory or the existing precautionary motive hypotheses in the literature. Therefore, examining the relationship between household savings rates and local sex ratios may be a particularly informative way to test our hypothesis.

We can also check the relationship between the savings rate by households with a daughter and the local sex ratio. On one hand, as the competition for a bride increases

[8] We put the limit of 40 years old for heads of households here to control for the possibility that some grown children may have married and lived outside the households, but the saving decisions of such households may not be comparable to others with only younger children.

among men, one might think that households with a daughter need to save less. On the other hand, as noted earlier, there can be a spillover effect through non-tradable goods that goes in the other direction. For example, local housing prices in a region with a highly skewed sex ratio could be bid up by households with a son. This would necessitate an increase in savings rates even by families with a daughter, in order for them to afford a house. This spillover channel is likely to operate more strongly in the urban areas as houses are more likely to be purchased from the local market (as opposed to be built on the occupants' own land in the rural areas). In short, due to this possible spillover channel, our hypothesis does not necessarily predict that the savings rate of households with a daughter declines with the local sex ratio, especially in the urban areas. However, one would expect the saving rates of households with a boy to be more sensitive to the local sex ratio than those of households with a daughter.

The regression results for the rural sample are presented in Table 6. With the full sample (the first four columns), it is clear that the savings rate by households with a son rises with the local sex ratio, exactly as our hypothesis predicts. Interestingly, the coefficients on the local sex ratio for the other types of households are also positive, but they are not statistically different from zero. In the last four columns, we exclude the outliers (those households with the smallest and the largest 5% of individual savings rates). The coefficient on the local sex ratio continues to be positive and statistically significant for households with a son. A 10 basis point increase in the local sex ratio (e.g., from 1.05 to 1.15) is associated with 13.6 percentage points increase in the savings rate by such households. The coefficients now become negative for households with one or two daughters, but still insignificant statistically.

The regressions for the urban sample are reported in Table 7. The coefficients on the local sex ratio are positive and significant both for households with a son and for those with a daughter. (Note that, since very few urban households have two children in the sample, it is not meaningful to do a regression on these households). The difference between the two is not statistically significant. The spillover effect could rationalize the positive coefficient for households with a daughter. At the same time, the quantitative effect of the spillover seems implausibly large.

In the last three columns, the top and the bottom 5% of outliers are removed. The results become more sensible. In particular, the local sex ratio continues to matter for households with a son: an increase in the local sex ratio by 10 basis points (e.g., from 1.05 to 1.15) would increase the savings rate of these households by about 11.6 percentage points. The coefficient for households with a daughter, while still positive, is no longer statistically different from zero. In any case, the point estimate for households with a son is more than twice as large as that for households with a daughter. Nonetheless, the fact that the coefficient is not negative for households with a daughter suggests that the spillover effect is likely present in urban areas. These patterns seem more in line with the prediction given by our hypothesis.

For the urban sample, we have opportunities to construct a set of proxies to detect a precautionary savings motive. We create an indicator variable for households with no access to public health insurance, one for those with at least one family member who has been laid off, one for those with at least one family member employed in a state-owned company, and one for those with at least one family member working in a company that has recently experienced a reorganization (and hence at risk of being laid off), and another one for households with a member working for an employer that has been losing money. In addition, we create a dummy for households that currently rent, rather than own, an apartment.

With the addition of these extra controls, the coefficient on the local sex ratio continues to be positive and significant: an increase in the sex ratio by 10 basis points (say, from 1.05 to 1.15) tends to be associated with an increase in the saving rate of about 10 percentage points. On the other hand, the coefficient on the same variable for households with a daughter becomes smaller (0.29) and is not statistically different from zero. This coefficient is still positive, consistent with the presence of a spillover effect.

With household level regressions, we can discuss and rule out some additional alternative explanations for the positive association between local sex ratios and local savings rate. Could the local sex ratio be correlated with some omitted or unobserved dimensions of quality of local social safety net, growth potential, or income uncertainty? A region with a high sex ratio imbalance may simply have an unusually high level of income uncertainty, or an unusually poor social safety net, that may be partially

addressed by having a son. This could be a problem for our earlier province-level panel regressions (before the IV approach). But in the current context, one may expect these variables to affect savings by all households (those with a daughter versus those with a son) in a given location in the same way. That is, all households should raise their savings rates in response to this region-wide shock. Yet, our results show that only savings by households with a son react strongly to local sex ratio, but savings by households with a daughter do not, especially after excluding outliers.

Could the omitted or unobserved variables be household (and location) specific? For example, a region may have an unusually high level of income uncertainty that is common to all households, but some households care about this uncertainty more than others in the same region. Those households which are particularly concerned with the local uncertainty are more likely to engage in a sex selective abortion in order to have a son and to save more at the same time. This may generate an appearance of a higher savings rate by households with a son. This is a much harder alternative to rule out. Ebenstein (2008) shows that sex ratio imbalance is overwhelmingly a result of sex selective abortions at higher orders of birth. That is, the sex ratio for first-born children has only a very mild degree of imbalance in China. However, the boys-to-girls ratio goes up substantially for the second-born children and become even more skewed for higher order births. This is consistent with the particulars of the family planning policy: since a couple in a rural area can legally have a second child if the first born is a girl, there is little need to perform sex selective abortion for the first born[9]. Yet, in Table 6, if we restrict attention to households with only one child, those with a son exhibit a positive elasticity of savings with respect to local sex ratio, but those with a daughter do not.

To gain further confidence in our interpretation that local sex ratios are not merely a proxy for other determinants of savings, it would be useful to check out how household savings may vary with respect to a wedding event. We turn to this next.

3.4 Additional evidence: wedding, housing and bank deposits

We provide three additional pieces of evidence on the time profile of household

[9] Ebenstein (2008) also reports a parental preference for "gender balance" when they have two children: while many parents prefer a son to a daughter, they prefer a combination of a son and a daughter to either two sons or two daughters.

savings rate with respect to the timing of a wedding event, regional variations in the size of a house, and regional variations in rural bank deposits.

We start with some suggestive evidence on how households savings change over time in relationship to a wedding event in the family, and how this differs for the groom's and the bride's families. The cultural norm is such that the groom's family provides a house or an apartment for newlyweds, or at least contributes the biggest chunk of the cost for a domicile. The groom's family is said to be responsible for paying the bride's family a one-time transfer that compensates the latter for rearing their daughter (Zhang and Chan, 1999). In addition, the groom's family bears most of the financial cost of holding a wedding ceremony although the bride's family may share some of the cost as well. Because weddings in China are occasions that call for significant cash outlays, families may have to save more before the weddings.

For 26 natural villages (or three administrative villages) in Guizhou Province, two rounds of household census were conducted in 2005 and 2007 by the International Food Policy Research Institute (IFPRI). In each round, all households were asked if there was a wedding event in the family any time over the previous three years and the related expenditures. This is addition to the usual list of questions asked in a household survey (such as household on demographics, income, consumption, and transfers). From this database, we are able to construct a time series profile of household savings rate with respect to the timing of wedding, i.e., the savings rate 3 years before the wedding, 2 years before the wedding, and so on. [In 2006, for example, the median wedding cost for the groom's family was 18,150 RMB, over eight times the per capita income in sample.]

Figure 3 plots the time profile of a representative household savings rate for groom and bride families, separately[10]. The horizontal axis stands for the number of years away from the time of a wedding. The vertical axis depicts the average household saving rate measured as (income-expenditure)/income*100. Because the number of households that have a wedding event is relatively small, we do not have enough statistical power to perform formal tests on the differences in the average savings rates across years or across household types. Nonetheless, the patterns in the figure are suggestive. First, the savings rate curves exhibit an inverse-V shape, peaking in the year before the wedding. (Because

[10] To reduce noise, the top and bottom of 5% outliers in terms of savings rates are dropped from the sample.

wedding event itself represents a major expenditure item, the savings rate for the year of wedding is lower). Household savings rates tend to be much lower after the wedding. Second, the savings rate curve for a groom's family lies almost everywhere above a bride's family except in the year after wedding. These features are consistent with the notion that a big chunk of household savings is motivated by preparation for a wedding, and savings for marriage is more important for a groom's family than for a bride's family. (The savings rate for a bride's family turns negative in some years after the wedding, possibly reflecting their consuming partly out of a transfer from the groom's family.)

A groom's family is generally expected to build a house in a rural area or buy an apartment in a city for a newly married couple. The logic of our hypothesis may imply that men or their families compete by providing a bigger house/apartment in a region with a more unbalanced sex ratio. We now examine this using data on average living space per household in the 2000 population census, with the regression results presented in Table 10.[11] It is clear that, holding constant household features (income and size in particular), the size of living space rises with the local sex ratio imbalance. This is true in both urban and rural areas. Somewhat surprisingly, the elasticity of living space with respect to local sex ratio is greater in the urban areas than in the rural areas. This could result from not having enough controls. But the paucity of control variables at the county level prevents us from investigating this further.

A key component of household savings takes the form of deposits at banks. For the year of 2002, we are able to compute actual bank deposits per person - or more precisely, local bank deposits in 2002, divided by local population in 2000 - for 1,972 rural counties. In Columns 1-4 of Table 10, we regress per-capita bank deposits by rural county on the local sex ratio and other controls. The first two regressions do not include province fixed effects, and the last two do. Columns 1 and 3 consider only the linear effect of the local income, whereas Columns 2 and 4 allow for a quadratic term for log income. The coefficients on the key regressor, sex ratio imbalance, are positive and statistically significant across all four specifications. Using Column 4 as an example, the point estimate is 1.09: an increase in the sex ratio by 10 basis points (e.g., from 1.10 to

[11] The summary statistics for the variables used in the county level analysis are provided in Table 1b.

1.20) is associated with an increase by ten percentage points in households' bank deposit balances.

4. Conclusions:

This paper proposes a new explanation for savings and tests it using data from China: a rising sex ratio imbalance, by increasing competition among men for potential wives, can stimulate households with a son to postpone consumption in favor of wealth accumulation. We provide supportive evidence from both panel regressions across Chinese regions and household level regressions.

Across Chinese provinces, there is a strong positive association between the local sex ratios for the marriage-age cohort and the local household savings rate. The association is stronger for rural areas than for urban areas. The point estimates suggest that approximately 68% of the increase in rural savings rates, and 18% of the increase in urban savings rates in the recent years can be attributed in the rise in sex ratios. This pattern continues to hold when we apply instrumental variables to the local sex ratios, by exploring regional variations in the financial penalties for violating the family planning policy and in the proportion of local population that is exempted from the birth quotas.

An examination of the household level savings provides additional confirmation of hypothesis. Households with a son tend to save more in regions with a more skewed male/female ratio, holding constant household size, income level, and the age and educational level of the household heads. We interpret this as reflecting extra savings stimulated by a greater competition for wives. On the other hand, households with a (or two) daughter(s) do not reduce their savings in response to an increase in the local sex ratio imbalance. We interpret this as reflecting a spillover effect that bids up the prices of goods and services such as housing for households with a daughter by competition among men (or their families).

Accumulating more wealth is not the only way for men to compete in the marriage market. Parents may also invest more in the human capital of their sons, and may also engage in more entrepreneurship, and other higher risk, higher returns activities. Finally, while the paper focuses on evidence from China, the basic mechanism can in

principle be applied to other countries. We leave a systematic examination of other consequences of sex ratio imbalance and of international data to future projects.

References:

Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

Aziz, Jahangir, and Li Cui, 2007**, "**Explaining China's Low Consumption: The Neglected Role of Household Income," IMF working paper 07/181. http://www.imf.org/external/pubs/cat/longres.cfm?sk=21026.0

Beck, Thomas and Ed Al-Hussainy, Financial Structure Dataset, The World Bank/DECRG-FI, Washington. August 2007.

Blanchard, Olivier J. and Francesco Giavazzi, 2005. "Francesco,Rebalancing Growth in China: A Three-Handed Approach." MIT Department of Economics Working Paper No. 05-32. Available at SSRN: http://ssrn.com/abstract=862524.

Chamon, Marcos, and Eswar Prasad, 2008, "Why are Saving Rates of Urban Households in China Rising?" working paper, IMF and Cornel University.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite, 1992, "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy*, 100(6): 1092-1125, December.

Department of Economic and Social Affairs. (2002). *Demographic Yearbook 2000*. New York: United Nations: 112-124.

Department of Economic and Social Affairs. (2007). *Demographic Yearbook 2005*. New York: United Nations: Table 3: 1-14.

Gu, Baochang and Krishna Roy, "Sex Ratio at Birth in China, with Reference to Other Areas in East Asia: What We Know," *Asia-Pacific Population Journal*, Vol. 10, No. 3 (1995), pp. 17-42.

Das Gupta, Monica, 2005, "Explaining Asia's 'missing women': A new look at the data," Population and Development Review 31(3): 529-535.

Ebenstein, Avraham, 2008, "Estimating a dynamic model of sex selection in China," working paper, Harvard University.

Edlund, Lena, 2001, "Dear son – expensive daughter: Do scarce women pay to marry?" Unpublished.

Edlund, Lena, Hongbin Li, xxx, 2008, "Sex ratios and crime: evidence from China's one-child policy," IZA DP no. 3214.

Guilmoto, Christophe Z. 2007. "Sex-ratio imbalance in Asia: Trends, consequences and policy responses." Available from www.unfpa.org/gender/docs/studies/summaries/regional_analysis.pdf.

He, Xinhua and Yongfu Cao, "Understanding High Saving Rate in China," *China & World Economy* / 1 – 13, Vol. 15, No. 1, 2007.

Hopkins, Ed, and Tatiana Kornienko, 2004, "Running to Keep in the Same Place: Consumer Choice as a Game of Status," American Economic Review 94(4): 1085-1107.

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.

Li, Shuzhuo, 2007. "Imbalanced Sex Ratio at Birth and Comprehensive Interventions in China," paper presented at the 4th Asia Pacific Conference on Reproductive and Sex Health and Rights, Hyderabad, India, October 29-31.

Lin, Ming-Jen and Ming-Ching Luoh, 2008, "Can Hepatitis B Mothers Account for the Number of Missing Women? Evidence from Three Million Newborns in Taiwan," American Economic Review, forthcoming, 2008.

Lin, Ming-Jen, and Ming-Ching Luoh, 2006, "Can Hepatitis B mothers account for the number of missing women? Evidence from three million newborns in Taiwan," National Taiwan University Faculty Working Paper.

Modigliani, Franco, 1970, "The Life Cycle Hypothesis of Saving and Intercountry Differences in the Saving Ratio," Introduction, *Growth and Trade, Essasys in Honor of Sir roy Harrod*, W.A. Elits, M.F. Scott, and J.N. Wolfe, eds. Oxford University Press.

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

Oster, Emily, 2005, "Hepatitis B and the case of the missing women," *Journal of Political Economy* 113(6): 1163-1216.

Oster, Emily, Gang Chen, Xinxen Yu, and Wenyao Lin, 2008, "Hepatitis B Does Not Explain Male-Biased Sex Ratios in China," University of Chicago working paper. http://home.uchicago.edu/~eoster/hbvnotecon.pdf

Population Divison, Department of Economic and Social Affairs. (2007). World Population Prospects: The 2006 Revision. New York: United Nations.

Qian, Nancy, 2008, "Missing Women and the Price of Tea in China: The Effect of Sex-specific Earnings on Sex Imbalance," *Quarterly Journal of Economics* 123(3), August.

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

Ravallion, Martin, 2008. "Geographic Inequity in a Decentralized Anti-Poverty Program: A Case Study for China," forthcoming in *World Bank Economic Review*.

Scharping, Thomas, 2003, *Birth Control in China 1949-2000*, New York, NY: Routledge Curzon.

Statistical Office, Department of Economic and Social Affairs. (1976). *Demographic Yearbook 1975*. New York: United Nations: 432-487.

Statistical Office, Department of International Economic and Social Affairs. (1988). *Demographic Yearbook 1986*. New York: United Nations: 474-525.

United Nations Statistics Division. (October, 2007) Gender Info 2007. Retrieved June 18, 2008, from http:// http://www.devinfo.info/genderinfo/.

Annoymous, 2008, "An expose of the cost of getting married in nine Chinese cities." http://news.sohu.com/20080604/n257268273.shtml.

Weckerly, Floyd W., 1998, "Sexual-Size Dimorphism: Influence of Mass and Mating Systems in the Most Dimorphic Mammals," *Journal of Mammalogy*, Vol. 79, No. 1 (Feb., 1998), pp. 33-52, Published by: American Society of Mammalogists. http://www.jstor.org/pss/1382840.

Zhang, J., and W. Chan, 1999. "Dowry and Wife's Welfare: A Theoretical and Empirical Analysis." *Journal of Political Economy* 107(4): 786-808.

Table 1a: Summary Statistics for Key Variables in Provincial Panel Regressions

	Rural					Urban				
Variables	Mean	Median	Min	Max	Std	Mean	Median	Min	Max	Std
1980										
Savings rate = log(income/expenditure)	0.16	0.17	-0.03	0.29	0.07	0.17	0.10	-0.05	2.27	0.43
Sex ratio for age cohort 6-24	1.06	1.07	0.98	1.15	0.04	1.08	1.08	1.03	1.16	0.04
1990										
Savings rate = log(income/expenditure)	0.19	0.19	0.01	0.52	0.11	0.14	0.156	-0.04	0.21	0.06
Sex ratio for age cohort 6-24	1.07	1.06	0.96	1.20	0.06	1.05	1.04	0.95	1.20	0.05
2000										
Savings rate = log(income/expenditure)	0.31	0.29	0.12	0.78	0.14	0.21	0.21	-0.08	0.343	0.07
Sex ratio for age cohort 6-24	1.09	1.07	1.03	1.21	0.04	1.07	1.06	0.92	1.17	0.04
2006										
Savings rate = log(income/expenditure)	0.25	0.26	0.03	0.62	0.12	0.30	0.29	-0.48	0.764	0.22
Sex ratio for age cohort 6-24	1.11	1.13	1.03	1.22	0.05	1.09	1.08	1.00	1.23	0.05
1980-2006										
Per capita income (log)	6.90	6.92	4.96	9.12	0.95	7.78	7.88	5.84	9.94	1.07
Share of population younger than 20	0.34	0.34	0.18	0.45	0.04	0.36	0.37	0.20	0.50	0.05
Share of population aged 20-59	0.41	0.41	0.18	0.65	0.11	0.43	0.44	0.15	0.66	0.13
Share of labor force enrolled in social security						0.07	0.00	0.00	0.97	0.15
Share of labor force employed in SOEs						0.21	0.16	0.05	0.58	0.12

Note: Various sources. The sex ratios for the age cohort 6-24 in all years are derived from the 2000 population census. To be precise, for the year 2000, we know the exact sex ratio for this age cohort from the census. For the age cohort 6-24 in 2001, we infer the sex ratio with data for the age cohort 5-23 in the 2000 census, since the two groups should theoretically be the same. Similarly, for the age cohort 6-24 in 2006, we match it with the cohort 0-18 in the 2000 census; for 1990, we match it with the cohort 16-34 in the 2000 census; and so on.

Table 1b: Summary Statistics on the Key Variables Used in County Analysis

Variables	Mean	Median	Min	Max	Std
Per capita deposit in 2002 (Yuan)	24,501	7,276	0.000	2,737,933	98,234
Per capita GDP in 1999 (Yuan)	7,442	4,171	347	1,113,476	30,573
Housing area per household in 2000 (sq meter)	81.011	78.131	19.968	268.785	22.955
Household size in 2000 (person)	3.779	3.730	2.357	7.362	0.556
Sex ratio for cohort of 0-9 in the 1990 census	1.082	1.072	0.778	2.509	0.059
Share of population under 20 in 2000 census	0.313	0.304	0.162	0.997	0.063
Share of population aged 20-59 in 2000 census	0.597	0.598	0.003	0.828	0.067

Source: Authors' calculations based on 1990 and 2000 censuses.

Table 2: Sex Ratios and Savings Rates across Provinces: Panel Fixed Effects Regressions 1978-2006

LHS variable = savings rate	1	2	3	4	5	6
	Rural	Rural	Urban	Urban	Urban	Urban
Per capita income (log)	0.521*	-0.617	0.469**	1.161***	0.458***	1.152***
	(0.31)	(1.00)	(0.19)	(0.26)	(0.17)	(0.26)
Per capita income (log) squared	0.009	0.062	-0.004	-0.051***	-0.004	-0.051***
	(0.01)	(0.06)	(0.01)	(0.02)	(0.01)	(0.02)
Sex ratio (6-24)	0.831***	2.195*	0.585***	0.523*	0.557***	0.540*
	(0.24)	(1.19)	(0.17)	(0.30)	(0.16)	(0.30)
Share of population aged 0-24	0.643**	0.668	0.601*	-0.618	0.506	(0.484)
	(0.30)	(0.79)	(0.36)	(0.42)	(0.48)	(0.49)
Share of population aged 25-60	1.872**	2.288	0.435	-1.391**	0.305	-1.237*
	(0.87)	(2.41)	(0.60)	(0.65)	(0.62)	(0.70)
Gini coefficient		0.621		0.058		0.024
		(0.86)		(0.16)		(0.17)
Share of labor force enrolled in social security					0.042	
					(0.07)	
Share of SOE employment in total labor force					0.156	-0.155
					(0.25)	(0.17)
Provincial fixed effects	yes	Yes	yes	yes	yes	yes
Year fixed effects	yes	Yes	yes	yes	yes	yes
Adjusted R-squared	0.441	0.205	0.397	0.816	0.396	0.816
AIC	-638.3	-9.7	-1406.0	-498.3	-1403.5	-497.3
N	811	141	798	141	798	141

Note: The Both provincial and year fixed effects are included but not reported here. Saving rate is defined as log(net income/living expenditure). The sex ratios for the age cohort 6-24 for all years are inferred from the 2000 census. For example, the cohort 6-24 in 2006 is the same as those of 0-18 in 2000; and the cohort 6-24 in 1990 is the same as those of 16-34 in 2000. The Gini coefficients for urban and rural areas by province are from Ravallion and Chen (2007) and only available for limited provinces in 1988, 1990, 1993, 1996 and 1999. The employment data prior to 1999 is from *Comprehensive Statistical Data and Materials on 50 Years of New China* (CNBS) while the data in later years are from various issues of *China Statistical Yearbooks*. Social security enrollment data since 2000 is available from *China Statistical Yearbooks*. Robust standard errors are in parentheses. *, **, and *** stand for significant level at 10%, 5%, and 1%, respectively.

Table 3: First Stage Regressions of 2SLS: Determinants of Local Sex Ratios

	1	2	3	4
	Rural	Rural	Urban	Urban
Share of minority in local population	-0.217***	-0.247***	-0.312***	-0.349***
(% population not subject to strict family planning)	(0.03)	-0.033	(0.04)	(0.04)
Penalty for violating family planning policy	0.019***	0.025**	0.016***	0.012
(% of local yearly income)	-0.003	-0.012	-0.004	-0.011
Dummy for extra penalty for higher order births	0.044***	0.038***	0.020***	0.016**
	(0.01)	-0.007	(0.01)	(0.01)
Log per capita income		0.110***		0.497***
		-0.032		(0.07)
Log per capita income squared		-0.005**		-0.032***
		-0.002		(0.00)
Share of population aged 0-24		0.157**		-0.007
		-0.07		(0.12)
Share of population aged 25-60		-0.217**		0.06
		-0.108		(0.18)
Provincial fixed effects	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes
Adjusted R-squared	0.827	0.843	0.668	0.730
AIC	-3959.34	-4011.29	-3530.13	-3669.71
N	811	811	798	798

Table 4: Sex Ratios and Saving Rates - IV Regressions, 1978-2006

LHS variable = local saving rate	1	2
	Rural	Urban
Per capita income (log)	0.539*	0.509***
	(0.31)	(0.17)
Per capita income (log) squared	0.008	-0.007
	(0.01)	(0.01)
Sex ratio for the age cohort 6-24	0.724*	0.501*
	(0.39)	(0.30)
Share of population aged 0-24	0.628**	0.601*
	(0.32)	(0.34)
Share of population aged 25-60	1.805*	0.431
	(1.00)	(0.58)
Provincial fixed effects	yes	yes
Year fixed effects	yes	yes
Adjusted R-squared	0.441	0.397
AIC	-638.2	-1420.0
Durbin-Wu-Hausman test for endogeneity	0.839	0.837
Hansen's J statistic for over identification	0.471	0.052
N	811	803

See footnotes to Table 2 for data sources. Robust standard errors are in parentheses. The symbols *, **, and *** stands for significant level at 10%, 5%, and 1%, respectively.

Table 5: Saving Rates in Rural and Urban Households in 2002

	Household type	Mean	Median	Max	Min	Standard deviation	N
Rural							
	One son	0.373	0.371	2.462	-2.986	0.597	599
	One girl	0.332	0.364	1.812	-3.559	0.613	243
	Two girls	0.277	0.348	1.973	-2.125	0.631	188
	One son and one girl	0.278	0.281	2.846	-3.709	0.571	835
	Two sons	0.293	0.327	1.851	-1.445	0.511	294
	Total rural	0.316	0.316	2.846	-5.026	0.582	9199
Urban							
	One son	0.279	0.270	1.174	-0.693	0.301	277
	One girl	0.261	0.264	1.327	-1.299	0.361	311
	Two girls	0.367	0.384	0.617	0.047	0.239	6
	One son and one girl	0.340	0.207	0.940	-0.354	0.358	15
	Two sons	0.369	0.111	1.502	-0.123	0.652	5
	Total urban	0.306	0.287	2.308	-2.441	0.378	6835

Note: The savings rate is defined as log(income/consumption). The data comes from the Chinese Household Income Project (2002), available from http://www.icpsr.umich.edu/cocoon/ICPSR/STUDY/03012.xml. To maximize comparability, we restrict the sample to households with both parents still alive, and head of households younger than 40.

Table 6: Household-level Savings in Rural China in 2002

	With all observations				With 10% outliers removed			
	Son	Daughter	Two daughters	Son and daughter	Son	Daughter	Two daughters	Son and daughter
Per capita income (log)	2.907***	3.359***	2.034***	1.891***	1.963***	2.808***	1.05	1.036***
	(0.57)	(0.66)	(0.55)	(0.39)	(0.63)	(0.79)	(0.78)	(0.36)
Per capita income (log) squared	-0.151***	-0.181***	-0.094**	-0.092***	-0.103***	-0.156***	(0.04)	-0.048**
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.05)	(0.02)
Household head age	-0.006	0.004	-0.013	-0.002	-0.001	0.002	-0.006	-0.003
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
Child aged 5-9	0.226***	0.199**	-0.017	0.079**	0.181***	0.112	0.049	0.068**
	(0.08)	(0.09)	(0.09)	(0.04)	(0.06)	(0.09)	(0.06)	(0.03)
Child aged 10-14	0.178***	0.127	-0.039	0.029	0.165***	0.050	-0.003	0.040
	(0.06)	(0.09)	(0.09)	(0.04)	(0.05)	(0.08)	(0.07)	(0.03)
Household head gender (Female =1)	-0.014	-0.282	-0.183	-0.136	-0.043	-0.076	-0.154	-0.057
	(0.08)	(0.21)	(0.17)	(0.09)	(0.07)	(0.16)	(0.17)	(0.07)
Household head year of schooling	0.002	-0.025*	-0.032**	-0.020**	-0.008	-0.023*	-0.040***	-0.017***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Household head as a minority	-0.242**	-0.245*	-0.076	-0.042	-0.089	-0.055	0.101	0.029
	(0.10)	(0.12)	(0.14)	(0.06)	(0.10)	(0.09)	(0.12)	(0.04)
Poor health	-0.011	-0.015	0.078	-0.124	0.030	-0.073	0.021	-0.106**
	(0.10)	(0.12)	(0.18)	(0.10)	(0.08)	(0.08)	(0.15)	(0.05)
Sex ratio at the county level	1.071**	0.464	0.064	0.295	1.355***	-0.052	-0.359	0.160
	(0.52)	(0.70)	(0.77)	(0.38)	(0.44)	(0.56)	(0.57)	(0.29)
Gini at the county level	-0.910**	-1.200**	-0.602	-0.801***	-0.462	-0.886**	-0.492	-0.863***
	(0.43)	(0.53)	(0.66)	(0.27)	(0.35)	(0.39)	(0.49)	(0.21)
Adj. R-squared	0.287	0.532	0.403	0.365	0.175	0.289	0.317	0.229
AIC	890.8	278.9	275.1	1064.5	472.2	124.1	108.4	487.2
N	599	243	188	835	533	220	163	770

Note: The savings rate is defined as log(income/consumption). The data comes from the Chinese Household Income Project (2002), available from http://www.icpsr.umich.edu/cocoon/ICPSR/STUDY/03012.xml. To maximize comparability, we restrict the sample to households with both parents still alive, and head of households younger than 40. "Poor health" is a dummy that takes the value of one if a household has at least one member with disability or extreme bad health. The sex ratio at the county level is calculated by authors based on the cohort of 0-9 from China Population Census 1990 (who aged 12-21 in 2002). Robust t statistics in parentheses; *, **, and *** denote significant at 10%, 5%, and 1% levels, respectively.

Table 7: Household Savings in Urban China in 2002

	With all the observations			Sample with 10% outliers removed		
	Son	Daughter	Combined	Son	Daughter	Combined
Per capita income (log)	0.968	1.02	1.056*	1.537**	0.168	0.611
	(0.84)	(0.65)	(0.56)	(0.62)	(0.49)	(0.44)
Per capita income (log) squared	-0.045	-0.045	-0.049	-0.081**	-0.001	-0.026
	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Household head age	0.008	-0.006	-0.002	-0.003	-0.006	-0.005
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)
Child aged 5-9	0.161*	0.155**	0.146**	-0.001	0.012	0.012
	(0.09)	(0.08)	(0.06)	(0.08)	(0.06)	(0.04)
Child aged 10-14	0.159**	0.172**	0.164***	0.010	0.059	0.042
	(0.08)	(0.08)	(0.06)	(0.06)	(0.05)	(0.04)
Household head gender (Female =1)	-0.069*	-0.119***	-0.097***	-0.053	-0.030	-0.037*
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
Household head year of schooling	-0.010	-0.013*	-0.012**	-0.009	-0.006	-0.008*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Household head as a minority	0.080	0.029	0.064	-0.048*	-0.034	-0.025
	(0.08)	(0.10)	(0.06)	(0.03)	(0.09)	(0.05)
Poor health	-0.109	-0.048	-0.079	-0.034	-0.015	-0.036
	(0.08)	(0.08)	(0.05)	(0.06)	(0.05)	(0.04)
Having a son			0.008			-0.011
			(0.03)			(0.02)
Sex ratio at the city level	1.815***	1.744***	1.800***	1.155**	0.459	0.723**
	(0.51)	(0.52)	(0.37)	(0.46)	(0.42)	(0.31)
Gini at the county level	-0.332	-0.697	-0.519	-0.128	-0.492	-0.315
	(0.42)	(0.58)	(0.35)	(0.37)	(0.39)	(0.27)
Adj. R-squared	0.105	0.161	0.149	0.054	0.090	0.079
AIC	102.7	205.5	296.8	4.2	13.3	3.4
N	277	311	588	259	287	546

Note: See previous table for variable definitions and data sources.

Table 8: Household-level Savings in Urban Areas in 2002, with Additional Controls

	With all the observations			Sample with 10% outliers removed		
	Son	Daughter	Combined	Son	Daughter	Combined
Per capita income (log)	1.126	1.266**	1.214**	1.625**	0.443	0.779*
	(0.88)	(0.62)	(0.55)	(0.65)	(0.46)	(0.43)
Per capita income (log) squared	-0.054	-0.059*	-0.057*	-0.086**	-0.016	-0.036
	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
Household head age	0.007	-0.005	-0.002	-0.005	-0.004	-0.004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Child aged 5-9	0.132	0.156**	0.139**	-0.023	0.018	0.006
	(0.09)	(0.07)	(0.06)	(0.08)	(0.06)	(0.04)
Child aged 10-14	0.142*	0.176**	0.162***	0.001	0.058	0.039
	(0.08)	(0.08)	(0.05)	(0.07)	(0.06)	(0.04)
Household head gender (Female =1)	-0.063*	-0.120***	-0.090***	-0.051	-0.027	-0.033
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
Household head year of schooling	-0.011	-0.014*	-0.013**	-0.010	-0.006	-0.008
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Household head as a minority	0.080	0.021	0.057	-0.046	-0.029	-0.022
	(0.08)	(0.09)	(0.06)	(0.03)	(0.09)	(0.05)
Poor health	-0.114	-0.055	-0.087*	-0.042	-0.022	-0.042
	(0.07)	(0.07)	(0.05)	(0.06)	(0.05)	(0.04)
Without public insurance	0.046	0.080*	0.060**	0.045	0.067**	0.053**
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
With a family member laid off	0.056	0.062	0.060	0.011	0.045	0.032
	(0.05)	(0.06)	(0.04)	(0.05)	(0.05)	(0.03)
With a family member in SOE	0.034	0.041	0.038	0.035	0.062**	0.051**
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
Experienced firm reorganization	0.010	-0.100**	-0.048	-0.002	-0.007	-0.010
	(0.04)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)
Employed firm losing money	-0.067	-0.045	-0.055*	-0.046	-0.037	-0.037
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.02)
Renting a house	-0.076*	-0.050	-0.062**	-0.068**	-0.091**	-0.082***
	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)
Having a son			0.008			-0.010
			(0.03)			(0.02)
Sex ratio at the city level	1.604***	1.706***	1.688***	1.052**	0.291	0.586**
	(0.48)	(0.50)	(0.36)	(0.44)	(0.39)	(0.29)
Gini at the city level	-0.182	-0.609	-0.378	-0.017	-0.309	-0.169
	(0.41)	(0.57)	(0.35)	(0.37)	(0.38)	(0.26)
Adj. R-squared	0.114	0.179	0.166	0.058	0.114	0.103
AIC	105.6	204.6	291.1	8.7	11.5	-5.1
N	277	311	588	259	287	546

Note: See previous table for variable definitions and data sources.

Table 9: The Impact of Sex Ratio on Per Capita Living Space

LHS Variable = Per Capita Size of Living Space in 2000

	City	City	County	County
Per capita GDP in 1999 (log)	0.10**	-0.44*	0.07**	-0.09
	(0.01)	(0.25)	(0.01)	(0.15)
Per capita GDP in 1999 (log) squared		0.03**		0.01
		(0.01)		(0.01)
Sex ratio	0.70**	0.71**	0.19**	0.20**
	(0.16)	(0.16)	(0.09)	(0.09)
Household size (log)	1.43**	1.47**	0.14**	0.14**
	(0.12)	(0.13)	(0.02)	(0.02)
Share of population aged 25-59	0.64**	0.70**	1.39**	1.39**
	(0.21)	(0.21)	(0.15)	(0.15)
Share of population aged 60 and above	3.30**	3.33**	3.67**	3.65**
	(0.34)	(0.33)	(0.31)	(0.30)
Adjusted R-squared	0.39	0.39	0.17	0.18
AIC	-323.3	-326.4	45.5	46.105
N	670	670	2089	2089

Note: Per capita living space is calculated by authors based on *China Population Census 2000*. The resident bank deposit and per capita GDP are from various issues of China *County Social and Economics Statistical Yearbooks*. The sex ratio is calculated based on the cohort of 0-9 in the 1990 census who aged 10-19 in 2000. The share of population aged 0-24 and 25-59 are calculated from the 2000 census. The symbols * and ** stand for significant level at 10% and 5%, respectively.

Table 9: Sex Ratios and Residential Bank Deposits across Rural Counties

LHS Variable = Per Capita Residential bank deposits in 2002

	2002	2002	2002	2002	1992-2002	1992-2002
Per capita GDP in 1999 (log)	0.63**	-2.24**	0.50**	-2.15**		
	(0.04)	(0.52)	(0.04)	(0.56)		
Per capital GDP in 1999 (log) squared		0.17**		0.16**		
		(0.03)		(0.03)		
Sex ratio for age cohort 12-21	2.22**	2.39**	0.97**	1.09**	0.49**	0.39**
	(0.75)	(0.79)	(0.41)	(0.43)	(0.19)	(0.18)
Share of population aged 0-24	11.57**	11.68**	15.30**	15.32**	-0.09	-0.20
	(1.14)	(1.14)	(1.14)	(1.14)	(0.23)	(0.30)
Share of population aged 25-59	22.7**	22.8**	28.0**	28.0**	0.03	-0.13
	(1.27)	(1.28)	(1.27)	(1.27)	(0.16)	(0.18)
Provincial fixed effects			yes	yes		Yes
Adjusted R-squared	0.41	0.42	0.50	0.51	0.01	0.07
AIC	5481	5456	5199	5177	2595	2494
N	1972	1972	1972	1972	1886	1886

Note: The residential bank deposit and per capita GDP are from *China County Social and Economics Statistical Yearbooks* (CNBS). For the first four regressions, the sex ratio is calculated based on the cohort of 0-9 in the 1990 census who aged 12-21 in 2002. The shares in population for age cohorts 0-24 and 25-59 are derived from the 2000 census. For the last two regressions, the sex ratio and share of population variables refer to their changes from 1990 to 2000 inferred from the 1990 and 2000 censuses.

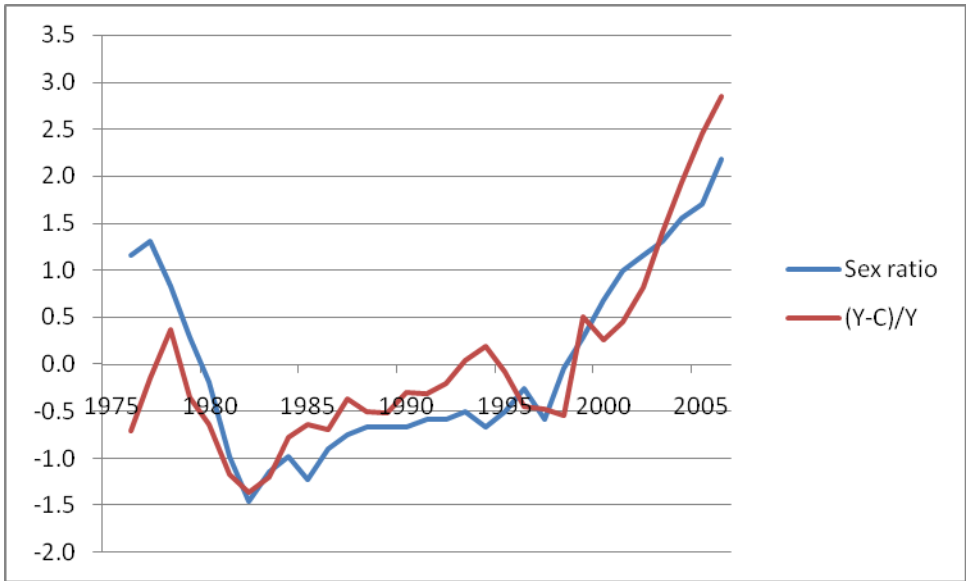

Figure 1 Sex Ratio and Saving Rate

Note: The sex ratio variable is defined as the sex ratio at birth 20 years ago. See the note of Table 1 for data sources. The saving rate is defined as the percentage of (GDP-private and government consumption) in total GDP, which is available from China Statistical Yearbook 2007. Both variables have been rescaled by subtracting the mean and dividing by the standard deviation.

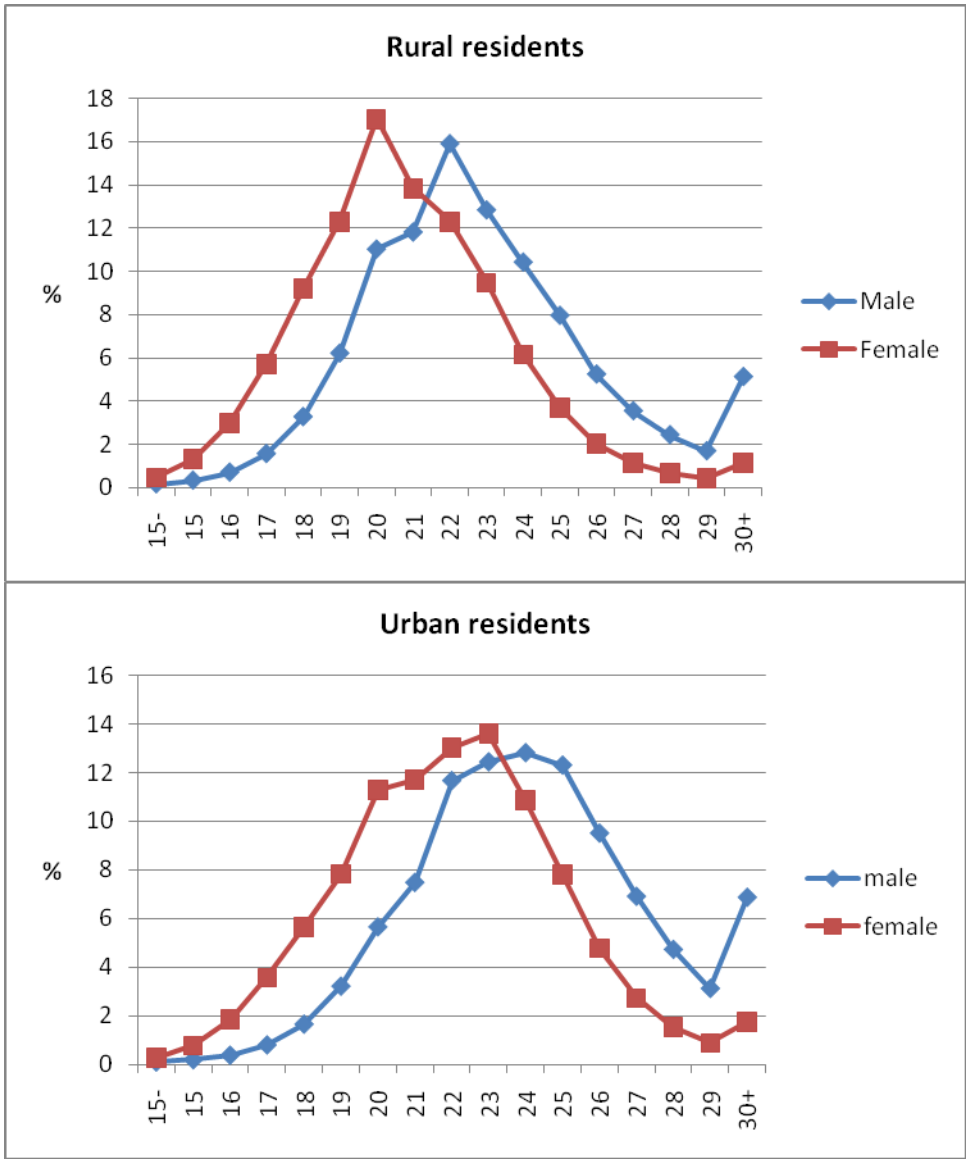

Figure 2 Age Distribution of First Marriage in China in 2000

Note: Calculated by authors based on *China Population Census 2000*.

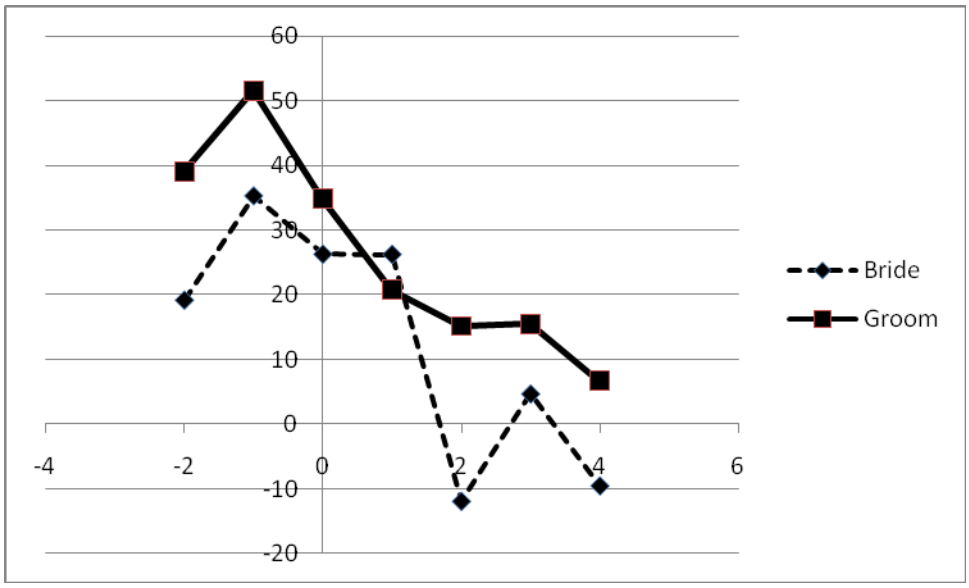

Figure 3: Time Profile of Household Savings Rate in relation to the Timing of a Wedding: Evidence from 26 Natural (3 Administrative) Villages in Guizhou Province

Note: Authors' calculation based on surveys conducted by IFPRI in 2005 and 2007.