Competitive Saving Hypothesis Revisited

Yongzhi Sun Liwei Yin*

September 16, 2019

Abstract

In Volume 119 of the Journal of Political Economy, Wei and Zhang (2011a) use data from the Chinese Household Income Project (CHIP) 2002 and find that a explanation called "competitive saving motive" that increasingly unbalanced premarital sex ratio raise household saving rate of son-families and the rapid increase in premarital sex ratio can potentially explain about half of China's household saving rate increasing during 1990-2007. This paper reexamines the competitive saving motive. We first use local sex ratio inferred from 2000 China population census and same dataset CHIP 2002 to find the competitive saving motive only holds for the household in rich counties. We then use data from the China Household Finance Survey (CHFS) to show that competitive holds for rural sample. The cross-regional evidence indicates that the competitive saving motive exist, but only in the rural area. By estimation and computation, an increase in sex ratio from 1985 to 2015 can explain about 28% of the actual increase of the increase of rural saving rate.

Keywords: China's household saving rate, sex ratio, competitive saving motive.

*Sun: Department of Economics, Texas A&M University, 4228 TAMU, College Station, TX 77843, United States. yongzhisun@tamu.edu. Yin: Department of Economics, Texas A&M University, 4228 TAMU, College Station, TX 77843, United States. yinliwei@tamu.edu. We would like to thank Li Gan, Yuzhe Zhang, Yonghong An, Jason Lindo, and Ren Mu. All errors are our own.

1 Introduction

The high aggregate household saving rate is one of the unique feature of the Chinese economy. According to National Accounts of OECD Countries, China's household savings as a percentage of household disposal income have been one of the highest in the world. The household saving rate has been increasing steadily from 28 percent in 2000 to 37 percent in 2015.[1] Over the period of China rapid income growth, China's sex ratio, which is defined as the ratio of the number of boys at birth to the number of girls at birth, has been also experienced a rapid a growth, from 1.07 in 1982 to 1.15 in 2015.[2] The common trend has captured economists' attention on linking the saving rate to the sex ratio.[3]

Can the sex ratio explain the high China's household saving rate? The first study that links the two variables together in explaining the "Chinese saving puzzle" is Wei and Zhang (2011a).[4][5] They argue a explanation called "competitive saving motive" that increasingly unbalanced premarital sex ratio require the both rural and urban son-families to save increasingly more to ensure the success in an increasingly competitive marriage market and increase the saving rate of the household factor. Wei and Zhang (2011a) use data from the Chinese Household Income Project (CHIP) data and provincial-level panel data in China to show that the rapid increase in premarital sex ratio during 1990-2007 can potentially explain about half of the sharp increase in China's household saving rate.

Why competitive saving motive need this replication In fact, birth and premarital sex ratio imbalance both rural and urban is still important in current China. Migration for working and schooling is another historical process for the youth during the rapid industrialization and urbanization in China. The pattern of unbalanced sex ratio for rural and urban can be reshaped during the intra-provincial and inter-provincial migration along with urbanization process over the last decades with Chinese rapid economy growth. The saving rate of Chinese household sector has been increasing steadily from 28 percent to 37

[1] See panel (1) of Figure 1 in Appendix.

[2] See panel (2) of Figure 1 in Appendix

[3] See Figure 2 in Appendix.

[4] Modigliani and Cao (2004) refers to the high China's household saving rate the "Chinese saving puzzle".

[5] To date, other explanations in the literature includes demographic changes (Modigliani and Cao, 2004; Horioka and Wan, 2007; Curtis, Lugauer, and Mark, 2015, etc), precautionary saving motive (Meng, 2003; Chamon and Prasad, 2010; He, Huang, Liu, and Zhu, 2017, etc), habit formation (Horioka and Wan, 2007; Carroll and Weil, 1994). Since our paper focuses on reconsidering the gender imbalance and the competitive saving motive, starting with Wei and Zhang (2011a), in explaining Chinese household saving rate, we do not expand space to review other important explanations and leave them in a review paper by Yang, Zhang, and Zhou (2012).

percent since Wei and Zhang (2011a). Does the competitive saving motive of premarital aged son-families still holds in China today? or the competitive effect is still a factor for some part of China? Because of the widely citation around the world and important implications for other Asian countries experiencing the same unbalanced premarital sex ratio other than China, Horioka and Terada-Hagiwara (2017) revisited the competitive saving motive and find a significant impact on the household saving rate using long-term time series data for the 1975–2010 period in both India and Korea. Many countries, including India, Korea, Singapore, Vietnam, Taiwan, in fact, have also exhibited unbalanced sex ratios in the premarital age cohorts. Accordingly, the competitive saving motive may have played a quantitatively important role in the evolution of these countries' savings rates as well. In addition, though it is not as easy to estimate its effect, the competitive saving motive can still be present and important in countries with a balanced sex ratio. Therefore, in this paper, we still focus on China's premarital sex ratio long-term pattern and reexamine the impact of local sex ratio on the household saving rate.

First, we use the CHIP data — the same sources of data as used by Wei and Zhang (2011a) — in an attempt to replicate and extend the estimates reported in Wei and Zhang (2011a). We first show that replication of Wei and Zhang (2011a) and analysis based on CHIP 2002 along with 1990 China population census. We also do the extensive robustness check with other county–level characteristics. Moreover, we use local sex ratio inferred from 2000 China population census and find the competitive saving motive only holds for the household in rich counties. The effects vanish for CHIP 2002 full sample with census data in 2000, especially for poor households and households in poor counties. We argue that unbalanced migration of premarital male and female is the major explanation for the vanishing effect.

Next, we turn to another nationally representative data set for estimating the effect of the sex ratio on household saving rate. In particular, we use the China Household Finance Survey (CHFS) 2015 wave, which covers a larger sample of individual and households and contains more detailed information about income, assets, debts, and expenditures. Estimated effects of the sex ratio on household saving rate based on the CHFS 2015 indicate that the competitive saving motive still hold, but only in the rural area. Specifically, in contrast to the estimates based on the CHIP 2002, in the rural, the effects on households with a son is much smaller, and the effects on households with a daughter are also significant positive. However, there is no effects on urban households according to the estimates based on CHFS 2015.

Finally, we use data from Population Census (1990, 2000, 2010), China Statistical

YearBook 1986–2016, and Comprehensive Statistical Data and Materials on 50/55/60 years of New China (CNBS) — also the same data sources as Wei and Zhang (2011a) used, and follow the same specification to reexamine the effect of sex ratio on the saving rate at province level. We adopt a more precise algorithm to compute the sex ratios of the 31 provinces for 7 years (1985, 1990, 1995, 2000, 2005, 2010, and 2015). Our results indicate that the competitive saving motive exist, but only in the rural area. The effect of sex ratio on the saving rate is insignificant different from zero for both full sample and the urban sample. By estimation and computation, an increase in sex ratio from 1985 to 2015 can explain about 28% of the actual increase of the increase of rural saving rate.

Related literature on sex ratio After two widely cited papers Wei and Zhang (2011a) and Du and Wei (2013), Wei and Zhang (2016) has a useful survey on recent literature on the competitive saving motive and broader economic implications. As a across country's evidence, Du and Wei (2016) report patterns of sex ratio imbalances and private-sector saving rate across countries and show that greater gender imbalance tends to correspond with higher saving rate. The result verifies the theory presented in the Du and Wei (2013). To go beyond the cross-country evidence and examine household-level evidence is quite useful. Wei, Zhang, and Liu (2017) test the effect of local sex ratio on home ownership and home prices. The paper finds son-families in area with more elevated sex ratio are more likely to own a home and concludes that rise in sex ratios contribute nearly half of the rise in cities' housing prices in China. Edlund, Li, Yi, and Zhang (2013) show that unbalanced sex ratios result in higher crime rates in property and indicate that rising sex ratios account for about one-seventh of overall crime increase in the period 1998–2004. Wei and Zhang (2011b) find that son-families in regions with higher sex ratios are more likely to become entrepreneurs, while daughter-families do not respond to local sex ratios. Using a natural experiment in Chinese Taiwan, Chang and Zhang (2012) and Chang and Zhang (2015) study the effect of mating competition on entrepreneurship in man's marriageable age. Using population census and elderly survey, they quantitatively indicate that young mens entrepreneurship is inspired by such a policy change. In addition, a few recent studies focus on the evidence that son-families have incentives to earn more income. Knight, Li, and Deng (2010) demonstrate that household having a son raises household income by more than 10% using CHIP data. Ding and Zhang (2014) report that son-families receive more remittances and invest more in productive activities compared to household with a daughter using a nationally representative rural household data in 300 villages. Li and Yi (2015) find that premarital son-families are more likely to migrate and generate more earnings than those with only a daughter and gender of the first child matters for

family finance and health consequence based on China Family Panel Survey. One of the interesting research is the impact of sex ratios on happiness. Using both the China Central TV postcard survey, which includes questions on happiness and China Family Panel Survey 2000, Tan, Wei, and Zhang (2015) show that higher sex ratios consistently lower people's self-reported happiness.

Roadmap The rest of the paper is organized as following: in the Section 2, we replicate Wei and Zhang (2011a). Section 3 provides new evidence from the CHFS. We reexamine the competitive saving motive at province level in Section 4. Section 5 concludes. All figures and tables are presented in Appendix.

2 Replication of the Wei and Zhang (2011a)

CHIP 2002 For our replication, we mainly use the household-level data from Chinese Household Income Project (CHIP) of 2002, the same sample as Wei and Zhang (2011a). This CHIP data consists of three distinct samples of urban, rural and migrant household groups. The CHIP sample are selected from a larger sample drawn by the National Bureau of Statistics of the People's Republic of China and conducted by the Chinese Academy of Social Sciences in 2003. The sample cover 22 provinces, 77 urban cities and 122 rural counties.

China population census 1990 and 2000 To reexamine the competitive saving motive, especially reconsider the county sex ratio pattern, we use the two waves of China population census in this section —in an attempt to replicate and extend Wei and Zhang (2011a) analysis. In Wei and Zhang (2011a) household-level regressions, for the cohort of ages 12–21 in year 2002, they infer county local sex ratios from the 1990 population census of cohort 0-9 years old. Taken the data unavailability into consideration, we also include and infer the same aged cohort from the 2000 population census (the cohort was 10-19 years old in 2000). These two local sex ratio inferred are considered as two measures for the key variable of interest, and the 2000 population census is possibly closer to the true value of 2002 county local sex ratio.

It is important to point out that the Wei and Zhang (2011a)'s local sex ratio is inferred from the county report of population census 1990 which includes the detailed population and detail composition of the population for all counties. That local population number is inferred from the population by the NBS. Unfortunately, county report of population

census 1990 is too far for current day, since cannot be assessed to us, we use the 1990 Population Census Sample Survey Data (0.095%) to calculate the local fraction of males or females for an specific range of ages in the county and then premarital local sex ratio. There will be a little difference for these two sources of sex ratio. Table 1 report the summary statistics with the average saving rate for each household type and the comparison with the Wei and Zhang (2011a). Associated with the table 4 of Wei and Zhang (2011a), we follow the Chamon and Prasad (2010) and Wei and Zhang (2011a) to define the household saving rate as log(disposable income/living expenditure). The disposable income is the household total income net of tax. The definition has two advantages according to Wei and Zhang (2011a): one is to make the error term more likely to satisfy the normality assumption and another is less susceptible to extreme values.

Empirical specification Following Wei and Zhang (2011a), we estimate the following empirical model

$$\texttt{saving rate}_{ij} = \beta_1 \cdot \texttt{local sex ratio}_j + \beta_2 \cdot \boldsymbol{X_i} + \beta_3 \cdot \boldsymbol{Y_j} + u_{ij}, \tag{1}$$

where i indicates households and j indicates counties. In this model, `saving rate` is defined as the *log* of the ratio of household disposal income to household consumption; `local sex ratio` is from the premarital cohort age form 0 to 9 in 1990 census data; $\boldsymbol{X_i}$ is a vector of household characteristics, including other determinants of the saving rate as household per-capita income, child's ages and characteristics of the household head: gender, age, age^2, year of schooling, a dummy variable denoting "poor health" if the family has severely ill member or a disabled as the health shocks to the household; $\boldsymbol{Y_j}$ refers to a vector of county characteristics. For the purpose of replication, we first control the county Gini coefficient as Wei and Zhang (2011a) did. Then we extend control variables for the counties to include county per-capita GDP, urbanization ratio, in-migration ratio, male and female unmarried ratio of age 15 and above. We also control the dummy variable defining a household without public insurance, a household with State-owned Enterprise employment, and household with a member in the reorganization firm or in a profit-loss firm or laid-off from the enterprise.

[Insert Table 1 here]

After micro-level data cleaning, we may not have the same sample used for analysis although in this replication case. Fortunately, use the similar nuclear household criteria

as Wei and Zhang (2011a)[6], we get our sample very close to (Wei and Zhang, 2011a) and the summary statistics of saving rate are nearly the same. In our sample of rural area, households with a son has average and median saving rate of 39.3% and 39.4%, respectively, higher than 31.8% and 35.3% for households with a daughter. For the urban area, household with a son and with a daughter has the similar average and median saving rate around 31% for our sample. Table 1 also report the summary comparison for the sex ratio, in our sample inferred from the 1990 Population Census Sample Survey Data: the average sex ratio for rural counties and urban cities is 1.09 and 1.08, with a standard deviation of 0.04 and 0.04, respectively. The smallest and largest values for rural counties and urban cities are 1.01 and 1.23, along with 1.02 and 1.24, respectively. For the sex ratio sample of (Wei and Zhang, 2011a), the sex ratio ranges from 1.01 to 1.23 with a mean of 1.09 for rural counties and 1.02 to 1.24 with a mean of 1.08 for urban cities which is similar to ours. However, the standard deviation are both 0.04 for rural and urban, is much smaller than ours due to the post-adjustment for the local population by NBS.

The key hypothesis of Wei and Zhang (2011a) focus on a particular regional variation in saving rate: holding constant household income and other family or household head level characteristics, household with a son should save more in a region with higher local sex ratio. And the daughter may have free rider effect that not responses to the higher local sex ratio in rural counties and have to save more in a region with higher local sex ratio due to the housing prices spillover channel in urban cities.

Empirical results comparison The table 2 present the regression results and comparison with the Wei and Zhang (2011a). Among them, the table 2a present the rural sample regression results and comparison with the Wei and Zhang (2011a). Column 1 and 2 of table 2a perform respectively the results comparison of the regression relating savings by household with a son and household with a daughter on a full sample. We find the similar results that local sex ratio has a strongly positive effect on the household saving rate, that raising 1 percent of local sex ratio can increase the average rural son-household saving rate by 0.54% which is statistical significant at the 5 percent level. The magnitude of competitive effect is smaller compared to the estimates of Wei and Zhang (2011a), however, the significance level is identical. The competitive saving motive effect is also economically large for the son-household and account large in the actual increase in the average rural

[6]Nuclear families: three-person household with both parents still alive, and mother's age less than 40, and no parents and other relatives living at home. The household survey cannot capture the moveing out children accurately, therefore the household is more likely to be a nuclear family by placing a limit on the mother's age. The sample size of this restricted sample is significant smaller compared to the orginal sample size of CHIP survey(Wei and Zhang, 2011a), the regression owns good statistical power.

household saving rate. Column 2 of table 2a report that the daughter-household does not responds to local sex ratio. The estimates of the effect is same as Wei and Zhang (2011a), and the coefficient on the local sex ratio is negative and not statistically significant. This result is also consistent with the theoretical implication in Du and Wei (2013). Same as Wei and Zhang (2011a), we remove the possible outliers through three different filters to do a sequence of additional regressions and perform and preserve the competitive saving effect in a smaller magnitude and same strong significance pattern. The same outliers removing method are also performing in the later pert, and in specific, the household whose reported annual household income or consumption are took out in the columns 3 and 4; the top and bottom 5 percent of households in terms of their saving rate are took out in the columns of 5 and 6; the top and bottom 5 percent of households in terms of their saving rate along with explicit marital status for the child are took out in column of 7 and 8. In all cases, we present the same patterns of results: the coefficient of the local sex ratio is around 0.5 to 0.6 and significant for son-families but the coefficient is negative and not statistically different from zero for daughter-families.

[Insert Table 2 here]

We now turn to urban household savings, and the table 2b present and compare them with the Wei and Zhang (2011a). In column 1 of table 2b, we contrast the son-families result and the coefficient on local sex ratio is positive and significant as the Wei and Zhang (2011a). In column 2, the point estimate for household with a daughter is 0.24 and significant at 10-percent level. We also get the magnitude of competitive effect is smaller compared to the estimates of Wei and Zhang (2011a), however, the significance level is similar. The effect on household with a son is larger than on daughter-families for full sample. We also attempt to remove the possible outliers as Wei and Zhang (2011a) since the big concern for noise in the data. Through a number of same filters as rural analysis, the effect of sex ratio on son-families is consistently larger than on daughter-families and this pattern is more robust than Wei and Zhang (2011a). The spillover effect of local housing cost on daughter-families still shows, but pressure and response is less significant than Wei and Zhang (2011a)'s estimates. In fact, in column 8, the coefficient for household with a daughter is no longer statistically different from zero.

Do other county characteristics affect the competitive motive? The 2000 population census data not only report the county-level basic situation and the various detailed composition of the population, and also local employment information, local

urbanized population share, local migration share from other provinces or from other cities, and also unmarried population share of aged 15 and above for male and female in each counties. Since the CHIP wave year 2002 is very close to census survey year 2000, we can use these other characteristics at 2000 as factors in our regression to check the robustness of competitive saving motive. In addition, we also get the local GDP and local from this data source: China Statistical YearBook for Regional Economy and to check whether the economy development or economy growth can vanish the competitive saving motive.

In any case, after holding constant family income and other household characteristics, the household saving rate should responds negatively significant to local per-capita GDP. The reason is that the consumer price and household consumption structure make the living expenditure cost is higher in economy developed county, therefore leads to a lower household saving rate. The urbanization ratio and the migration population share from other province or cities are both an indicator for local economy development, we also expect the negative effect on household saving rate after control characteristics like household income. The share of male and female unmarried population which describe the situation and condition for the local marriage market can also be a factor for saving behavior of household with a son.

[Insert Table 3 here]

We do the same robustness check as Wei and Zhang (2011a) including that median regression, using the sample with different mother age threshold, the sample with extend household with other relatives and also using interaction term in the son and daughter whole sample regression. The results are also presented in consistent patterns: local sex ratio has positive significant effect for the son-families but not for the counterpart. For the purpose of saving space and focusing on our point, we don't report these tables. In table 3, we report several additional robustness checks rather than Wei and Zhang (2011a) using the other county-level factors mentioned above. The econometric purpose is to check whether other county characteristics are omitted variables in model 1. We first present the results with the logarithm of per-capita GDP, and then table 3b—3e report the sequence of results. In all cases, the results are still performed in the same pattern and robust. The effect of 1 percent local sex ratio increase on son-families saving rate is around 0.48% to 0.65% for the rural and around 0.16% to 0.38% for the urban. Therefore, in this sub-section, the competitive saving motive still holds for son-families.

The coefficient of sex ratio is around 0.2 and significant at the 10 percent level for urban daughter-families, and negative insignificant for rural counterpart. The coefficient of the related county factors meets the expectation except for the variable "share of male and female unmarried population" with an insignificant effect, and the potential reason is the unmarried population share is not the premarital age specific.

What if use year 2000 sex ratio rather than 1990? Unlike the county report of population census 1990 is beyond our available data, the 2000 census county report which includes post-adjustment for the local population, the detailed composition of the population for all counties by NBS is among our data availability. Another reason is that local population number of 2000 is possibly closer to the true value of 2002 county local sex ratio. Therefore, we can use sex ratio inferred from year 2000 to do the same regression. To ensure the precision and maximize the comparability across the two waves of population census, we match administrative division code for all counties of year 1990 to year 2000 due to the change of local administrative jurisdiction. The process is not that easy. We use the communique of the State Council of the People's Republic of China from year 1990 to year 2000, to get the history of the change of local administrative jurisdiction code, and match them together after that.

In this subsection, we use the same rural and urban CHIP 2002 sample. The only difference is the data source of sex ratio. From 1990 census data, we use the age cohort of 0-9 years old; from 2000 census data, we use the age cohort of 10-19 years old. Therefore, the difference between population census 1990 and 2000 is mainly because the migration for the premarital aged population between 10-19 years old in each county. Resident with more education could be more mobile such as college students with age 17 and above may stay in the city of their college after graduation. Therefore, the results may be different.

[Insert Table 4 here]

Table 4 reports the regression results with same specification as Wei and Zhang (2011a) using sex ratio of year 2000. The top row is the result of rural sample and the bottom row is for urban sample. And surprisingly, the competitive saving motive effect vanish with sex ratio of year 2000. The coefficient of local sex ratio on household saving rate of rural son-families is around 0.18 to 0.42 and insignificant at 10 percent level, and also no longer statistically from zero using the full sample or the sample after removing the possible outliers. The effect on urban son-families is also negligible and insignificant. The point estimates of sex ratio on rural and urban household with a daughter is around -0.20 to

0.38 and -0.59 to -0.00 and almost insignificant, respectively. The effects vanish for CHIP 2002 both urban and rual son-families and daughter-families with 2000 China population census.

Potential explanation: migration As we mentioned above, the only difference is the data source of sex ratio, and the difference between population census 1990 and 2000 is mainly because the migration for the premarital aged population between 10-19 years old in each county. Therefore, unbalanced migrate-in and migrate-out for premarital male and female is our major potential explanation.

[Insert Table 5 here]

Table 5 report the distribution for county sex ratio of rural counties and urban cities. Since census 2000 and census 1990 of Wei and Zhang (2011a) come from the same data source in NBS county report of population census, we mainly compare these two distribution. The standard deviation and mean value of two groups is very close. However, since the potential premarital migration occurs within 10-years period from 1990 to 2000, the maximum value of rural counties and urban cities of year 2000 is 1.27 and 1.23, slightly bigger than the corresponding maximum value of year 1990. Moreover, the minimum value of census 2000 in rural counties is 0.77, which is significant smaller than 1.01 in 1990. The pattern is also applied to urban counties with 0.92 compared to 1.02. This distribution change confirm the explanation of migration potentially.

[Insert Figure 3 here]

Next, we turn to the local sex ratio comparison between year 2000 and year 1990 using from Population Census Sample Survey Data in figure 3. The top two sub-figures in blue show the distribution of 1990 Sample Survey and the bottom two show the distribution of county report 2000. Census Sample Survey should have larger standard deviation due to no adjustment in local population in general, however, the minimum value of rural counties in 2000 is smaller rather than 1990 Sample Survey. The premarital female migrating-out from rural counties can explain this results. Figure 4 presents cross-section scatter plot using per-capita income calculated from rural and urban county sample in CHIP 2002 with sex ratio in 1990 and 2000, respectively. For the rural part, there is clear positive and significant relationship between per-capita income and sex ratio in 1990, however, the relationship turn to negative and significant, means that the premarital aged female tend

to migrate out from poor rural counties. The plot for urban cities is in a more ambiguous way but the negative relationship is rather stronger in year 2000. Same pattern are confirmed by the figure 5. The higher per-capita GDP or per-capita income, the more sex ratio reducing from 2000 to 1990.

[Insert Figure 4 here]

[Insert Figure 5 here]

To go a further step, we focus on checking competitive motive of son-families and group the son-nuclear-household sample into three parts by different criterion. The first rule is based on the per-capita income of each county. Since the CHIP data separated into rural and urban sample, we divide the rural counties and urban cities into three parts respectively.[7] We can also group the son-families by household-income—the second rule.[8] These two pattern are presented in figure 6 and figure 7, respectively. Along with the unconditional effect presented, table 6 presents the conditional causal effect on local sex ratio including the daughter-families. The top half of figure 6 and above panel of the table 6 present the similar correlation that for rich rural, the coefficient of county sex ratio in 1990 and 2000 on household saving rate is 1.1 and 0.9 respectively and both significant at 10 percent level. The results are robust both using the full sample or sample without outliers. However, for the poor rural counties which per-capita income is among bottom 25 percent, the competitive effect vanishes again using county sex ratio in 2000, but not for sex ratio in 1990. The bottom panel of figure 6 present ambiguous correlation in a similar way. Both above panel of table 7 in rural and urban regression also reports the similar interpretation: using 1990 county sex ratio, the effects of full rural sample and rich rural counties are positive significant and the effects are marginal significant for middle and poor rural counties. The middle and poor urban cities do present the significant competitive saving motive effect and very rich urban cities do not. However, for local sex ratio in 2000, only rich rural and rich urban can present the positive significant effect, but other sub-sample along with the full sample give the disappearance of competitive saving motive. In specific, all daughter-families sub-sample give insignificant results except that in the poor urban cities due to the housing price spillover effect. The sub-sample analysis

[7] We divide rural counties as poor rural, middle rural and rich rural counties; similarly, the urban cities can be separated into three parts that poor urban, middle urban and rich urban.

[8] We divide rural households as poor rural, middle rural and rich rural households; similarly, the urban households can be separated into three parts that poor urban, middle urban and rich urban.

and comparison using the second rule are presented in figure 7 and table 7. Only sample according to middle-class household income group shows the positive effect and significant at 10-percent level with sex ratio in 2000, and the poor household and household in poor counties consistently do not respond to local sex ratio. One major explanation is following the logic from the point focusing on sex ratio comparison and relation between sex ratio change and local economy development mentioned above. If the young female migrates out from the poor rural counties which leads to more unbalanced sex ratio in poor county, the son-families in this county will face more unprecedented fierce competition in marriage market. In addition, no matter where is the very poor son-families in, the household has no incentive to save. One thing, they cannot compete very well in the rich or middle cities or counties; another, the situation becomes worse in poor rural from 1990 to 2000. In a extreme case, there will be no available premarital aged female in the county, then the son-families could choose to migrate out or quit the market directly. Therefore, the migration of aged 10-19 household during the 10–years–long period could strongly affect the saving behavior and marriage motive for both son- and daughter- families. Reconsidering the sex ratio in year 2000 is crucial for explaining the competitive saving motive and also understand the migration profile and income-competition profile in rural and urban China.

[Insert Figure 6 here]

[Insert Figure 7 here]

Table 8 presents one of the potential reason about the premarital female migration out from the poor rural counties. Especially in rural China, marriage migration is the major form of migration and cause of female migration. In China, daughter has less responsibility to look after elder parent compared to son, young female have a higher probability to migrate out for working or schooling and stay in the new place for marriage. The numbers are all computed from census 1990 and Population Census Sample Survey. It implies that young women could be much more mobile than young man because of the marriage migration in both intra-provincial or Inter-provincial way, especially for that from poor rural one. We can conclude our point that unbalanced migration of premarital male and female is the major explanation causing the competitive motive effects vanish in the full CHIP sample. The competitive saving motive could hold for the household with sufficient disposable income in the economy developed counties, but not for the poor counterpart.

[Insert Table 8 here]

3 Competitive saving motive: evidence from CHFS

In this section, we turn to an alternative data set — the China Finance Household Survey (CHFS) — to examine the impact of sex ratio on household saving rate. The CHFS, conducted by Southwestern University of Finance and Economics in China, is a nationally representative longitudinal household survey data, contains sufficient information about individual and household income, expenditures, assets, and debts. It fits well with our research purpose. The survey started in 2011, and since then there are another three waves: 2013, 2015, and 2017. We mainly use the 2015 wave in this analysis. The CHFS 2015 includes 353 counties from 29 provinces (excluding Tibet Autonomous Region and Xinjiang Uighur Autonomous Region) in China, and is a sample of approximately 38,000 household and 140,000 individuals.

The CHFS survey have two potential advantages. First, the survey contains much more detailed information on household finance, and also covers a much larger sample size. These advantages have important implications for the reliability of the analysis. The rich information on household finance is important for getting relative accurate income and consumption statistics; and having a large number of households means that we can control for province fixed effect so as to tackle with the geographical difference in economic conditions and culture, given that there emerges growing divergence across regions in China. Another advantage is the high data quality especially. The sampling design in the CHFS focuses on a large range of households whereas the CHIP sample consists of majority households from the the state-owned enterprise (SOE) sector. In addition, to deal with the issue of only reporting income range in the sample (including for the high-income households), the CHFS infer the actual income from taxes reports. This approach would eliminate the effect of the top coded, the common issue in the survey data, on the analysis.

[Insert Table 9 here]

Using the CHFS 2015 and the *local sex ratio* for the cohort of ages 10-24 in 2015 inferred from the cohort for age 5-19 years in the 2010 population census, we follow the same method and control for the same households and counties characteristics as used in Wei and Zhang (2011a) to reconsidering the evidence of the competitive saving motive.

Table 9 reports the regression results from the CHFS 2015 sample. Column 1 and 2 of the table perform the comparison results for households with a son and with a daughter on a full sample without removing possible outliers. respectively. Column 3 and 4 show the comparison results on a sub-sample with removing bottom and top 5 percent saving rates.

The above panel of the table presents the results for the rural sample. For households with a son in the rural full sample, the estimated effect of the local sex ratio of cohort for age 5–14 on household saving rate is 0.66 and statistical significant at 5 percent level, which means that 1 percent increase in local sex ratio would lead a 0.66 percent increase on average in household saving rate. The estimate is larger than our replicated result using the CHIP (0.54 in Table 2a), but is still much smaller than 1.34 in Wei and Zhang (2011a). For households with a daughter in the rural full sample, we get a larger effect than households with a son, with raising 1 percent sex ratio would increase 0.75 percent household saving rate. By contrast, both our and Wei and Zhang (2011a)'s results in Table 2a suggests that there is no significant effect for households with a daughter in the CHIP rural sample. Using the sex ratio of cohort for age 10–19 as an alternative measure, the effect of sex ratio on household saving rate is 0.54, although it is not statistically significant, for household with a son in the rural sample, while the effect is 0.78 for households with a daughter and it is statistical significant at 5 percent level. After removing possible saving rate outliers, the estimates in column 3 show that the effects of sex ratio for cohort of age 5–14 on household saving rate for household with a son is 0.68, and it is 0.61 when use the sex ratio for cohort of age 10–19. Both the estimated coefficients are a little bit larger than in column 1 and are statistically significant. Column 4 shows that for households with a daughter, the competitive saving motive preserve in a little bit smaller magnitude at 10 percent level.

The results for the urban sample are presented in the below part of the table. It is surprising that all estimated effects of the sex ratio on household saving rate in column 3 and 4 become much smaller than the rural sample, and the most important is that they are no longer statistically different from zero. These results are totally different with them in Table 2b.

4 Competitive saving motive: across-regions

Sex ratios and saving rates across province In order to identify the effects of the sex ratio on household saving rate at province level, we estimate the following regression equation

$$\texttt{saving rate}_{mt} = \beta_0 + \beta_1 \cdot \texttt{province sex ratio}_{mt} + \beta_2 \cdot \boldsymbol{Y}_{mt} + \lambda_t + \eta_m + u_{mt}, \quad (2)$$

where m indicates provinces and t indicates years. In this model, `saving rate` is defined as the value of the aggregate household disposal income net aggregate household consumption dived by aggregate household disposal income; The income and consumption are computed from the Comprehensive Statistical Data and Materials on 50/55/60 years of New China (CNBS); $\boldsymbol{Y_m}$ refers to a vector of province characteristics, including per capita GDP (calculated from the CNBS), the young and old dependent ratio (calculated from the China Statistical YearBook 1986–2016), and the share of SOE in total labor force (calculated from the China Statistical YearBook 1986–2016); The year and province fixed effects are captured by λ_t and η_m, respectively; u_{mt} is the error term. $\texttt{province sex ratio}_{mt}$ is the variable of interest, which is calculated from the premarital cohort for age 5 to 19 (or for age 10 to 19 as a robustness check) in census data 1990, 2000 and 2010. The estimate of β_1 identifies the causal effect of the province sex ratio on the province aggregate household saving rate. The model 2 is estimated by the pooled OLS estimation, and the standard errors are clustered at the province level.

[Insert Table 10 here]

In Table 10, we report the main regression results. First three columns show the results with the sex ratio for age cohort 5–10 as the key regressors. In Column 1, the results presented are from full sample. The effect of the *log* of province per capita income on the province aggregate household saving rate is 0.20 and statistical significant, which means that a 1 percent increase in the per capita income would lead to a 0.20 percent increase in the saving rate. This result is consistent with Wei and Zhang (2011a). In Table 14 of that paper, they estimate a similar coefficient. However, the estimated coefficient of the sex ratio on the province aggregate household saving rate is -0.03 and statistically insignificant different from zero. In other words, there is no significant effect of the sex ratio on the province saving rate. This result contradicts with the finding in Wei and

Zhang (2011a) where they argue that the effect of the local sex ratio on the local saving rate is significant positive with a coefficient of 0.28.

In column 2, we extend our analysis to examine whether the competitive motive exist in the urban sample. The estimated association between the local sex ratio and the local saving rate is still insignificant different from 0, although the magnitude now become positive at 0.05. This result reflect that in the urban sample, the competitive may not exist. As the full sample regression, the per capita income has a significant positive impact on the local saving rate with a 1 percent increase in per capita income leading to a 0.14 percent increase in the local saving rate.

Column 3 reports the results when we restrict to the rural sample. The estimated effects of the province per capita income on the province saving rate is significant positive. A 1 percent increase in the per capita income would lead to a higher saving rate by 0.39 percent. For the competitive saving motive, the estimated coefficient of the sex ratio is 0.24 and statistically significant different from zero at the 5 percent level. This result indicates that the province saving rate is higher in the region with a higher sex ratio, which implies that the competitive saving rate exists in the rural.

Column 3 to 6 of the table presents the results using the sex ratio for age cohort 10–19 as the key explanatory variable. We find pretty close estimates of the coefficients for both the per capita income and the sex ratio. There is a significant positive effect of the per capita income on the saving rate, with the coefficient 0.20 for the full sample, 0.15 for the urban, and 0.39 for the rural. Although the estimated coefficients for the full and the rural sample are almost the same as use the sex ratio for age 5–19, they are insignificant different from zero. For the rural, the estimated marginal effect of the sex ratio is 0.21 and significant at 5 percent level. Overall, the results change a little bit and support that the competitive saving motive only exist in the rural.

Discussion about the results In this section, we use data from Population Census (1990, 2000, 2010), China Statistical YearBook 1986–2016, and Comprehensive Statistical Data and Materials on 50/55/60 years of New China (CNBS) — the same data sources as Wei and Zhang (2011a) used, and follow the same specification, but why are our estimated results different with that in Wei and Zhang (2011a)?

The main reason is probably that we use a different approach to define the sex ratio for premarital cohort and consider a shorter time span. In Wei and Zhang (2011a), they use data from 31 provinces for 27 years (1990 – 2007) in the panel data regression. Except for the sex ratio, the data contain all other variables information for each province in each year. To get the statistics for the sex ratio for each province in each year, they use the

2000 population census as basis to infer sex ratio in other years. Specifically, they infer the sex ratio of the cohort for age 7–21 (their main focus) in 1990 from the cohort for age 17–31 in the 2000 census, and the sex ratio of the cohort for age 7–21 in 1991 from the cohort for age 18–32 in the 2000 census, and so on; similarly, the sex ratio of the cohort for age 7–21 in 2001 is inferred by the cohort for age 6–20 in the 2000 census, and the sex ratio of the cohort for age 7–21 in 2002 is inferred by the cohort for age 5–19, and so on. For the 2000 sex ratio, it is drawn directly from the 2000 census.

The major advantage of the approach in Wei and Zhang (2011a) is to enlarge the sample size. As they motioned in the paper: "Ideally, we would like to know sex ratio for a fixed age cohort in every region and in every year. ... we make the following shortcut:..." Indeed, they have over 800 observations to do the panel regression. However, their approach also has a significant drawback: the algorithm assumes that there is no change in the sex ratio between 7–21 age group in year i and [7+(2000-i)]–[21+(2000-i)] age group in 2000. It is not convincing that the cohort for age 7–21 in 1990 are the same as the cohort for age 17–31 in 2000, similarly for the cohort for age 7–21 in 2007 being the same as the cohort for age 0–14 in 2000. In fact, as discussed in the previous sections, in China's big urbanization progress, the internal migration is one of the world's most extensive migration: from the rural to urban, or from the inland to the the coast, or for the poor area to the rich area, etc. The effect of the huge internal movement would not be ignored. Migration do affect local sex ratio. Therefore, we adopt a relatively accurate way to compute sex ratio so as to minimize the effect due to measurement error on sex ratio. Specifically, we consider the sample of 31 provinces only for 7 years (1985, 1990, 1995, 2000, 2005, 2010, and 2015, except for Chongqing, which has only 3 years data). The sex ratio of each province for the year 1990, 2000, and 2010 is directly drawn the corresponding census data; For the sex ratio in 1985, it is calculated from the census 1990. 1995 and 2005's sex ratio are computed by the average of before- and after- current year's sex ratio; For the sex ration in 2015, it is inferred from the census 2010. Our algorithm alleviate the concern and thus show that the competitive saving motive exists only in the rural area. By estimation and computation, an increase in sex ratio from 1985 to 2015 can explain about 28% of the actual increase of the increase of rural saving rate.

5 Conclusion

The high aggregate household saving rate is one of the unique feature of the Chinese economy. Over the period of China rapid income growth, China's sex ratio, which is

defined as the ratio of the number of boys at birth to the number of girls at birth, has been also experienced a rapid a growth. The common trend has captured economists' attention on linking the saving rate to the sex ratio. Can the sex ratio explain the high China's household saving rate? Starting with Wei and Zhang (2011a), they argue a explanation called "competitive saving motive" that increasingly unbalanced premarital sex ratio require the both rural and urban son-families to save increasingly more to compete in the marriage market and show that the rapid increase in premarital sex ratio during 1990-2007 can potentially explain about half of the sharp increase in China's household saving rate.

In this paper, we reexamine the impact of the sex ratio on the household saving rate. First, we use the CHIP data — the same sources of data as used by Wei and Zhang (2011a) — in an attempt to replicate and extend the estimates reported in Wei and Zhang (2011a). We use local sex ratio inferred from 2000 China population census and find the competitive saving motive only holds for the household in rich counties. The effects vanish for CHIP 2002 full sample with census data in 2000, especially for poor households and households in poor counties. We argue that unbalanced migration of premarital male and female is the major explanation for the vanishing effect.

Next, we turn to another nationally representative data set China Household Finance Survey (CHFS) 2015 wave for estimating the effect of the sex ratio on household saving rate. The result indicates that the competitive saving motive still hold, but only in the rural area. Specifically, in the rural, the effects on households with a son is much smaller, and the effects on households with a daughter are also significant positive. However, there is no effects on urban households.

Finally, we use China population census and provincial panel data as the same data sources as Wei and Zhang (2011a) used, and adopt a more precise algorithm to compute the sex ratios of the 31 provinces for 7 years (1985, 1990, 1995, 2000, 2005, 2010, and 2015). Our results report that the competitive saving motive exist, but only in the rural area. By estimation and computation, an increase in sex ratio from 1985 to 2015 can explain about 28% of the actual increase of the increase of rural saving rate.

References

Carroll, C. D., and D. N. Weil (1994): "Saving and growth: a reinterpretation," in *Carnegie-Rochester Conference Series on Public Policy*, vol. 40, pp. 133–192. Elsevier.

Chamon, M. D., and E. S. Prasad (2010): "Why are saving rates of urban households in China rising?," *American Economic Journal: Macroeconomics*, pp. 93–130.

Chang, S., and X. Zhang (2012): "The economic consequences of excess men," Discussion paper, IFPRI Discussion Paper.

——— (2015): "Mating competition and entrepreneurship," *Journal of Economic Behavior & Organization*, 116, 292–309.

Curtis, C. C., S. Lugauer, and N. C. Mark (2015): "Demographic patterns and household saving in China," *American Economic Journal: Macroeconomics*, 7(2), 58–94.

Ding, W., and Y. Zhang (2014): "When a son is born: The impact of fertility patterns on family finance in rural China," *China Economic Review*, 30, 192–208.

Du, Q., and S.-J. Wei (2013): "A theory of the competitive saving motive," *Journal of International Economics*, 91(2), 275–289.

——— (2016): "A Darwinian perspective on exchange rate undervaluation," *European Economic Review*, 83, 111–138.

Edlund, L., H. Li, J. Yi, and J. Zhang (2013): "Sex ratios and crime: Evidence from China," *Review of Economics and Statistics*, 95(5), 1520–1534.

He, H., F. Huang, Z. Liu, and D. Zhu (2017): "Breaking the iron rice bowl: Evidence of precautionary savings from the chinese state-owned enterprises reform," *Journal of Monetary Economics*.

Horioka, C. Y., and A. Terada-Hagiwara (2017): "The impact of sex ratios before marriage on household saving in two Asian countries: The competitive saving motive revisited," *Review of Economics of the Household*, 15(3), 739–757.

Horioka, C. Y., and J. Wan (2007): "The determinants of household saving in China: a dynamic panel analysis of provincial data," *Journal of Money, Credit and Banking*, 39(8), 2077–2096.

Knight, J., S. Li, and Q. Deng (2010): "Son preference and household income in rural China," *The journal of development studies*, 46(10), 1786–1805.

Li, W., and J. Yi (2015): "The Competitive Earning Incentive for Sons: Evidence from Migration in China," Discussion paper, IZA Discussion Paper.

Meng, X. (2003): "Unemployment, consumption smoothing, and precautionary saving in urban China," *Journal of Comparative Economics*, 31(3), 465–485.

Modigliani, F., and S. L. Cao (2004): "The Chinese saving puzzle and the life-cycle hypothesis," *Journal of economic literature*, 42(1), 145–170.

Tan, Z., S.-J. Wei, and X. Zhang (2015): "Deadly Discrimination: Implications of "Missing Girls" for Workplace Safety," Discussion paper, IFPRI Conference Paper.

Wei, S.-J., and X. Zhang (2011a): "The competitive saving motive: Evidence from rising sex ratios and savings rates in China," *Journal of political Economy*, 119(3), 511–564.

——— (2011b): "Sex ratios, entrepreneurship, and economic growth in the People's Republic of China," Discussion paper, National Bureau of Economic Research.

——— (2016): "The competitive saving motive: Concept, evidence, and implications," *Frontiers of Economics in China*, 11(3), 355–366.

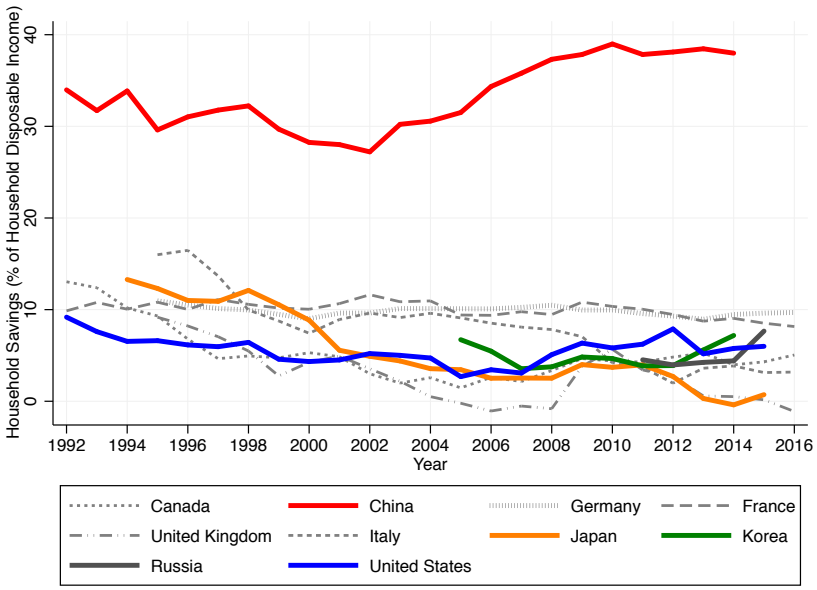
Wei, S.-J., X. Zhang, and Y. Liu (2017): "Home ownership as status competition: Some theory and evidence," *Journal of Development Economics*, 127, 169–186.

Yang, D. T., J. Zhang, and S. Zhou (2012): "Why are saving rates so high in China?," in *Capitalizing China*, pp. 249–278. University of Chicago Press.

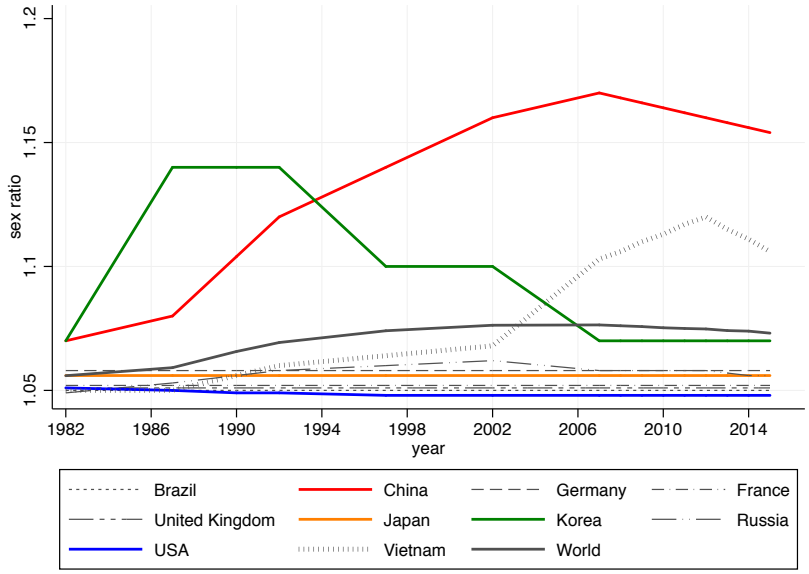
Appendix

A Figures

Figure 1: Aggregate Saving Rate and Sex Ratio At Birth Across Countries



(1) Saving rate comparison across countries



(2) Sex ratio comparison across countries

Figure 2: China's Aged 10-19 Sex Ratio and Household Saving Rate

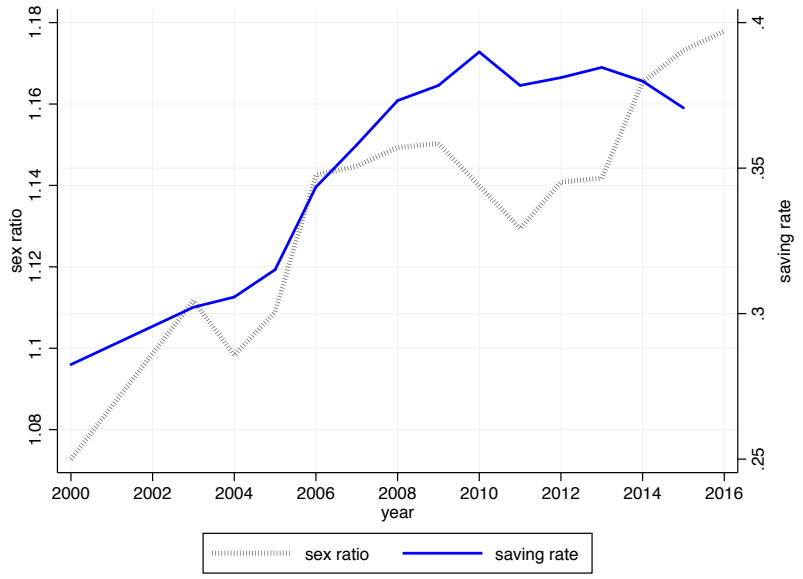


Figure 3: Change in the distribution of sex ratio between 1990 and 2000: distribution comparison

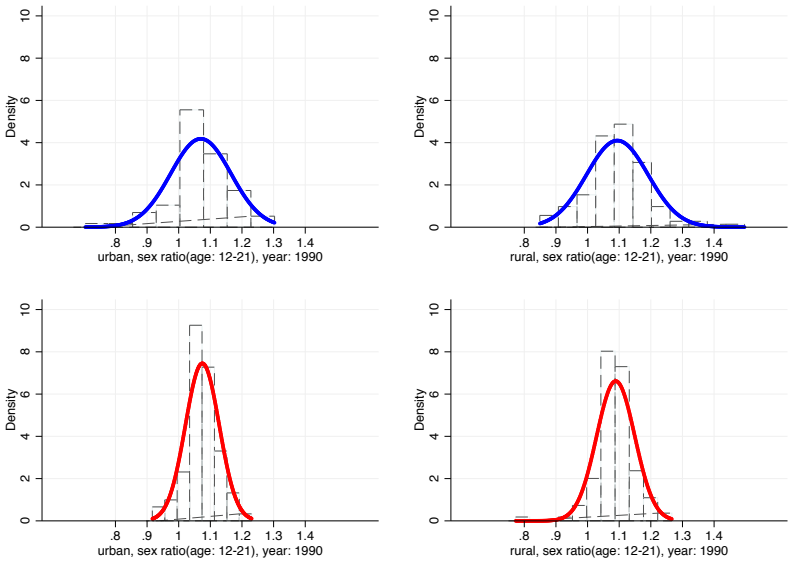
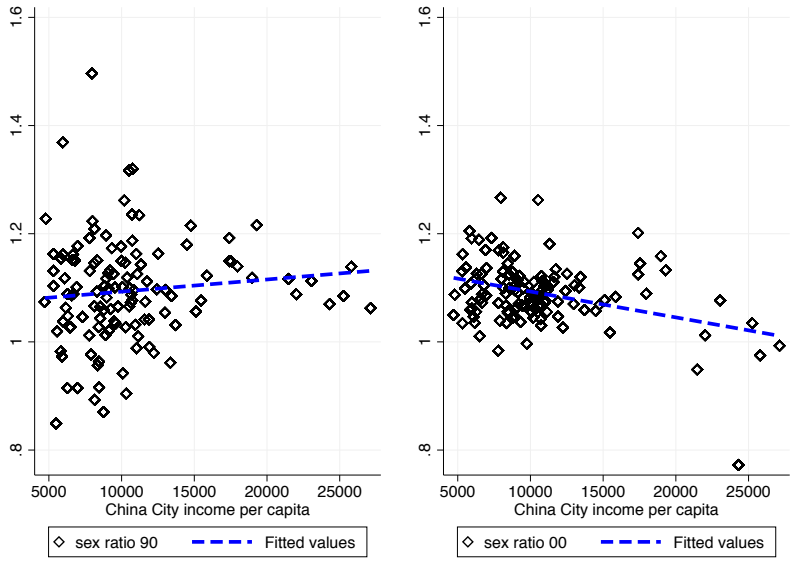
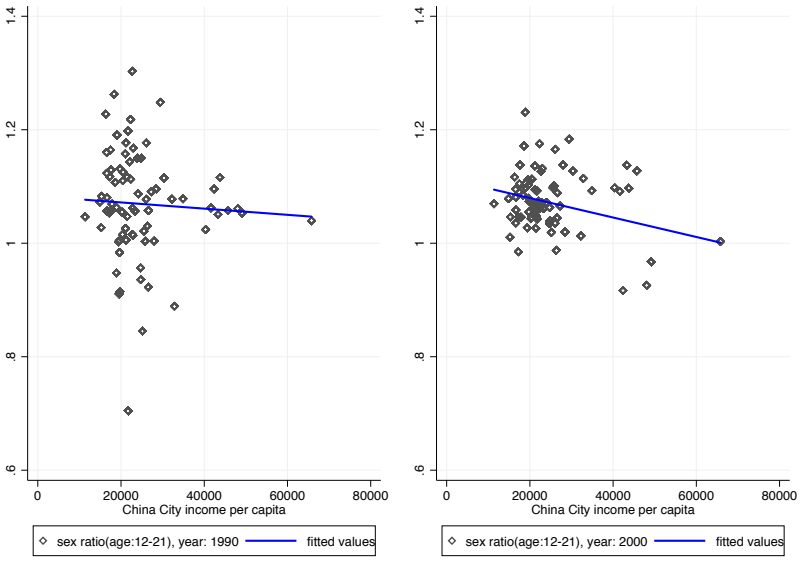


Figure 4: Sex ratio vs. per capita income 1990 and 2000



(1) Rural



(2) Urban

Figure 5: Change in the rural county sex ration vs. per capita GDP and income

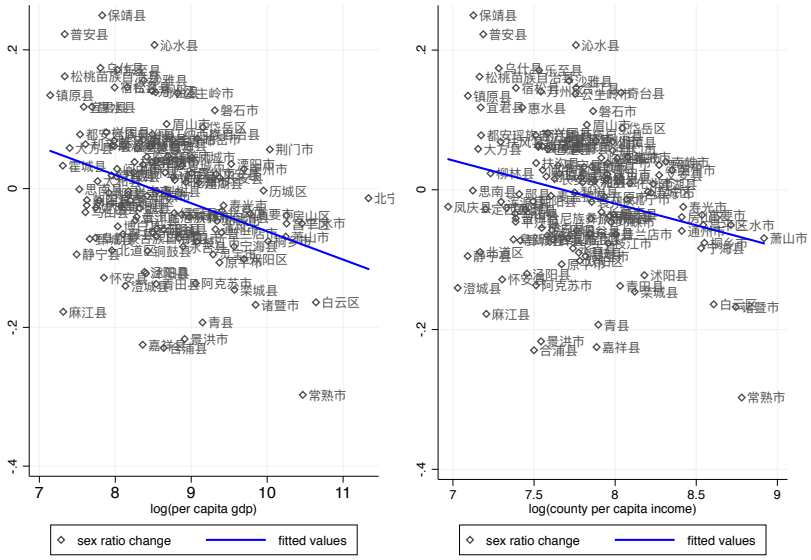


Figure 6: Saving rate and sex ratio by households in rich and poor counties between 1990 and 2000

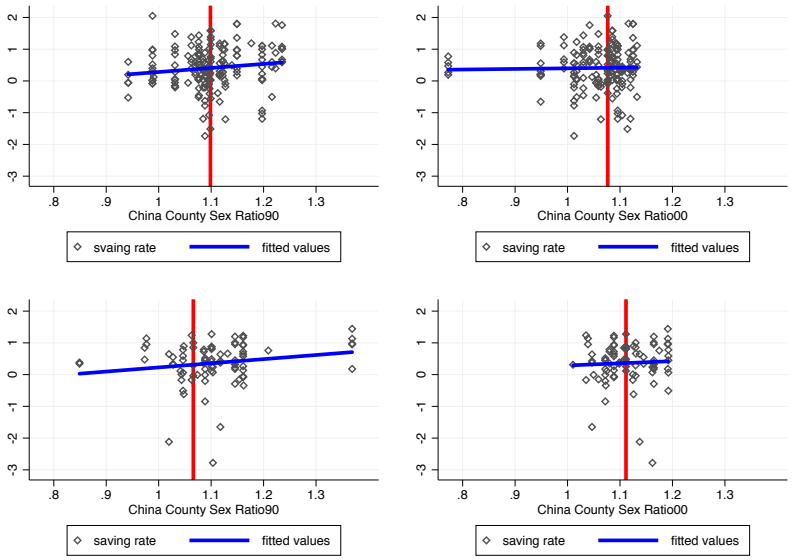
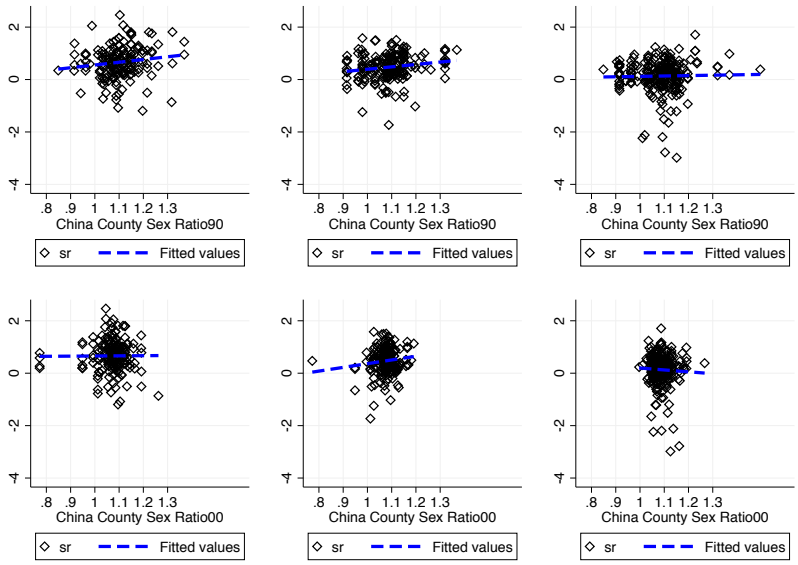
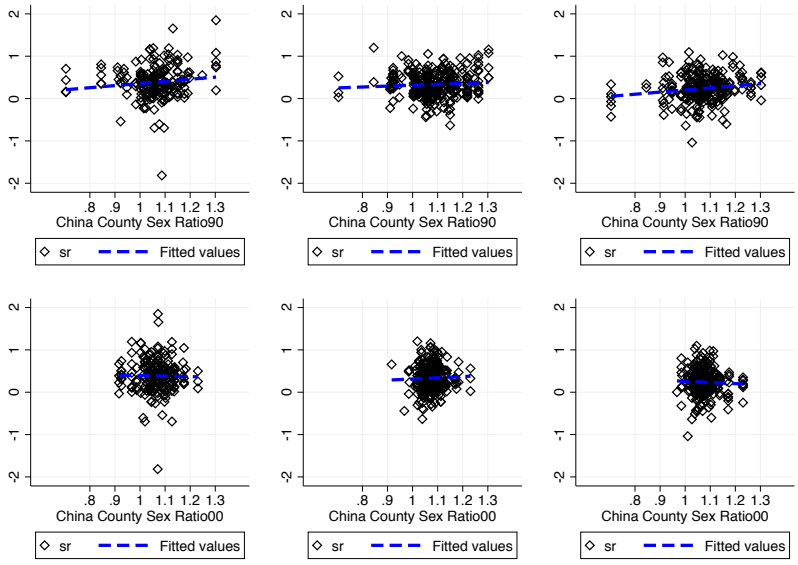


Figure 7: Saving rate and sex ratio by rich, middle-income, and poor households rich between 1990 and 2000



(1) rural



(2) urban

B Tables

Table 1: Summary Statistics: a comparison with Wei and Zhang (2011, JPE)

Household type	Mean	Median	Max	Min	SD	Observations
Rural (Wei and Zhang, 2011)						
One son	0.393	0.394	2.462	-2.986	0.625	580
One daughter	0.318	0.353	1.812	-3.559	0.626	326
sex ratio 1990	1.090		1.230	1.010	0.040	122
Urban (Wei and Zhang, 2011)						
One son	0.312	0.306	1.849	-1.816	0.333	769
One daughter	0.302	0.308	2.153	-1.299	0.356	766
sex ratio 1990	1.080		1.240	1.020	0.040	77
Rural (this paper)						
One son	0.387	0.392	2.462	-2.986	0.624	587
One daughter	0.325	0.365	1.821	-3.559	0.622	329
sex ratio 1990	1.090	1.094	1.490	0.850	0.970	122
Urban (this paper)						
One son	0.314	0.31	1.849	-1.816	0.332	764
One daughter	0.305	0.311	2.153	-1.299	0.357	763
sex ratio 1990	1.070	1.063	1.300	0.700	0.090	77

Table 2: Results Comparison with Wei and Zhang (2011, JPE): local sex ratios for the age cohort 12-21 in 2002 from 1990 census 0-9 years old

(a) Rural household-level evidence

	Subsample that removes the following potential outliers: Rural							
	Full sample		inc. or expend. < 2,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
sex ratio	0.54**	-0.13	0.52**	-0.27	0.62***	-0.24	0.65***	-0.21
(county)	(0.23)	(0.27)	(0.23)	(0.23)	(0.17)	(0.19)	(0.20)	(0.26)
Obersvations	586	329	569	318	528	297	446	242
sex ratio	1.34**	-0.17	1.38**	-0.18	1.10***	-0.23	1.20***	-0.32
(county)	(0.52)	(0.55)	(0.51)	(0.54)	(0.44)	(0.43)	(0.43)	(0.44)
Observations	580	326	564	315	522	292	489	269

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

(b) Urban household-level evidence

	Subsample that removes the following potential outliers: Urban							
	Full sample		inc. or expend. < 3,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
sex ratio	0.38***	0.24*	0.36***	0.22*	0.17**	0.13*	0.21**	0.10
(county)	(0.12)	(0.13)	(0.12)	(0.13)	(0.07)	(0.08)	(0.09)	(0.13)
Observations	764	763	762	759	688	687	377	396
sex ratio	1.54**	1.85**	1.16**	1.07**	0.74***	0.65**	0.98**	0.47
(county)	(0.29)	(0.33)	(0.30)	(0.37)	(0.24)	(0.26)	(0.31)	(0.44)
Obersvations	769	766	604	605	691	688	384	399

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Robustness checks using other county characteristics

(a) Per capita GDP may have no effect on the competitive motive

	Subsample that removes the following potential outliers: Urban							
	Full sample		inc. or expend. < 2,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
sex ratio 1990	0.53**	-0.09	0.51**	-0.24	0.61***	-0.21	0.65***	-0.20
(county)	(0.23)	(0.26)	(0.22)	(0.23)	(0.17)	(0.19)	(0.19)	(0.26)
ln_pcgdp	-0.08**	-0.12***	-0.07**	-0.11***	-0.06**	-0.07**	-0.04	-0.07*
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Observations	586	329	569	318	528	297	446	242
sex ratio 1990	0.36***	0.23*	0.34***	0.21*	0.16*	0.12*	0.26**	0.11
(county)	(0.12)	(0.13)	(0.12)	(0.13)	(0.09)	(0.07)	(0.13)	(0.13)
ln_pcgdp	-0.04***	-0.06***	-0.05***	-0.06***	-0.02	-0.04***	-0.03	-0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Observations	764	763	762	759	688	687	377	396

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

(b) Urbanization may have no effect on the competitive motive

	Subsample that removes the following potential outliers: Urban							
	Full sample		inc. or expend. < 2,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
ratio_1221_90	0.49**	-0.13	0.48**	-0.28	0.58***	-0.23	0.62***	-0.22
	(0.23)	(0.27)	(0.22)	(0.22)	(0.17)	(0.19)	(0.19)	(0.26)
urban_ratio	-0.47**	-0.45**	-0.45**	-0.51**	-0.36**	-0.34**	-0.30**	-0.30
	(0.19)	(0.21)	(0.18)	(0.20)	(0.14)	(0.17)	(0.14)	(0.19)
Observations	586	329	569	318	528	297	446	242
ratio_1221_90	0.38***	0.23*	0.36***	0.21*	0.16*	0.04	0.25**	0.12
	(0.12)	(0.13)	(0.12)	(0.13)	(0.09)	(0.10)	(0.13)	(0.13)
urban_ratio	-0.30***	-0.27***	-0.30***	-0.27***	-0.16***	-0.15***	-0.10*	-0.12*
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.07)
Observations	764	763	762	759	688	687	377	396

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

(c) Provincial migration share may have no effect on the competitive motive

	Subsample that removes the following potential outliers: Urban							
	Full sample		inc. or expend. < 2,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
ratio_1221_90	0.53**	-0.12	0.52**	-0.26	0.62***	-0.21	0.64***	-0.20
	(0.23)	(0.27)	(0.22)	(0.23)	(0.17)	(0.19)	(0.20)	(0.26)
migration_prov	-4.02***	-2.06***	-3.93***	-2.21***	-2.23***	-1.60***	-1.76**	-1.59***
	(0.96)	(0.58)	(0.94)	(0.57)	(0.69)	(0.45)	(0.69)	(0.49)
Observations	586	329	569	318	528	297	446	242
ratio_1221_90	0.37***	0.23*	0.36***	0.21	0.16*	0.04	0.24*	0.10
	(0.12)	(0.13)	(0.12)	(0.13)	(0.09)	(0.10)	(0.13)	(0.13)
migration_prov	-0.64**	-0.58**	-0.70**	-0.62**	-0.46**	-0.25	0.04	-0.19
	(0.29)	(0.29)	(0.28)	(0.29)	(0.20)	(0.21)	(0.22)	(0.24)
Observations	764	763	762	759	688	687	377	396

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

(d) City migration share have no effect on the competitive motive

	Subsample that removes the following potential outliers: Urban							
	Full sample		inc. or expend. < 2,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
ratio_1221_90	0.56**	-0.11	0.54**	-0.25	0.63***	-0.21	0.65***	-0.20
	(0.23)	(0.27)	(0.22)	(0.23)	(0.17)	(0.19)	(0.19)	(0.26)
migration_city	-3.06***	-1.65***	-2.96***	-1.77***	-1.87***	-1.24***	-1.55***	-1.25***
	(0.73)	(0.49)	(0.71)	(0.48)	(0.53)	(0.34)	(0.52)	(0.36)
Observations	586	329	569	318	528	297	446	242
ratio_1221_90	0.37***	0.22*	0.36***	0.21	0.16*	0.04	0.24*	0.10
	(0.12)	(0.13)	(0.12)	(0.13)	(0.09)	(0.10)	(0.13)	(0.13)
migration_city	-0.52***	-0.42**	-0.56***	-0.44**	-0.34**	-0.21	-0.00	-0.11
	(0.18)	(0.20)	(0.17)	(0.20)	(0.13)	(0.14)	(0.15)	(0.17)
Observations	764	763	762	759	688	687	377	396

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

(e) Unmarried population share may have no effect on the competitive motive

	Subsample that removes the following potential outliers: Urban							
	Full sample		inc. or expend. < 2,000		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
ratio_1221_90	0.55**	-0.01	0.53**	-0.18	0.59***	-0.23	0.63***	-0.19
	(0.24)	(0.29)	(0.23)	(0.24)	(0.18)	(0.21)	(0.20)	(0.27)
unmarried_15	-0.14	-1.17*	-0.06	-0.81	0.36	-0.08	0.31	-0.24
	(0.67)	(0.60)	(0.64)	(0.56)	(0.49)	(0.54)	(0.53)	(0.58)
Observations	586	329	569	318	528	297	446	242
ratio_1221_90	0.34***	0.19	0.32***	0.17	0.17*	0.01	0.25**	0.10
	(0.12)	(0.14)	(0.12)	(0.13)	(0.10)	(0.10)	(0.13)	(0.13)
unmarried_15	0.45	0.45	0.55	0.36	-0.05	0.31	-0.52	-0.09
	(0.36)	(0.36)	(0.36)	(0.36)	(0.27)	(0.29)	(0.38)	(0.40)
Observations	764	763	762	759	688	687	377	396

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: What if use year 2000 sex ratio rather than 1990? same specification as Wei and Zhang (2011, JPE)

	Full sample		Subsample that removes the following potential outliers: Rural & Urban inc. or expend. < 3,000 (< 5,000 for urban)		Bottom & Top 5%		Bottom & Top 5% no explicit marriage status	
	son	daughter	son	daughter	son	daughter	son	daughter
sex ratio 2000	0.19	0.38	0.42	0.17	0.23	-0.20	0.18	-0.09
(county level)	(0.45)	(0.46)	(0.43)	(0.44)	(0.31)	(0.38)	(0.33)	(0.43)
Observations	586	329	569	318	528	297	446	242
sex ratio 2000	0.17	-0.02	0.15	-0.00	-0.04	-0.37**	-0.02	-0.59**
(county level)	(0.20)	(0.23)	(0.20)	(0.23)	(0.16)	(0.18)	(0.21)	(0.24)
Observations	764	763	762	759	688	687	377	396

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Summary Statistics of sex ratio 2000 v.s. sex ratio 1990

Rural Counties	Mean	Max	Min	SD	Observations
ratio age0-9, census 1990	1.09	1.49	0.85	0.09	122
ratio age0-9, census 1990 (Wei and Zhang, 2011)	1.09	1.23	1.01	0.04	122
ratio age10-19, census 2000	1.09	1.27	0.77	0.06	122
Urban Cities	Mean	Max	Min	SD	Observations
ratio age0-9, census 1990	1.07	1.30	0.70	0.09	77
ratio age0-9, census 1990 (Wei and Zhang, 2011)	1.08	1.24	1.02	0.04	77
ratio age10-19, census 2000	1.07	1.23	0.92	0.05	77

Table 6: Migration for the specific age12-21 group: the rich rural v.s the poor rural

	Full sample		Subsample that removes outliers	
	son	daughter	sr outliers of son	sr outliers of daughter
sex ratio 0-9	1.11*	0.01	1.13*	-0.18
(county 1990)	(0.64)	(0.56)	(0.64)	(0.55)
sex ratio 10-19	0.90*	0.50	0.87*	0.46
(county 2000)	(0.52)	(0.57)	(0.53)	(0.59)
sex ratio 0-9	0.93*	-1.13	1.49**	-0.84
(county 1990)	(0.80)	(0.77)	(0.71)	(0.69)
sex ratio 10-19	-2.34*	-0.69	-1.26	-1.02
(county 2000)	(1.31)	(1.45)	(1.31)	(2.07)

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Results comparisons by county and household income

(a) Rural

rural	data sample variables	sex ratio 1990 census son s.r.	sex ratio 1990 census daughter s.r.	sex ratio 2000 census son s.r.	sex ratio 2000 census daughter s.r.
full sample	ratio_1221_90/00	0.54**	-0.13	0.19	0.38
		(0.23)	(0.27)	(0.45)	(0.46)
rich county	ratio_1221_90/00	1.08**	0.80	1.15***	0.53
		(0.53)	(0.66)	(0.43)	(0.54)
middle county	ratio_1221_90/00	0.44	-0.58	-2.22**	-1.22
		(0.33)	(0.41)	(0.95)	(0.98)
poor county	ratio_1221_90/00	0.56	0.01	-2.29	0.03
		(0.54)	(0.50)	(1.49)	(1.38)
rich household	ratio_1221_90/00	1.00*	1.48	-0.07	0.46
		(0.56)	(1.01)	(0.64)	(0.60)
middle household	ratio_1221_90/00	0.90**	-0.07	1.44*	-1.68
		(0.35)	(0.45)	(0.83)	(1.97)
poor household	ratio_1221_90/00	0.00	-0.43	-1.48	0.69
		(0.33)	(0.35)	(0.93)	(0.80)

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

(b) Urban

urban	data sample variables	sex ratio 1990 census son s.r.	sex ratio 1990 census daughter s.r.	sex ratio 2000 census son s.r.	sex ratio 2000 census daughter s.r.
full sample	ratio_1221_90/00	0.38***	0.24*	0.17	-0.02
		(0.12)	(0.13)	(0.20)	(0.23)
rich county	ratio_1221_90/00	-0.23	0.10	0.68*	0.26
		(0.26)	(0.25)	(0.36)	(0.38)
middle county	ratio_1221_90/00	0.53***	-0.03	0.56	-0.08
		(0.18)	(0.15)	(0.48)	(0.45)
poor county	ratio_1221_90/00	0.64***	1.00***	-0.60	-0.86**
		(0.24)	(0.28)	(0.39)	(0.41)
rich household	ratio_1221_90/00	0.57*	0.30	0.45	0.40
		(0.30)	(0.26)	(0.30)	(0.39)
middle household	ratio_1221_90/00	0.16	0.23	0.22	0.13
		(0.18)	(0.22)	(0.43)	(0.48)
poor household	ratio_1221_90/00	0.48**	0.23	-0.01	-0.62*
		(0.19)	(0.19)	(0.38)	(0.36)

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Marriage migration is a major form of migration and cause of female migration

Marriage as a Cause of Migration (1990)(%)

Destination	Both sexes	Men	Women
Intra-provincial			
All	14.0	2.0	28.0
City	3.0	1.0	6.0
Town	8.0	1.0	15.0
Rural areas	19.0	3.0	35.0
Inter-provincial			
All	14.0	2.0	30.0
City	2.0	1.0	6.0
Town	8.0	1.0	18.0
Rural areas	20.0	3.0	40.0

Table 9: Competitive Hypothesis: Evidence from CHFS 2015

	Full sample		Subsample that removes outliers	
	son	daughter	sr outliers of son	sr outliers of daughter
sex ratio 5-14	0.66**	0.75**	0.68**	0.66*
(county, 2010)	(0.31)	(0.37)	(0.32)	(0.37)
sex ratio 10-19	0.54	0.78*	0.61*	0.77*
(county, 2010)	(0.36)	(0.43)	(0.37)	(0.43)
Observations	796	635	752	604
sex ratio 5-14	0.04	0.22	0.03	0.19
(county, 2010)	(0.32)	(0.24)	(0.30)	(0.23)
sex ratio 10-19	0.04	0.35	0.04	0.37
(county, 2010)	(0.23)	(0.26)	(0.23)	(0.26)
Observations	2,010	1,556	1,915	1,484

Notes: 1. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Sex ratios and saving rates across province: Left-hand-side variable=(Y-C)/Y

	full	urban	rural	full	urban	rural
sex ratio for age cohort 5-19	-0.03	0.05	0.24*			
	(0.10)	(0.09)	(0.14)			
sex ratio for age cohort 10-19				-0.03	0.04	0.21*
				(0.08)	(0.07)	(0.12)
Per capita income (log)	0.20***	0.14***	0.39***	0.20***	0.15***	0.39***
	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.05)
share of population of aged 0-14	-0.25	0.04	0.80	-0.25	0.04	0.81
	(0.49)	(0.42)	(0.70)	(0.49)	(0.42)	(0.70)
share of population of aged 15-64	-0.37	0.03	0.22	-0.37	0.02	0.21
	(0.46)	(0.39)	(0.65)	(0.46)	(0.39)	(0.65)
share of SOE emplyoment in total labor force	-0.05	-0.21***	0.04	-0.05	-0.21***	0.03
	(0.07)	(0.06)	(0.10)	(0.07)	(0.06)	(0.10)
Constant	-0.74	-0.87*	-2.76***	-0.74	-0.86*	-2.74***
	(0.55)	(0.49)	(0.77)	(0.55)	(0.49)	(0.77)
provincial fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	214	214	214	214	214	214
R^2	0.77	0.85	0.74	0.77	0.85	0.74

Notes: 1. Robust standard errors (cluster at provincial level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1.