The Expansion of Higher Education and Household Saving in China*

Christopher Bollinger† Xiaozhou Ding‡ Steven Lugauer§

Abstract

We examine whether access to higher education impacts household saving rates. A 2-period model of household saving decisions demonstrates why increased college opportunities induce households with children to save more. We examine this theory using survey data from Chinese households during the unprecedented education expansion. Using estimates of the change in the expected probability of college attendance, we estimate the effect on household saving rates by comparing households before and after the reform. We find that a 10-percentage point increase in the probability of going to college raises the saving rate by 5.9 percentage points.

Keywords: Higher Education Expansion, Saving, China

JEL Classification Codes: D12, E21, I22, J13

*For helpful comments and suggestions, we thank Shiyu Cheng, Chad Curtis, Donna Ginther, Hao Guo, Chenxu Hu, Olga Malkova, Jenny Minier, Min Qiang Zhao, Shunan Zhao, editor Suqin Ge, two anonymous referees, and seminar and conference participants at the University of Kentucky Macro Group Meeting, CES North America Annual Conference, Kentucky Economic Association Annual Conference, SEA Annual Meetings, WEAI Conference, SOLE Annual Meeting, Chongqing Technology and Business University, China Meeting of the Econometric Society at Jinan University, Nanjing Audit University, Biennial Conference of China Development Studies at Shanghai Jiaotong University, University of Kentucky Applied Micro Group Meeting, and International Symposium on Human Capital and Labor Markets at Central University of Finance and Economics.

†Department of Economics, University of Kentucky, Lexington, KY 40506. Email: crboll@email.uky.edu.

‡Department of Economics, Dickinson College, Carlisle, PA 17013. Email: dingx@dickinson.edu.

§Department of Economics, University of Kentucky, Lexington, KY 40506. Email: steven.lugauer@uky.edu.

1 Introduction

Beginning in 1999, the Chinese government rapidly expanded the previously constrained Chinese higher education system. The expansion increased the number of students allowed into Chinese universities by over 40 percent in the first year alone. While education in China is heavily subsidized, costs to families of the students were high relative to overall income and student loans were uncommon, requiring households to pay these costs out of accumulated savings or current income (Shen and Li, 2003). In recent years, education expenditures have accounted for nearly 10 percent of Chinese household expenditures (Chi and Qian, 2016). Thus, we examine evidence for an unintended consequence of the expansion: an increase in the rate of household savings. China's household saving rate has increased dramatically since economic reforms began in 1978. Before the reforms, households typically saved less than five percent of their income. Saving gradually trended up during the 1980s and 1990s, before rapidly accelerating after 2000. Now, households save over twenty-five percent of their income, on average. Several theories have been put forward to explain the high saving rates, including demographic changes (Modigliani and Cao, 2004; Curtis, Lugauer and Mark, 2015; Imrohoroglu and Zhao, 2018*a*; Ge, Yang and Zhang, 2018), income uncertainty (Chamon, Liu and Prasad, 2013), private expenditures (Chamon and Prasad, 2010), economic reforms (He et al., 2018), and gender-related issues (Wei and Zhang, 2011; Zhou, 2014). None of the theories fully explains the changes. We provide evidence that the college expansion was an additional factor in the rise of household savings rates in China.

To motivate our empirical approach and to clearly articulate the potential connection between education expansion and saving rates, we present a simple two-period theoretical model of household saving. The model is designed to highlight college saving decisions in a two period world where there is uncertainty in college admission. The key implication is that when the probability of admission to college rises, saving rises as well, for some households. This leads us to a testable hypothesis that, conditional on other factors which would determine the baseline saving rates prior to the expansion, families experiencing a larger rise in expectation of attendance will have higher saving rates. We exploit this intuition in our empirical estimation.

We use China's 1999 higher-education expansion as a natural experiment and estimate how the policy-induced changes in the likelihood of college affected saving rates. Guided by our theoretical findings, we focus on identifying families with school aged children whose expectation of college opportunities for those children changed. These families are likely to have higher saving rates - in part to catch up - than other families, even those who would have expected to send their child to college prior to the expansion. We use 1995 and 2002 data from the China Household Income Project (CHIP). We use four distinct samples: families with college aged children before and after the expansion and families with school aged children before and after the expansion. We use the college aged samples to estimate models of college attendance and then use the primary school aged samples to estimate the savings model, using the change in probability to measure the intensity of treatment. We estimate the savings models both prior to the expansion and after the expansion. Much of our variation derives from differences in the timing and magnitude of the expansion across provinces. While the expansion has been large across all provinces, in 2002, there was substantial variation.

Our approach is consistent with other literature examining the impact of the higher education expansion. For example, Che and Zhang (2018) use an approach similar to ours to show that the Chinese college expansion increased human capital and improved productivity. Feng and Xia (2018) examine how the college expansion affected technology adoption in Chinese firms. Also, see Li and Xing (2010), Knight, Deng and Li (2017), Li et al. (2017), and Hu and Bollinger (2017, 2021). These papers primarily focus on labor market implications, while we connect college expansion to saving behavior.

Consistent with the theoretical model, our main finding is that the expansion has an economically and statistically significant effect on household saving rates. A 10 percentage point increase in the college probability increases the average household's saving rate by 5.9 percentage points. This result provides evidence that the education expansion increased saving for households with school-age children. Our results appear to be driven by households experiencing a large increase in the possibility of college, while the few households who might always have planned on college had similar savings. Our main findings are robust to a variety of alternative specifications including interactions with important demographic variables, omitting provinces with the largest changes in enrollment rates and including households without a dependent child. We also carefully consider other policy changes that occurred in China, including housing reform, public healthcare reform, and the reorganization of State-Owned Enterprises.

The rest of our paper is organized as follows. Section 2 documents the saving rates among Chinese households and provides background on the education expansion. Section 3 presents the model of household saving that motivates our empirical analysis. Section 4 introduces the data, and Section 5 discusses our estimation strategy. Section 6 presents the main empirical findings. Section 7 contains additional analysis, and Section 8 concludes.

2 Saving and Higher Education in China

We are interested in how the expansion of higher education affects household saving decisions. In this section, we document the large increase in Chinese household saving rates over time and provide background on the relevant changes to education policy. The remainder of the paper links these two phenomena, both theoretically and empirically.

2.1 Household Saving Rates in China

Throughout the paper, we define the household saving rate as

$$\text{Saving rate} = \frac{\text{Income} - \text{Expenditure}}{\text{Income}} = 1 - \frac{\text{Expenditure}}{\text{Income}}. \tag{1}$$

Income is measured by self-reported total annual income, including wages, other income, and asset returns. Expenditure is measured by total consumption expenditure. Both variables are available in China Household Income Project surveys. We use current (unadjusted) prices throughout, except where noted. Figure 1 plots the urban household saving rate in China from 1981 to 2011 using National Bureau of Statistics data. Saving slowly trended up before 2000; after that, though, the saving rate exploded and the slope of saving rate increased as well. In recent years, the typical Chinese household has been saving 30 percent of their income, a truly astonishing number. There is wide-spread interest in understanding why Chinese households save so much.

Several theories have been put forward. One strand of the literature has focused on the economic reforms that have been ongoing since 1978. Chamon, Liu and Prasad (2013) argue that the reforms have led to uncertainty over income, pensions, and healthcare, inducing Chinese households to save more for precautionary reasons.[1] As an example of this, He et al. (2018) show that the reform of state-owned enterprises in the 1990s increased unemployment and other risks for Chinese households, generating a dramatic increase in saving. Similarly, Choi, Lugauer and Mark (2017) show that high income growth coupled with high income uncertainty helps to explain why Chinese households currently save so much

[1] See Chamon and Prasad (2010) for more on how the burden for many expenditures shifted away from collectives and onto households. Also, see Yoo and Giles (2007).

more than US households. Other papers, including Song and Yang (2010), Lugauer and Mark (2013), Song et al. (2015), and Curtis (2016), have studied the interaction between saving and aggregate growth within the context of China.

A second strand of the literature has focused on the implications of China's fertility policies. Modigliani and Cao (2004) and Curtis, Lugauer and Mark (2015) appeal to the life-cycle hypothesis of household saving to argue that the large demographic changes, caused in part by the One Child Policy, have increased saving rates over time.[2] Several papers have documented specific examples of how demographics impact saving in China. Imrohoroglu and Zhao (2018*b*) argue that increases in expected old-age health costs and the aging population act together to increase savings. Rosenzweig and Zhang (2014) present evidence showing that the increase in the prevalence of co-residing elderly parents helps to explain household saving behavior. Wei and Zhang (2011) and Zhou (2014) show that a gender imbalance in the sex-ratio has interacted with life-cycle considerations to impact saving.

Our paper is related to these previous studies, in that the cost for higher education has shifted to families and saving for college varies over the life-cycle.[3] China radically and very rapidly altered its higher-education policies, which allows us to study households that experienced a sudden change in their expectations for sending their children to college. We next provide background on these policy changes.

2.2 Background on Higher Education Enrollment Policy

College admission in China is based purely on the Gaokao (college entrance exam) administered every year in June. Students are admitted on a provincial basis, and the Ministry of Education decides the number of students from each province through the coordination between local governments. The enrollment rates vary substantially across provinces. So, students compete for admission with other applicants within their province. If a high school graduate fails to get into a college, they can retake the exam and reapply in the following year. Due to the Cultural Revolution, the entrance exam was suspended between 1968 and 1978. In 1978, China initiated sweeping economic reforms, including reinstating the college entrance exam and enrolling new students into universities.

Figure 2 shows the number of students matriculating, total number enrolled, the number of college

[2]Also, see Bairoliya, Miller and Saxena (2018), Banerjee et al. (2015), Chao, Laffargue and Yu (2011), Choukhmane, Coeurdacier and Jin (2018), Curtis, Lugauer and Mark (2017), Jia et al. (2021), and Lugauer, Ni and Yin (2019).

[3]In a paper related to ours, Chen and Yang (2012) use the education reform in China to test the theory of precautionary saving. Also, note that housing in China shares these characteristics; see Chen, Yang and Zhong (2020).

teachers and staff, and the number of higher institutions by year. There were over 400,000 new students enrolled in 1978, and about 850,000 students total. The number of students gradually increased until 1999. Then, China instituted the higher education expansion, and enrollment exploded. In 1999 alone, the number of newly enrolled students increased by more than 40 percent. Over 1.5 million high school students began college in 1999, and by 2012 nearly 7 million students were starting college each year. Figure 3 plots this unprecedented increase in college enrollment alongside the national enrollment rate. The jump in the enrollment rate after 1999 (from below 40 percent to over 60) is readily apparent. Our analysis focuses on this sudden policy change.

The central Chinese government designed and controlled the 1999 higher education expansion, and it reflected the government's political objectives. China had experienced various social and economic headwinds in the 1990s, including a sizable reduction in employment at State-Owned Enterprises. The children of the large Chinese baby-boom generation were facing a bleak employment outlook, and the government was urged to find a solution.[4] According to lore, Min Tang, an economist at the Asian Development Bank in China, proposed enrollment expansion in a November 1998 letter to Premier Zhu Rongji. The hope was that doubling college enrollment within three years would stimulate investment in services, construction, and other related industries and would ultimately increase consumption (Wan, 2006). It was also suggested that households could use their savings to pay for college tuition and expenses. The plan was implemented almost immediately. So fast was the enrollment policy changed that typical Chinese households could not have anticipated its timing nor size. Moreover, as we show below, the policy had a differential impact across households. The increase in enrollment rates differed by province as well as by other observable household characteristics. We exploit the massive policy change to estimate how the expansion in higher education affected household saving behaviors.

The link between education and saving that we propose is a simple one. Households had to save in order to afford college tuition and expenses. Student loans were not readily available (Wang et al., 2014), and while the prevalence of college increased, so did the cost.[5] Figure 4 plots the total annual tuition and fees per student (in current prices) for different types of education over time. Due to compulsory education laws, the cost of attending elementary school and middle school remained close to zero. In contrast, the

[4]Even though the One Child Policy went into effect before 1980, the sheer size of the previous generation meant that in aggregate they still had many children reaching adulthood.

[5]Credit constraints often bind even when loans can be obtained by some. See Keane and Wolpin (2001), Cameron and Taber (2004), Carniero and Heckman (2009), Brown, Karl Scholz and Seshadri (2012), Sun and Yannelis (2016) and Malkova and Braga (2018) for example.

already high tuition and fees for college went up by a factor of four. Education was (and remains) a major expenditure category for Chinese households (Chamon and Prasad, 2010). According to Yao et al. (2011), most Chinese households save for education motives, and Wei and Zhang (2011) show that 76 percent of single-female-child households save for education.

The college expansion reform implemented by the Ministry of Education of China serves as a good natural experiment for analyzing the impact of education expansion on household saving. First and foremost it was a large exogenous shock. Second, the magnitude of the expansion varied by province even though it was implemented nation-wide. Each province received a different "quota" to determine how many students were to be admitted, and families could not easily relocate due to the Hukou registration system. Thus, the change in the expected probability of attending college depended on geography; however, as we show below, other household characteristics mattered, as well. Our empirical strategy leverages the exogenous variation in expected enrollment rates at the household level. Before detailing our approach for analyzing the data, we next present a theory for why an expansion in higher education relates to household saving behavior.

3 A Model of Households Saving for College

This section contains a simple two-period model of household saving decisions. Our goal is to motivate our empirical analysis by presenting an explicit theory showing how enrollment rates can impact saving decisions. The key choice by households in the model is whether or not to save enough to cover college tuition and expenses. An expansion in college opportunities induces some families to increase their savings.

For tractability, we consider households with log utility over consumption C in period 1 and expected consumption C' in period 2.

$$U = \ln C + \mathrm{E}[\ln C']. \tag{2}$$

The household receives income Y in period 1, which is split between consumption and saving S. Thus, the budget constraint in the first period is

$$C + S = Y, \tag{3}$$

where households begin life with no assets and saving must be positive. We abstract from discounting future utility and set the return on saving to zero, as these considerations do not materially alter the analysis.

The household saves in order to consume in period 2, but also to potentially pay college tuition for their child. The saving decision is made before knowing whether or not their child will be admitted into college. However, the household does know the probability p that their child gets in (where $0 < p \leq 1$). Within the model, we interpret the education expansion as an increase in p. When a child is accepted into college, the household can pay tuition and expenses τ out of savings. If the household has not saved enough, $S < \tau$, then the child cannot attend college. The household's second period budget constraint is

$$C' = S, \tag{4}$$

if $S < \tau$. If $S \geq \tau$, then, with probability $1 - p$ the budget constraint remains $C' = S$. But, with probability p, if $S \geq \tau$, then the second period budget constraint becomes

$$C' = S + \theta - \tau, \tag{5}$$

where θ is the benefit from sending a child to college.

We assume that the benefits of sending a child to college outweigh the costs; thus, $\theta > \tau > 0$, and if $S \geq \tau$, an admitted child always enrolls in college. There are many other ways to model the benefits from college, but this modeling choice is tractable and easy to interpret. The household must decide whether it wants to consume less in period 1 in order to have enough saved up to potentially pay college tuition τ. The benefit θ can be taken literally as the incremental increase in old-age support from having a college-educated child, but also more broadly to include non-monetary benefits.

With this set-up, we can examine how household saving decisions change in response to the policy of increasing the enrollment rate p. Consider a household that is saving too little to send their child to college, $S = Y/2 < \tau$. For this household, the marginal utility of consuming in period 1 exceeds the expected marginal utility of saving for college. Even if their child is accepted into college, the household will not obtain θ in the second period. However, for this low-saving household, an increase in p could induce the household to save more (enough to cover tuition τ). That is, an increase in p increases the expected utility of saving for college. The exact p in which the household is indifferent between saving τ (and possibly sending their child to college) and continuing to save less than τ is given by the following proposition.

Proposition 1 (Threshold p**)** *For a household with* $Y < 2\tau$*, the expected utility from saving* τ *and saving less than* τ *become exactly equal at* $p^t > 0$*, where the threshold* p^t *is given by:*

$$p^t = \frac{\ln\left[\frac{Y^2}{4(Y-\tau)\tau}\right]}{\ln\frac{\theta}{\tau}}. \tag{6}$$

The key implication from Proposition 1 is that as higher education is expanded (p increases), a low-saving household (saving less than τ) will start saving τ once p reaches the threshold p^t.[6] This theoretical result helps to motivate our main empirical strategy, in which we estimate the idiosyncratic increases in p across households and quantify how changes in p impact saving.

An increase in p may not impact all households in the same way, though. Households with older children or no children might be affected differently, or not at all. The model could be expanded to consider these additional household types; however, the remainder of the paper focuses on households with young children. Even among households with young children, the impact could differ. Some families, experiencing a smaller increase in p, may not hit their threshold, leaving their saving behavior unchanged.[7] Our empirical analysis exploits variation in the (estimated) changes in p across households (as suggested by the theoretical model) to estimate how the change in college probability affects household saving. We introduce the data next and then provide the full details of our empirical strategy.

4 The Data

We use household level information from the China Household Income Project (CHIP).[8] The CHIP consists of repeated cross-sections of data from household surveys that were conducted in five waves across 12 provinces. It is the most widely-used micro data set on Chinese households. The survey contains questions on income and expenditures, as well as other household characteristics such as geographic

[6]We derive Proposition 1 in Section A of the Online Appendix in a straightforward way. We find the threshold p by equating the lifetime utilities from either saving for college or not. The Appendix also shows that the threshold p^t decreases with the benefits derived from college θ, increases with college costs τ, and decreases with income Y, over the relevant range of variable values.

[7]Another set of households (e.g. with $Y > 2\tau$) may have been saving enough to send their children to college, $S \geq \tau$, prior to the education expansion. These high-income households could even reduce their saving (but only as low as τ) as p goes up. We show this in the Appendix, and the result is intuitive. An increase in p for a household already saving more than τ merely increases the chance of obtaining θ. This increase in expected income in period 2 lets the household consume more (and save less) in period 1.

[8]As mentioned above, the education expansion likely impacted some types of households differently than those with young children. Therefore, using micro data is critical to our identification strategy; it also allows us to net out other factors affecting saving that might be correlated with the education expansion (e.g. income). We return to these issues below.

location, number of children, and educational attainment. We use the 1995 and 2002 waves of urban households, which bracket the 1999 college expansion.

The raw data contains 6,929 household observations in 1995 and 6,835 in 2002. We focus on households with school-age children, but we also make use of households with college-age children. We define school-age as between 6 and 18 because it covers the usual elementary through high school years. We define college-age as between 18 and 23, and consider all households containing any college-age child as a college household. Households with children between 6 and 18 (but none between 18 and 23) count as school-age children households. We do not use households in which all children are older than 23 or younger than 6. To hone in on prime-age workers, we drop households whose head's age is less than 25 or above 60.

Our key outcome of interest is the saving rate, as calculated for each household using Equation (1). We drop the few households that changed Hukou after the expansion but before their children took the college entrance exam. We also drop households from the 2002 sample in which a child took the exam before 1999 (when the expansion began) because we will use the households with college-age children to estimate the change in the likelihood of college enrollment (due to the policy). The final data set consists of 2,900 school-age children households and 1,218 college-age households in 1995 and 2,357 school-age children households and 973 college-age households in 2002.

Table 1 reports summary statistics by household type (college or school-age children) for 1995 and 2002, using current (unadjusted) prices. The average saving rate for school-age children households rises considerably between 1995 and 2002, more than a 30 percent increase from 10.7 to 14.2. As the large standard deviations attest, there is ample variation in saving across households to exploit. In the remainder of the paper, we link the saving behavior to the expansion of higher education. Note, that over the same time period, the average saving rate among households with college-age children actually decreased. We do not directly study this decrease, but it might be related to the larger share of college-age children attending college in 2002 (18.3 percent) versus 1995 (10.1 percent). We use this observed increase in college attendance to estimate the changes in the expectation of attending college for younger children.

The remainder of Table 1 reports the means and standard deviations for the control variables that we include in our regressions.[9] Many of the controls also had large changes over time. For example, China's rapid growth pushed up incomes, assets, and expenses. Other changes can be traced to specific policies. The decreasing number of children was likely due to fertility policies set in the 1970s. While changes in

[9] Appendix B provides definitions for each variable.

home ownership and employment at State-Owned Enterprises (SOEs) can be traced to privatization policies enacted in the late 1990s (Chen, Yang and Zhong, 2020; Berkowitz, Ma and Nishioka, 2017; Chen and Wen, 2017). Importantly, we not only control for variables related to these other policy initiatives, but we also allow their impact on saving rates to vary over time. We also run several additional robustness checks to show that these other policies do not drive our results.

As mentioned, we use the sample of households with college-age children to estimate the change in college probability. However, since quotas for college admittance were set at the province level, we also use province specific information on enrollment rates. For each year and province, we approximate the enrollment rate with the ratio of new college students to the number of senior high school graduates.[10] Even though it is possible that students go out of their home provinces for college, the vast majority of them remain in their home provinces. To support the idea that college enrollments are correlated with college probabilities, we supplement our analysis using the Census 2000 data. Appendix Table C1 shows the percentage of college students in a province that are from this province. We use number of college students from the province divided by total number of students attending college in that province. In most cases, more than 70% of college students in that province were originally from that province. Municipalities like Beijing have smaller shares due to a larger higher education system therefore they admit more students from other provinces. We address this issue in robustness checks by dropping households reside in Beijing. Using the Census 2000 data we find supportive evidence that the provincial level college enrollment rates are good predictors for college probabilities. We select school-age people (18-22) and merge college enrollment rates in the college-going year with each individual. Then we perform probit analyses with different specifications to see whether they can meaningfully determine the likelihood of going to college. The results are in Appendix Table C.2.[11] The estimates for provincial enrollment rates across columns are statistically significant, which provide a consistent story that enrollment rates at the province level can predict college status.

Figure 5 shows the estimated college enrollment rates from the 12 provinces represented in our dataset from 1990 to 2002. In general, there is an upward trend, although many provinces dip just prior to the expansion due to the cohort effect where more high school students graduated starting in 1996. After

[10] As far as we know, publicly available data on provincial college admission rates does not exist. Fan et al. (2017) uses university-specific cutoff scores collected from newspapers and websites to examine college expansion in China, and Hu and Bollinger (2017) use the number of people taking the college entrance exam and subsequent enrollment at the national level.

[11] Linear regressions have similar results.

1998, enrollment rates increase for all provinces. But neither the enrollment rate levels nor the changes in enrollment rates are uniform across provinces. These differences in enrollment changes were driven by policy, and they therefore help us to identify the impact on savings. The next section (Equation 11, in particular) provides further details.

5 Empirical Approach

The theoretical model in Section 3 emitted a straightforward relationship between the policy-induced change in a household's expectation of college attendance and its saving rate. However, the available data (comprised of repeated cross-sections) does not track the same households over time, and it does not contain a measure of the expected probability of attending college. Ideally, we would estimate a difference-in-differences model where we would have an indicator for families who were affected by the policy. Some families were still not able to send their child to college (e.g. if their child still did not meet the academic criterion). Some families met the criterion for entrance prior to the expansion. Thus, our ideal specification would be a differences-in-differences specification:

$$SR_{ijt} = \alpha Post_t + \gamma Treat_{ij} + \pi Treat_{ij} Post_t + X'_{ijt}\delta + \lambda_j + u_{ijt}. \tag{7}$$

The dependent variable is the saving rate (SR_{ijt}) for household i from province j in year t. $Post_t$ is a dummy variable that equals one if the household is observed in year 2002 after the expansion, i.e., $Post_{2002} = 1$. $Treat_{ij}$ is also a binary variable that equals one if the household was affected by the policy. The vector X_{ij} not only contains the control variables listed in Table 1 but also a quadratic function of household income due to the nonlinear relationship between income and saving rate (Huggett and Ventura, 2000). Vector λ_j represents province fixed effects, and u_{ijt} captures measurement error. The coefficient γ measures any difference between the treated and control group prior to the treatment. The coefficient α measures a shift in savings before and after treatment not associated with treatment. The coefficient π is the treatment effect. Alternatively, estimates could be obtained by two regressions, one for each time period with the treatment indicator in each. The treatment effect is then the difference between the two coefficients on treatment (Post - pre). This latter specifications is algebraically equivalent to interacting the $Post$ variable with all X's, and thus allows for differences in response to other variables.

The problem, however, is that we cannot directly identify, in families with school aged children, who would find their expectation of college attendance changed. Thus we do not have a treatment indicator readily measured. To address this we estimate models of college attendance using the sample of families with college-age children. The difference in the predicted probabilities serves as a measure of the treatment variable, we label as Δp_{ij}. We use the provincial enrollment rates for the families to identify the expansion impact. We then estimate the following Equation (8) twice; once using the 1995 sample of households with school-age children and once using the 2002 sample of households with school-age children.

$$SR_{ij} = \pi \Delta p_{ij} + X_{ij}'\delta + \lambda_j + u_{ij}. \tag{8}$$

The coefficient (π) on the change in college probabilities (Δp) is our key parameter of interest. We obtain two estimates, one for 1995 (prior to the policy change) and one for 2002 (after the expansion). In each separate regression, π measures how much of a difference having a higher Δp makes for saving. The difference between the two estimates, $\pi^{2002} - \pi^{1995}$, then, is similar to a differences-in-differences estimate, with the difference attributable to the education expansion.

To see how this compares to a difference-in-differences estimator, we consider Equation 7 in each of the two time periods. In the post period, Equation 7 resolves to

$$SR_{ij1} = \alpha + (\gamma + \pi)\,\Delta p_{ij} + X_{ij1}'\delta + \lambda_j + u_{ij}. \tag{9}$$

Thus the coefficient on our proxy for treatment, Δp_{ij}, in the post period is $\gamma + \pi$. For the pre-period, Equation 7 resolves to

$$SR_{ij0} = \gamma \Delta p_{ij} + X_{ij0}'\delta + \lambda_j + u_{ij}. \tag{10}$$

The coefficient corresponding to π is the difference between the coefficient on Δp_{ij} in each period. As noted above, we do not have the ideal treatment variable. We use the difference in probabilities as a measure of the likelihood that a particular household would be affected by the change in test score cutoff.

Note that the coefficients (δ and λ's) are estimated separately for 1995 and 2002. The impact from the controls can vary over time.[12] Recall, for example, that the housing market changed during the late 1990s.

[12]Thus, the use of micro data allows us to net out changes in the relationship between other observables and saving behavior by households with young children in a way that would not be possible with only aggregated provincial data. Relatedly, as we discuss below, we can estimate a household specific change in enrollment probability.

Our regressions allow the relationship between home ownership and savings to change, accordingly. We return to housing (and other) reforms in Section 7.1. Alternatively we could have simply regressed the saving rates on the probability of attendance in each year, and differenced the coefficients. Unsurprisingly, the results are similar, however this is not exactly a difference-in-differences estimate, as it forces the coefficient on treatment to be zero prior to the treatment period. We prefer our approach, and focus on that methodology here. The alternative approach is available from the authors, and indicates a larger effect.

To construct Δp_{ij} we estimate Equation (11) two times, once using households with college-age children in the 1995 sample and separately using the 2002 college-age households.

$$p_{ij} = X'_{ij}\beta + \gamma f(ER_i) + \zeta_j + \nu_{ij}. \tag{11}$$

The dependent variable (p_{ij}) equals one if the household's college-age child is enrolled in college, and zero otherwise. The share of college-age children attending college jumps from about 10 percent in 1995 to over 18 percent in 2002; see Table 1. The control variables are the same as above, and ζ represents a full set of province fixed effects. $f(ER_i)$ is a polynomial function of enrollment rates. Note that ER_i is the household specific enrollment rate corresponding to the year that the household's child took the college entrance exam in province j. The enrollment rate is not only different across provinces, but also different for households if their children's college-going years are different.

With estimates of β, γ, and ζ from both 1995 and 2002 in hand, the second step is to calculate the predicted probability of college enrollment, p, for each household in the subset of families with young (age 6 to 18) children. To do so, we plug in the observable data (X_{ij} and the province) into the estimated version of Equation (11) to get an estimate of p for each household with young children.[13] We do this twice for each household (and for both the 1995 and 2002 samples) - once using the vector of coefficient estimates based on 1995 college-age households and once using the estimates from 2002. We then use the difference in the two estimated probabilities as our measure of the household's change in the expected probability of sending their children to school. Specifically, ER_j^{2002} is assigned to households with school-age children in province j in 2002 and ER_j^{1995} is assigned to households with school-age children in province j in 1995.

$$\Delta p_{ij} = X'_{ij} \cdot \left(\hat{\beta}^{2002} - \hat{\beta}^{1995}\right) + \left(\hat{\gamma}^{2002} ER_j^{2002} - \hat{\gamma}^{1995} ER_j^{1995}\right) + \left(\hat{\zeta}_j^{2002} - \hat{\zeta}_j^{1995}\right). \tag{12}$$

[13] These children had not yet taken the college entrance exam, so we use the survey year enrollment rate for ER.

We assume that the variation in Δp across households was exogenously driven by the higher education expansion. If this were strictly true, then we would only need to calculate Δp for the 2002 households and estimate Equation (8) one time. However, the treatment intensity could have been correlated with unobservable household characteristics. Thus, we estimate Δp and π for the 1995 households in order to difference out any spurious (or pre-existing) correlation between household saving and Δp. In a sense, the Δp estimates for the 1995 households act as a placebo to check the efficacy of the treatment on the 2002 households.

Our other identifying assumption is that households experiencing different treatments (low changes in college probability versus high) would have had similar trends in their saving rates if the college expansion had not occurred. In Section 6.2, we use aggregate data to provide evidence that the saving rate trends were in fact similar across households in the years leading up to the college expansion.

6 Higher Education Expansion's Effect on Saving

This section reports our main findings, which are obtained by estimating Equation (8) via ordinary least squares and Equation (11) as a probit model. According to our estimates, a 10 percentage point increase in the expected probability of attending college leads to a 5.9 point increase in the average household's saving rate. The estimates are statistically significant and robust to a host of specification alterations. We begin by discussing the probit model.

6.1 Who Goes to College?

The 1999 higher education expansion increased college opportunities for nearly all families. To calculate household-specific changes in the expectation of college attendance, we first estimate Equation (11) using the subset of families with children in their college years (age 18 to 23), separately for 1995 and 2002. The dependent variable is the dummy variable indicating whether a household has a college child, and the controls include all those listed in Table 1, as described above. The Appendix reports the probit estimate details.

We then use the resulting coefficient estimates from Equation (11) to calculate an expected college probability for each household with young children (ages 6 to 18). Figure 6 shows the distribution of predicted college probabilities for the 2002 households using both sets of estimated coefficients. The distribution shifts to the right when using the 2002 coefficients. After the 1999 college expansion, the predicted probability of attending college increases markedly. The increase is mostly driven by the increase in provincial enrollment rates, but other factors matter, too.

Finally, we calculate the change in college probability using Equation (12). Again, we do this for both the 1995 and 2002 households with young children. To get an initial sense of the relationship between the change in college probability and saving rates, we also regress the household specific saving rates on the controls listed in Table 1 (for the 2002 households). Figure 7 plots a fitted curve (smoothed by a local polynomial regression (Fan and Gijbels, 1996)) of the residuals against the change in college probability. The curve lines up closely with the predictions of our theoretical model. Small increases in college probability appear to be less correlated with saving. However, once the change becomes large enough, saving rates are higher, too. We next estimate this relationship using a quasi-difference-in-differences approach, in order to net out unobservable factors for households experiencing different changes in college probability.

6.2 The Main Empirical Findings

Table 2 reports the main regression results based on Equation (8). The estimate of π (the marginal effect of the change in college probability) for the 2002 sample equals 0.486. The effect on saving is quantitatively large and statistically different from 0 at better than the 5 percent significance level. Households for which the probability of attending college went up saved more. The estimate for the 1995 sample is much smaller (-0.108). Recall, the 1995 sample was not (yet) actually subject to the college expansion, while the 2002 sample was. Thus, we interpret the difference in the two coefficient estimates as the effect due to the policy change. This 'difference-in-differences' (DD) estimate equals 0.594, and it is statistically different from zero. Taken literally, the DD estimate implies that, on average, a 10 percentage point (policy driven) increase in the expected probability of college increased the typical households saving rate by about 5.9 percentage points. This effect is very large, but given the dramatic increase in saving observed in China after 2000, it is not implausible. Also note, the impact on aggregate saving is not straightforward to calculate because households with older children or no children likely reacted to the policy changes, possibly through general equilibrium effects on future employment or wages. The alternative specification is pooling the two years together and interact each variable with the year dummy including provincial variables such as population, urban employment, and GDP growth rate. We provide the result in Column (3) and it shows similar estimate compared to the more preferred flexible specification. Chamon and Prasad (2010) compare households with different age children and find that the effect of having a school age child increases saving rate by 2 percentage points before 1999 and 4 percentage points after 1999. The difference is approximately 2 percentage points. However, they do not utilize cross-province differences and examine the saving rate in an aggregate level.

Table 3 reports two sets of robustness checks.[14] All the regressions include the full set of controls, but we do not report the coefficient estimates to save space. In the first set (columns 1-4), we drop households residing in Beijing (columns 1 and 2) and Chongqing (columns 3 and 4). We omit Beijing because it has the largest higher education system in China. The overall enrollment rates exceed one which means it absorbs more college students than its own high school graduates. It can also be seen from the Appendix Table C1 that local students have much smaller percentage compared to other provinces. We omit Chongqing because by the year 1995 it was still part of Sichuan Province. It became a municipality in 1998, in order to ensure that our estimate is not driven by relative larger size of colleges after Chongqing became a municipality, we omit it to test the robustness. Dropping Beijing and Chongqing is somewhat arbitrary, but our estimates for the changes in college probability are largest for households in these two provinces, on average (see Figure 5 and Appendix Table C.4). The resulting DD estimates are smaller than our main DD estimates; however, the estimated effect is still very large and statistically significant. We retain the households from Beijing and Chongqing in the remainder of our analysis.

The second set of robustness checks adds households with zero dependent children in the sample. If it is the case that households with dependent children were expecting increases in college probability while in contrast households without children were not affected by such expansion, we should expect an even larger effect.[15] Again, we find that the increase in college probability leads households to save more.[16] We assign $\Delta p = 0$ to all households without dependent children, which has 686 households in 1995 and 633 households in 2002. The inclusion of this group results in an estimated impact of 8.34 percentage points, compared to our baseline result of 5.94. Overall, the results across the robustness checks reported in Table 3 reinforce our main results. The number of observations is less than in the main regressions, but the effects remain statistically significant and quantitatively large. The empirical results in Tables 2 and 3 are consistent with

[14] As an additional robustness check, we also have conducted a standard placebo check by randomly assigning households to different provincial enrollment rates. We find no significant impact on saving, providing additional support for our main results.

[15] Age-savings profiles show supportive evidence for households without children do not respond to the college expansion. The figures are in Appendix E.

[16] We also conducted the analysis using zero dependent child households as control group and households with child as treatment group. We found small positive impact. The results are available upon request.

the intuition coming out of our structural model that low-saving households ($S < \tau$) experiencing a large enough increase in the likelihood of college (p) will start to save more (i.e. Proposition 1).

We end this section by revisiting our identification strategy. A critical assumption for the difference-in-differences framework is that the differentially treated groups were not trending (before the treatment) in ways correlated with the treatment.[17] In our case, the concern is that the types of households experiencing large increases in their college probability were also the types of households already increasing their savings, anyway. This concern is valid, and we can even imagine that the policy makers might have wanted to target the treatment towards households with increasing rates of savings because these households would be most able to pay for college. With the cross-sectional data at hand, we cannot directly check the micro-data for pre-trends; however, we can take a more aggregated look at the data. Figure 8 plots changes in the average household saving rate before the education expansion (from 1995 to 1998) against enrollment rates during the enrollment increase (1998 to 2002), for each province. There exists little correlation, and the fitted line actually reveals a slightly negative relationship. In other words, the policy did not selectively increase enrollment rates more for the provinces in which saving rates were already growing the most before the policy was enacted. Recall that the household-level estimates for the change in college probability (Equation (11)) were a function of the provincial enrollment rates. So, from this perspective, the treatment on individual households does not seem to depend on their pre-policy saving behavior. Thus, we conclude that our analysis is not picking up pre-trends, but, instead, the saving behavior is driven by the response to changing college opportunities.

7 Other Explanations for China's High Saving Rates

This section examines the other main explanations for China's high household saving rates put forth in the literature (see Section 2). First, we examine policy changes enacted around the same time. Then, we examine demographic changes. As noted above, our regressions already include a number of controls aimed at accounting for these related factors.[18] Moreover, we think these other factors were unlikely to be linked to the pattern of increased college opportunities. However, since the policy reforms and demographic

[17] Another critical assumption is that households did not select into experiencing a higher increase in college probability, for example, by moving across provinces. This is unlikely to have occurred because of the speed of the policy implementation and the Hukou restrictions on migration. Just to be sure, we have checked that our results do not change when we remove the few households in the sample that changed their Hukou after the policy.

[18] Our main regressions also allow the coefficient estimates on these controls to change over time.

changes were so big and also impacted household saving, we now go beyond simple control variables and allow for interactions with our variable of interest. Throughout the many specifications, the impact of higher education expansion on saving rates remains large and statistically significant.[19]

7.1 Other Reforms

We begin by interacting the change in college probability (allowing for heterogeneous effects) with the variables related to three other large reforms: SOE reform, housing reform, and healthcare reform. Specifically, we estimate

$$SR_{ij} = W_{ij}'\delta + \pi \cdot \Delta p_{ij} + \phi \cdot Z_{ij} + \psi \cdot \Delta p_{ij} \cdot Z_{ij} + \lambda_j + u_{ij}, \tag{13}$$

where Z_{ij} represents a dummy variable capturing either SOE job status, public health coverage, or private house ownership (each is considered separately). Vector W is the vector X, including all covariates except Z.[20] Then, $\partial SR_{ij}/\partial \Delta p = \pi + \psi \cdot Z_{ij}$ is the marginal effect for each household. We report the average marginal effects, based off of the marginal effects across different Z_{ij}'s. We are interested in whether controlling for the heterogeneous impact across Z (i.e. the exposure to other reforms) alters our estimate (π) for the education expansion effect, and we find that it does not.

7.1.1 Employment at a State-Owned Enterprise (SOE)

Public sector jobs used to be part of the so-called "iron rice bowl" of social support. The SOE reform in the 1990s, however, led to massive layoffs and made even incumbent SOE workers less secure in their employment. As discussed above, several papers have shown that the new risk led to an increase in precautionary saving. If the households that experienced large increases in the likelihood of college were also facing greater employment risk, then our estimates could be conflating the two channels. Columns (1) and (2) of Table 4 provide evidence against this possibility. The coefficient estimate for Δp (0.424) in 2002 is close to the main estimate, even after controlling for the interaction with SOE status. Note, the interaction term is positive in both years, but small and statistically insignificant. The bottom line is that the DD estimate (5.97 percentage points) remains large and close to our baseline result.

[19]The Appendix also reports triple difference-in-difference estimates, which leverage these other factors affecting saving rates to further difference out potential trends coming from a third, omitted, dimension (Imbens and Wooldridge, 2009).

[20]To save space, the tables below do not report the coefficient estimates for the controls.

7.1.2 Public Health Coverage

Healthcare is another motive for precautionary saving (Chamon and Prasad, 2010), as well as life-cycle saving. Access to public health coverage has been in flux in China due to policy changes, migration and Hukou regulations, and population aging. Columns (3) and (4) of Table 4 report estimates controlling for heterogeneous effects based on having public health coverage. The resulting DD estimate of 0.618 is actually slightly higher than our baseline estimate. Interestingly, the coefficient estimate for the interaction term flips signs and becomes statistically insignificant after the college expansion.

7.1.3 Home Ownership

China also reformed its housing market in the 1990s. By 1998, most families were allowed to own their homes, and many households saved in the form of housing or in order to buy housing (required down payment rates were high). Again, as with the other reforms, we think that a connection between housing reform and exposure to the higher education expansion is unlikely at the household level. Columns (5) and (6) in Table 4 indicate that controlling for the interaction between the change in college probability and home ownership has little impact on our main estimate. The DD estimate (0.628) remains very large.

7.2 Household Demographics

We next consider regressions that control for interactions with variables capturing household demographic characteristics. We still apply the approach embodied in Equation (13), letting Z represent the various demographic factors of interest.

7.2.1 Number of Children

China enacted the One Child Policy in 1978, although enforcement of the policy varied over time and by location. Several papers have shown that Chinese households with fewer children (i.e. one) tend to save more and have drawn a connection between China's fertility policies and high saving rates. One reason why single child households might save more is to invest in their child's college education (a quantity/quality tradeoff similar to Becker and Lewis (1973)). In order to examine the potential heterogeneous effect across households with different numbers of children, we define a new variable *Single* that is equal to one if there is only one dependent child and zero if there are two or more children. We re-estimate Equation (13), interacting the new variable (Z = *Single*) with the change in the probability of college.

Table 5 columns (1) and (2) present the results. The resulting DD estimate for the average marginal effect equals 0.616, close to our baseline result. Note, the coefficient estimate on *Single*, while not statistically significant before the education expansion, is close to what Lugauer, Ni and Yin (2019) estimated, a 2.4 percentage point decrease in saving rates due to each additional child using a different data set and different estimation methodology.

7.2.2 Sex Composition

Some Chinese households may prefer having a son over a daughter because sons traditionally support their parents in old age. Relatedly, the One Child Policy may have led to the gender imbalance now prevelant in China. Households with sons invest differently for their child's education and also for marriage and housing purposes. Wei and Zhang (2011) show that these factors have had a large impact on household saving rates in China.[21] Therefore, we next interact college probability with a variable (Z = *Male*) that equals one if a household has only male children and is zero otherwise. We count mixed gender households (of which there are few) as non-male households. Columns (3) and (4) of Table 5 suggest that the male child effect is quite small. Moreover, our estimate of the college expansion effect (DD=0.605) remains large.

7.2.3 Life-Cycle Saving and Age Effects

Simple life-cycle theory predicts that households of different ages can have different saving patterns. The younger households in our sample (household head with an age closer to 25) might be net borrowers, while older households (age close to 60) rapidly accumulate assets in anticipation of retirement. The older the household head, the closer the child is to college age, on average. To address this, we next interact the household head's age ($Z = Age$) with the change in college probability. Columns (5) and (6) of Table 5 present the results. Once again, the DD estimate (0.611) remains similar to our baseline result.

In summary, the 1999 higher education expansion had a large effect on household saving rates, and this finding holds across households with different exposures to other reforms and different demographic characteristics. While the economic reforms and demographic factors likely also impacted household saving (our regressions do not rule out these alternative stories), the increase in college opportunities had an additional effect above and beyond these other explanations.

[21] Cai et al. (2019) and Zhou (2014) also discuss related issues.

8 Conclusion

In this paper, we exploit the policy-induced increase in college enrollment to estimate the impact on household saving rates in China. We find that the expansion in higher education resulted in higher household saving rates, especially for previously low-saving households experiencing a large increase in the probability of sending their children to college. Through this saving channel, the college expansion likely affected China's economic growth rate and international capital flows. Our findings are robust to a host of specification modifications, as well as to controlling for other concurrent policy changes and on-going demographic changes.

China's education expansion was unique in several ways. The policy change was large and swiftly implemented, and college enrollment levels were relatively low before the reforms. Also, China had strict fertility controls at the time, effectively shutting down the quantity channel in the classic fertility theory of a quantity/quality trade-off. These characteristics helped inform our estimation strategy, but it remains a question as to whether our findings are applicable to other countries and situations. Similarly, our regressions necessarily capture the short run responses, from when the policy was new and unexpected, rather than the long-run response. We leave an exploration of these important issues to future research. Understanding the link between saving and education is important and informative for policy makers for many reasons. First, understanding saving behavior itself is, of course, important, and education expenses are a key part of household expenditures, over 10 percent on average in China. Second, many countries including the U.S. are debating whether to increase access to higher education, and it is natural to question how the additional schooling will be financed. Research is needed on how the expansion of education affects the economy and how households respond to the increase in college opportunities. Our story is straightforward. Households save to pay for college. Expanding higher education means that more families will expect their children to attend university; hence, these households save more.

References

Bairoliya, Neha, Ray Miller, and Akshar Saxena. 2018. "The Macroeconomic Impact of Fertility Changes in an Aging Population." *Working paper.*

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics*, 7(1): 22–53.

Becker, Gary S., and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2, Part 2): S279–S288.

Berkowitz, Daniel, Hong Ma, and Shuichiro Nishioka. 2017. "Recasting the Iron Rice Bowl: The Reform of China's State-Owned Enterprises." *Review of Economics and Statistics*, 99(4): 736–747.

Brown, Meta, John Karl Scholz, and Ananth Seshadri. 2012. "A New Test of Borrowing Constraints for Education." *Review of Economic Studies*, 79(2): 511–538.

Cai, Xiqian, Yi Lu, Jessica Pan, and Songfa Zhong. 2019. "Gender Gap Under Pressure: Evidence from China's National College Entrance Examination." *Review of Economics and Statistics*, 101(2): 249–263.

Cameron, Stephen V., and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *Journal of Political Economy*, 112(1): 132–182.

Carniero, Pedro, and James J. Heckman. 2009. "The Evidence on Credit Constraints in Post-Secondary Schooling." *The Economic Journal*, 112(482): 705–734.

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

Chamon, Marcos D., Kai Liu, and Eswar S. Prasad. 2013. "Income Uncertainty and Household Savings in China." *Journal of Development Economics*, 105: 164–177.

Chao, Chi Chur, Jean Pierre Laffargue, and Eden Yu. 2011. "The Chinese Saving Puzzle and the Life-Cycle Hypothesis: A Revaluation." *China Economic Review*, 22(1): 108–120.

Chen, Binkai, and Rudai Yang. 2012. “Understanding China’s Rising Saving Rate: The Role of Higher Education Reform.” In *The New Chinese Economy*. 137–153. Springer.

Chen, Binkai, Xi Yang, and Ninghua Zhong. 2020. “Housing demand and household saving rates in china: Evidence from a housing reform.” *Journal of Housing Economics*, 49: 101693.

Chen, Kaiji, and Yi Wen. 2017. “The Great Housing Boom of China.” *American Economic Journal: Macroeconomics*, 9(2): 73–114.

Che, Yi, and Lei Zhang. 2018. “Human Capital, Technology Adoption and Firm Performance: Impacts of China’s Higher Education Expansion in the Late 1990s.” *The Economic Journal*, 128(614): 2282–2320.

Chi, Wei, and Xiaoye Qian. 2016. “Human Capital Investment in Children: An Empirical Study of Household Child Education Expenditure in China, 2007 and 2011.” *China Economic Review*, 37: 52–65.

Choi, Horag, Steven Lugauer, and Nelson C. Mark. 2017. “Precautionary Saving of Chinese and U.S. Households.” *Journal of Money, Credit and Banking*, 49(4): 635–661.

Choukhmane, Taha, Nicholas Coeurdacier, and Keyu Jin. 2018. “The One-Child Policy and Household Savings.” *Working paper*.

Curtis, Chadwick C. 2016. “Economic Reforms and the Evolution of China’s Total Factor Productivity.” *Review of Economic Dynamics*, 21: 225–245.

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

Curtis, Chadwick C., Steven Lugauer, and Nelson C. Mark. 2017. “Demographics and Aggregate Household Saving in Japan, China, and India.” *Journal of Macroeconomics*, 51: 175–191.

Fan, Elliott, Xin Meng, Zhichao Wei, and Guochang Zhao. 2017. “Rates of Return to Four-Year University Education: An Application of Regression Discontinuity Design.” *The Scandinavian Journal of Economics*.

Fan, Jianqing, and Irene Gijbels. 1996. *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability*. Vol. 66, CRC Press.

Feng, Shuaizhang, and Xiaoyu Xia. 2018. "Endogenous Skill-Biased Technology Adoption: Evidence from China's College Enrollment Expansion Program." Human Capital and Economic Opportunity Working Group Working Papers 2018-099.

Ge, Suqin, Dennis Tao Yang, and Junsen Zhang. 2018. "Population Policies, Demographic Structural Changes, and the Chinese Household Saving Puzzle." *European Economic Review*, 101: 181–209.

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

Hu, Chenxu, and Christopher Bollinger. 2017. "China's Higher Education Expansion and College Premium: IV Estimations." *Working paper.*

Hu, Chenxu, and Christopher Bollinger. 2021. "Effects of cohort size on college premium: Evidence from China's higher education expansion." *China Economic Review*, 70: 101700.

Huggett, Mark, and Gustavo Ventura. 2000. "Understanding Why High Income Households Save More than Low Income Households." *Journal of Monetary Economics*, 45(2): 361–397.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1): 5–86.

Imrohoroglu, Ayse, and Kai Zhao. 2018*a*. "The Chinese Saving Rate: Long-Term Care Risks, Family Insurance, and Demographics." *Journal of Monetary Economics*, 96: 33–52.

Imrohoroglu, Ayse, and Kai Zhao. 2018*b*. "Intergenerational Transfers and China's Social Security Reform." *Journal of the Economics of Ageing*, 11: 62–70.

Jia, Ning, Steven Lugauer, Jinlan Ni, and Zhichao Yin. 2021. "The Effects of Relaxing the One-Child Policy." *Working Paper.*

Keane, Michael P., and Kenneth I. Wolpin. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, 42(4): 1051–1103.

Knight, John, Quheng Deng, and Shi Li. 2017. "China's Expansion of Higher Education: The Labour Market Consequences of a Supply Shock." *China Economic Review*, 43: 127–141.

Li, Hongbin, Yueyuan Ma, Lingsheng Meng, Xue Qiao, and Xinzheng Shi. 2017. "Skill Complementarities and Returns to Higher Education: Evidence from College Enrollment Expansion in China." *China Economic Review*, 46(March): 10–26.

Li, Shi, and Chunbing Xing. 2010. "China's Higher Education Expansion and its Labor Market Consequences." *IZA Discussion Paper*, , (4974).

Lugauer, Steven, and Nelson C. Mark. 2013. "The Role of Household Saving in the Economic Rise of China." *Global Interdependence, Decoupling, and Recoupling*, 131.

Lugauer, Steven, Jinlan Ni, and Zhichao Yin. 2019. "Chinese household saving and dependent children: Theory and evidence." *China Economic Review*, 57: 101091.

Malkova, Olga, and Breno Braga. 2018. "Hope for the Family: The Effects of Merit Aid Programs on Parental Outcomes." *Working paper.*

Modigliani, Franco, and Shi Larry Cao. 2004. "The Chinese Saving Puzzle and the Life-Cycle Hypothesis." *Journal of Economic Literature*, 42(1): 145–170.

Rosenzweig, Mark R., and Junsen Zhang. 2014. "Co-residence, Life-Cycle Savings and Inter-Generational Support in Urban China." *NBER Working Paper.*

Shen, Hong, and Wenli Li. 2003. *A review of the student loans scheme in China.* Bangkok: UNESCO Bangkok.

Song, Zheng, and Dennis T. Yang. 2010. "Life Cycle Earnings and the Household Saving Puzzle in a Fast-Growing Economy." *Manuscript.*

Song, Zheng, Kjetil Storesletten, Yikai Wang, and Fabrizio Zilibotti. 2015. "Sharing High Growth across Generations: Pensions and Demographic Transition in China." *American Economic Journal: Macroeconomics*, 7(2): 1–39.

Sun, Stephen Teng, and Constantine Yannelis. 2016. "Credit Constraints and Demand for Higher Education : Evidence From Financial Deregulation." *Review of Economics and Statistics*, 98(March): 12–24.

Wang, Xiaojun, Belton M. Fleisher, Haizheng Li, and Shi Li. 2014. "Access to College and Heterogeneous Returns to Education in China." *Economics of Education Review*, 42: 78–92.

Wan, Yinmei. 2006. "Expansion of Chinese higher education since 1998: Its causes and outcomes." *Asia Pacific Education Review*, 7(1): 19–32.

Wei, Shang-Jin, and Xiaobo Zhang. 2011. "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China." *Journal of Political Economy*, 119(3): 511–564.

Yao, Rui, Feifei Wang, Robert O. Weagley, and Li Liao. 2011. "Household Saving Motives: Comparing American and Chinese Consumers." *Family and Consumer Sciences Research Journal*, 40(1): 28–44.

Yoo, Kyeongwon, and John Giles. 2007. "Precautionary Behavior and Household Consumption and Savings Decisions: An Empirical Analysis Using Household Panel Data from Rural China." *Review of Economics and Statistics*, 89(3): 534–551.

Zhou, Weina. 2014. "Brothers, Household Financial Markets and Savings Rate in China." *Journal of Development Economics*, 111: 34–47.

9 Tables

Table 1: Summary Statistics by Year and Children's Age

	School-Age		College-Age	
	1995	2002	1995	2002
Saving rate	0.107	0.142	0.141	0.128
	(0.226)	(0.342)	(0.238)	(0.334)
Child college attainment dummy			0.101	0.183
			(0.301)	(0.387)
Controls				
Number of kids	1.111	1.076	1.631	1.159
	(0.323)	(0.267)	(0.630)	(0.404)
Number of elderly	0.0838	0.0878	0.0755	0.0647
	(0.313)	(0.340)	(0.279)	(0.262)
Private house	0.410	0.782	0.475	0.757
	(0.492)	(0.413)	(0.500)	(0.429)
Housing accumulation fund	0.441	0.543	0.424	0.540
	(0.497)	(0.498)	(0.494)	(0.499)
Age	39.74	39.94	50.10	47.27
	(5.058)	(4.911)	(4.638)	(3.273)
Gender	0.633	0.652	0.664	0.638
	(0.482)	(0.477)	(0.472)	(0.481)
Currently employed	0.980	0.923	0.860	0.853
	(0.140)	(0.267)	(0.347)	(0.354)
College degree	0.066	0.110	0.100	0.050
	(0.247)	(0.313)	(0.300)	(0.219)
SOE job	0.822	0.325	0.825	0.342
	(0.383)	(0.468)	(0.380)	(0.475)
Public health	0.703	0.636	0.709	0.680
	(0.457)	(0.481)	(0.454)	(0.467)
Employment tenure	15.80	13.58	21.41	16.32
	(7.425)	(8.613)	(10.23)	(11.60)
Years of schooling	10.45	11.46	10.32	10.29
	(2.984)	(3.004)	(3.651)	(2.881)
Spouse years of schooling	9.657	10.67	8.714	9.582
	(3.476)	(3.571)	(4.148)	(3.439)
Annual income	1.349	2.244	1.655	2.335
	(0.705)	(1.408)	(0.900)	(1.394)
Annual expenses	1.175	1.805	1.379	1.901
	(0.624)	(1.222)	(0.797)	(1.133)
Total assets	1.076	3.954	1.454	4.056
	(1.473)	(10.05)	(3.932)	(5.415)
Observations	2,900	2,357	1,218	973

Notes: This table reports means for the 1995 and 2002 CHIP data, by household type. School-age households report having at least one child aged 6-18 and none older. College-age households have at least one child aged 18-23. Standard deviations are in parentheses. Income, expenses, and assets are in 10,000s of yuan at current (unadjusted) prices.

Table 2: College Probability's Effect on Household Saving Rates

	(1) 1995	(2) 2002	(3) Pooled
Δp	-0.108	0.486**	-0.068
	(0.077)	(0.241)	(0.081)
Post			0.069
			(0.481)
Post × Δp			0.489**
			(0.240)
No. kids	0.008	0.014	0.006
	(0.010)	(0.016)	(0.009)
No. elderly people	-0.004	0.006	-0.004
	(0.006)	(0.006)	(0.005)
Private house	0.007	-0.036**	0.006
	(0.006)	(0.016)	(0.006)
Housing accumulation fund	0.002	-0.001	0.001
	(0.003)	(0.012)	(0.003)
Head gender	0.005	0.028***	0.005
	(0.007)	(0.008)	(0.006)
Currently working	0.008	0.022	0.008
	(0.014)	(0.019)	(0.012)
Head tenure	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)
Head age	-0.001	-0.001	-0.000
	(0.000)	(0.001)	(0.000)
Head with college degree	-0.020	-0.023	-0.018
	(0.018)	(0.016)	(0.016)
Head years of schooling	-0.000	0.001	0.000
	(0.000)	(0.001)	(0.000)
Spouse years of schooling	0.000	0.005***	0.000
	(0.001)	(0.002)	(0.001)
SOE job	0.008	0.025*	0.007
	(0.005)	(0.013)	(0.005)
Public health	-0.010	0.019*	-0.007
	(0.008)	(0.010)	(0.007)
Annual income	0.602***	0.359***	0.602***
	(0.093)	(0.077)	(0.080)
Annual income square	-0.028*	-0.011	-0.028**
	(0.015)	(0.009)	(0.012)
Annual expense	-0.551***	-0.333***	-0.549***
	(0.084)	(0.038)	(0.075)
Total asset	-0.001	-0.003**	-0.001
	(0.002)	(0.001)	(0.002)
Population, millions			-0.009
			(1.614)
Urban employment, millions			0.119
			(1.048)
GDP growth rate			-0.023
			(2.745)
Observations	2,900	2,357	5,257
R^2	0.787	0.713	0.739
DD	0.594		
	(0.253)		

Notes: This table reports regression estimates of Equation 8 for 1995 and 2002. Income, expenses, and assets are in 10,000s of yuan and current prices. Row DD reports the difference in the coefficient estimates (π) on Δp. The regressions include province fixed effects, and the parentheses report bootstrapped robust standard errors clustered at province level with 500 repetitions.

Table 3: Robustness Checks

	No Beijing		No Chongqing		Add zero child	
	(1)	(2)	(3)	(4)	(5)	(6)
	1995	2002	1995	2002	1995	2002
Δp	-0.130	0.585***	-0.106	0.553**	0.065	0.899***
	(0.137)	(0.215)	(0.077)	(0.262)	(0.128)	(0.311)
Observations	2,690	2,225	2,900	2,268	3,586	2,987
R^2	0.780	0.709	0.787	0.710	0.773	0.380
DD	0.715		0.659		0.834	
	(0.254)		(0.273)		(0.318)	

Notes: Each regression is based on Equation 8. The dependent variable is saving rate. Each regression includes the full set of controls and province fixed effects. The parentheses report bootstrapped robust standard errors clustered at province level with 500 repetitions.

Table 4: Other Reforms

	SOE job		Public health		Private house	
	(1) 1995	(2) 2002	(3) 1995	(4) 2002	(5) 1995	(6) 2002
Δp	-0.132	0.424**	-0.146	0.521**	-0.092	0.477**
	(0.164)	(0.216)	(0.158)	(0.248)	(0.072)	(0.235)
Δp×SOE job	0.023	0.187				
	(0.148)	(0.158)				
Δp×Public health			0.036	-0.035		
			(0.123)	(0.132)		
Δp×Private house					-0.124	0.010
					(0.147)	(0.160)
SOE job	0.008	0.029**	0.008	0.025*	0.008	0.025*
	(0.007)	(0.013)	(0.005)	(0.013)	(0.005)	(0.013)
Public health	-0.010	0.019*	-0.009	0.018	-0.010	0.019*
	(0.008)	(0.010)	(0.010)	(0.011)	(0.008)	(0.010)
Private house	0.007	-0.036**	0.007	-0.037**	0.002	-0.036*
	(0.007)	(0.016)	(0.007)	(0.016)	(0.007)	(0.018)
Observation	2900	2357	2900	2357	2900	2357
R-square	0.787	0.713	0.787	0.713	0.787	0.713
Marginal effect	-0.113	0.484	-0.120	0.498	-0.143	0.485
	(0.080)	(0.240)	(0.092)	(0.235)	(0.104)	(0.239)
DD	0.597		0.618		0.628	
	(0.253)		(0.252)		(0.260)	

Notes: Each regression includes province fixed effects and the control variables. The parentheses report bootstrapped robust standard errors clustered at province level with 500 repetitions.

Table 5: Demographics

	Number of kids		Gender of kids		Life-cycle	
	(1) 1995	(2) 2002	(3) 1995	(4) 2002	(5) 1995	(6) 2002
Δp	-0.429	0.602*	-0.174	0.478*	-0.040	0.929*
	(0.331)	(0.347)	(0.130)	(0.247)	(0.577)	(0.501)
Single	0.032	0.003				
	(0.060)	(0.058)				
$\Delta p\times$ Single	0.340	-0.122				
	(0.309)	(0.247)				
Male			0.004	0.003		
			(0.003)	(0.008)		
$\Delta p\times$ Male			0.110	0.012		
			(0.125)	(0.124)		
Head age	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
$\Delta p\times$ Head age					0.004	-0.029
					(0.011)	(0.009)
Observation	2,900	2,357	2,900	2,357	2,900	2,357
R^2	0.787	0.713	0.787	0.713	0.787	0.713
Marginal effect	-0.126	0.490	-0.121	0.484	-0.096	0.515
	(0.086)	(0.249)	(0.086)	(0.241)	(0.140)	(0.249)
DD	0.616		0.605		0.611	
	(0.2638)		(0.256)		(0.286)	

Notes: Each regression includes province fixed effects and the control variables. The parentheses report bootstrapped robust standard errors clustered at province level with 500 repetitions.

10 Figures

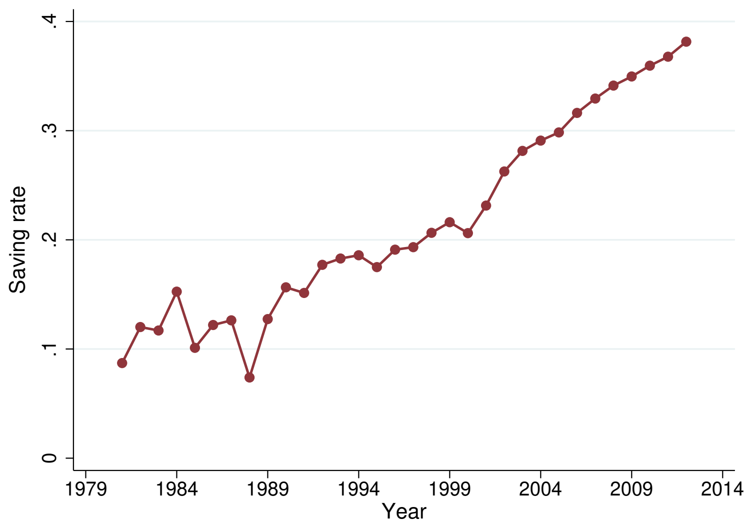

Figure 1: Household Saving Rates

Notes: The data is from the China Yearly Statistical Book and the National Bureau of Statistics. The variables used to construct the saving rate are per capita total income of urban households (y) and per capita annual cash consumption expenditure of urban households (e) at current prices. The household saving equals $(y - e)/y$.

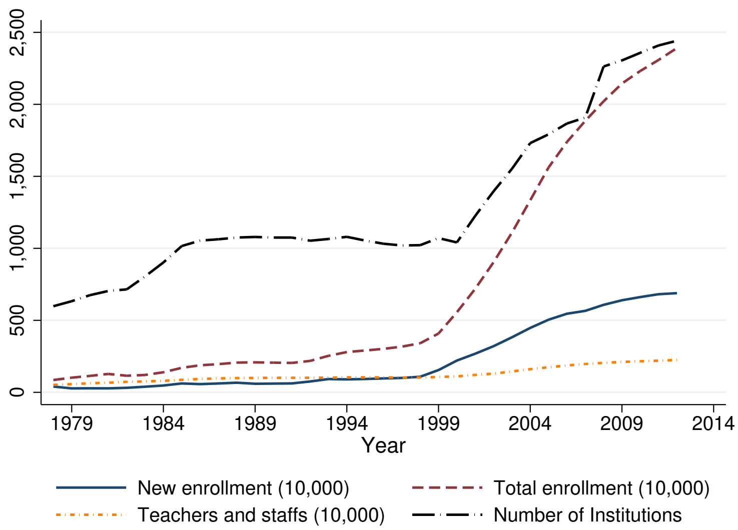

Figure 2: China Higher Education Expansion

Notes: The data comes from the China Yearly Statistical Book and the National Bureau of Statistics. The variables are in units of 10,000, except the number of institutions, which is unadjusted.

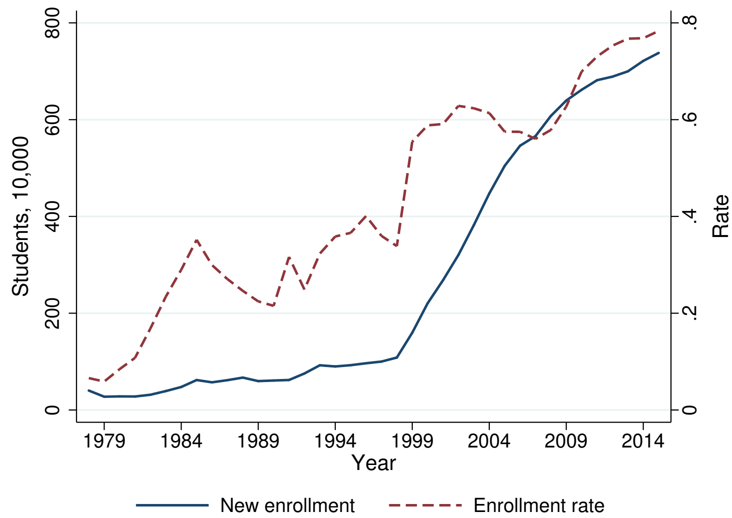

Figure 3: New Students Enrolled and the Enrollment Rate

Notes: The enrollment data is from the China Yearly Statistical Book and the National Bureau of Statistics. The enrollment rate is calculated by dividing the number of newly enrolled students by the number of national college entrance examination takers. The number of exam takers was collected from http://news.koolearn.com/20180606/1152629.html.

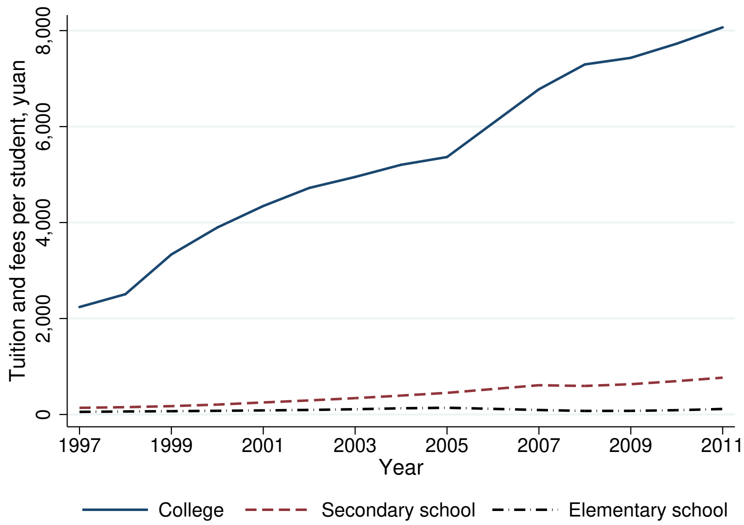

Figure 4: Cost of Education

Notes: The data comes from the China Yearly Statistical Book and the National Bureau of Statistics. The reported values are per-student at current (unadjusted) prices.

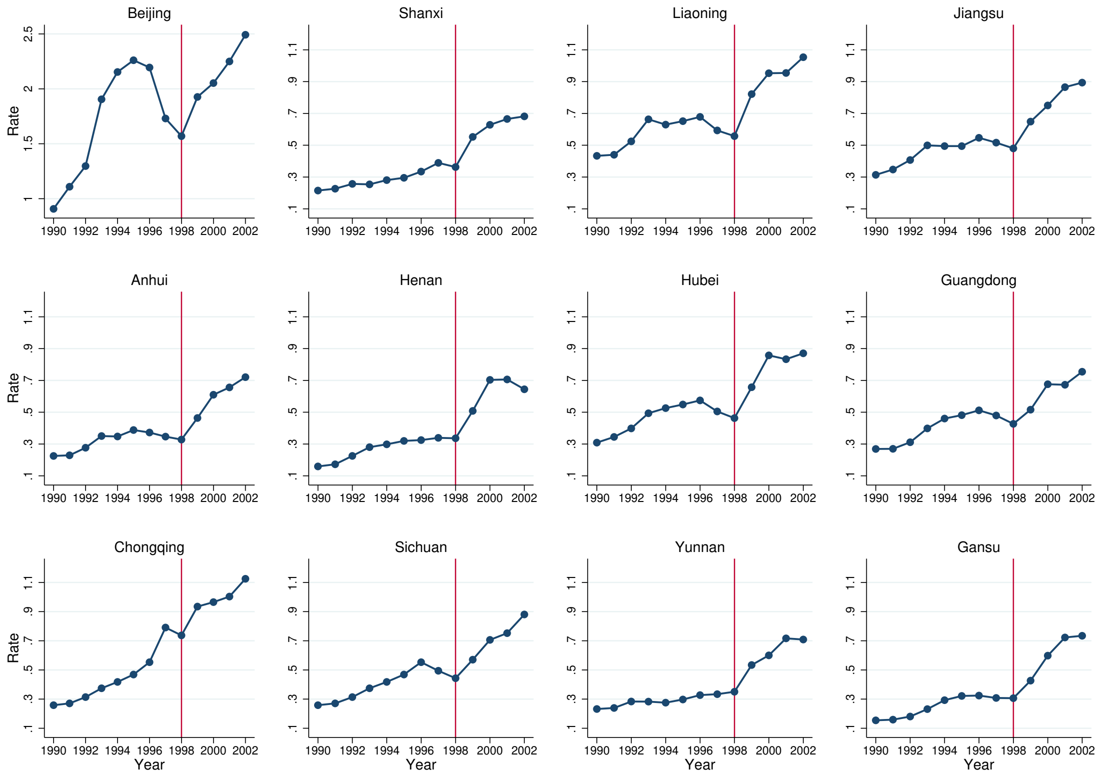

Figure 5: Enrollment Rate by Province

Notes: We obtained the underlying data from the National Bureau of Statistics. This figure plots the annual rate of planned new enrollment in regular higher institutions divided by the number of senior high school graduates, from 1990 to 2002 by province. Prior to 1997, Chongqing was part of Sichuan, so we use the Sichuan rate.

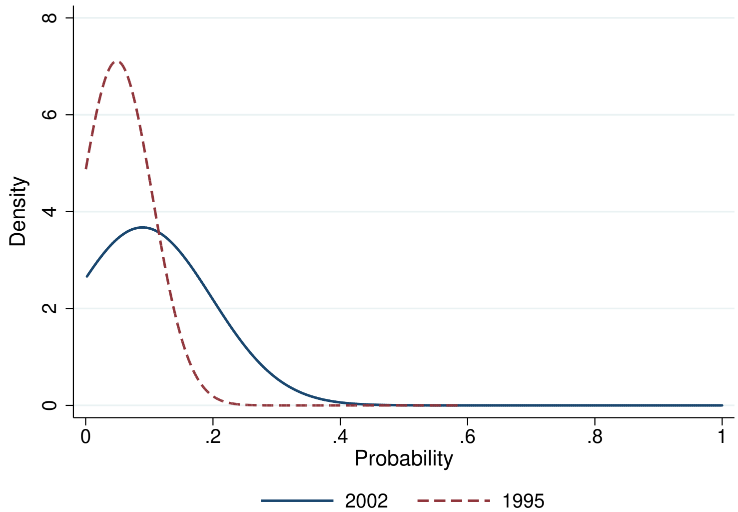

Figure 6: Estimated College Probability for 2002 Households

Notes: This figure shows the distribution of estimated college probabilities across the households surveyed in 2002, based on the coefficient estimates from both the 1995 and 2002 probit regressions of Equation 11.

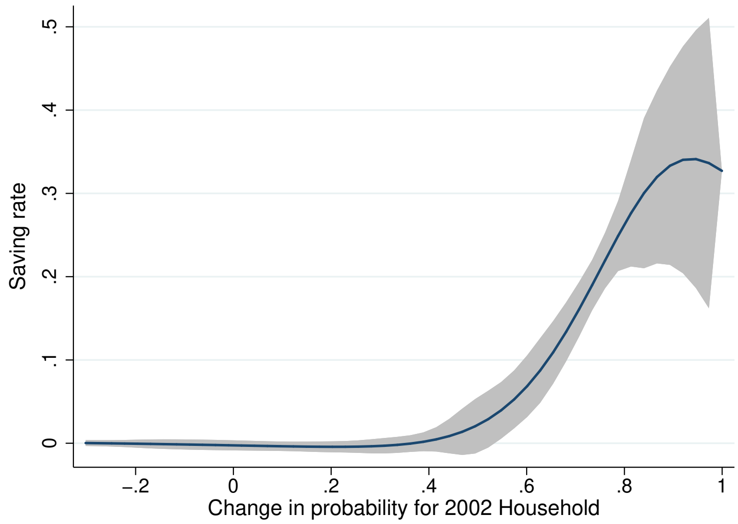

Figure 7: Saving Rate and Change of Probability for 2002 Households

Notes: This figure shows the relationship between saving rates and the change in college probability for 2002 households via local polynomial regressions.

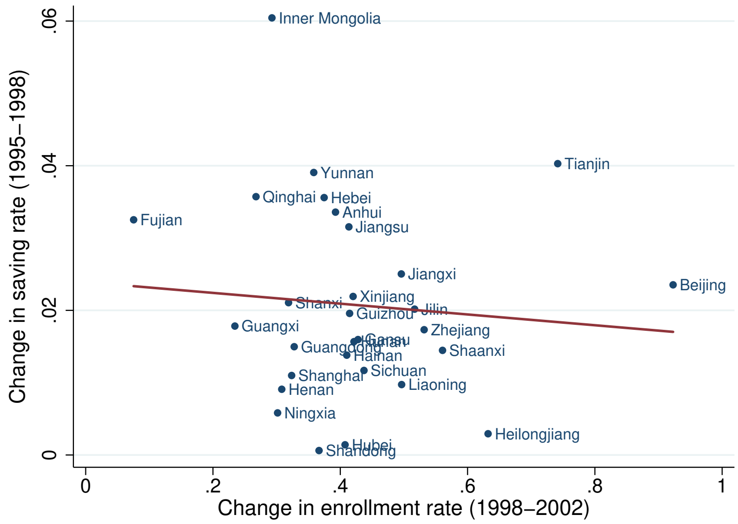

Figure 8: Test for Pre-Trends

Notes: The vertical axis measures the saving rate change between 1995 and 1998. The horizontal axis measures the subsequent enrollment rate changes between 1998 and 2002. The underlying data comes from the National Bureau of Statistics Yearly Statistical Book.

Online Appendix
The Expansion of Higher Education and Household Saving in China

A Structural Model

This section derives the results for the structural model that are discussed in the main text. The first sub-section derives Proposition 1 and the related properties of the household's college probability threshold. The second sub-section shows that saving can decrease with college probability for high-saving households.

A.1 Derivation of Proposition 1

We derive the threshold p^t in three steps. First, we calculate lifetime utility for a household (with $Y < 2\tau$) choosing $S < \tau$, at their optimal choice of S. Second, we calculate lifetime utility assuming the household saves enough to pay for college ($S \geq \tau$). However, we note that low income households (with $Y < 2\tau$) will not choose $S > \tau$; at most, they will choose $S = \tau$. These households will be at a 'corner solution'. Finally, we find the p that equalizes the two lifetime utilities (i.e. for $S < \tau$ and $S = \tau$).

Step 1: finding the optimal choice of saving (with $S < \tau$) is straightforward by substituting C and C' from the budget constraints into Equation (2)

$$U = \ln(Y - S) + \ln S.$$

[1] The first order condition is

$$-\frac{1}{Y-S} + \frac{1}{S} = 0. \tag{A.1}$$

Solving for the optimal choice of S gives

$$S^* = \frac{1}{2}Y. \tag{A.2}$$

Lifetime utility then equals

$$U = 2\ln\left(\frac{1}{2}Y\right). \tag{A.3}$$

Call this $U1^*$. Obviously, then, for a household with $Y < 2\tau$, the household cannot send its child to college with this saving choice.

[1]We note here, that a household with $Y > 2\tau$ would have no solution in the case of $S < \tau$, since they would be at a corner solution and the set is not closed.

Step 2: finding optimal saving for a household choosing $S \geq \tau$ (and that also has $Y < 2\tau$) requires two parts. First, we consider the constrained or 'corner solution' of $S = \tau$. Second, we argue that, for a low-income household ($Y < 2\tau$), an interior solution ($S > \tau$) always gives lower utility than the corner solution. Hence, a low-income household will save at most $S = \tau$.

First, finding the corner solution only requires setting $S = \tau$. So, the household receives θ in period 2 with probability p and has only its savings τ with probability $1 - p$. The resulting lifetime utility is

$$U = \ln(Y - \tau) + p \cdot \ln \theta + (1 - p) \cdot \ln \tau. \tag{A.4}$$

We call this $U2^*$.

Second, we argue that a low-income household will not choose $S > \tau$. In some sense, this is self-evident because, in the absence of the possibility of college, their utility was maximized by $S < \tau$. For a household (again, with $Y < 2\tau$) choosing $S > \tau$, the utility function is

$$U = \ln(Y - S) + p \ln(\theta + S - \tau) + (1 - p) \ln(S).$$

The marginal utility with respect to additional savings (first order condition) is

$$\frac{\partial U}{\partial S} = \frac{-1}{Y - S} + \frac{p}{\theta + S - \tau} + \frac{1 - p}{S}.$$

Setting this equal to zero and solving for $S > \tau$ (while imposing $Y < 2\tau$) emits no solution for any p from zero to one. The marginal utility from additional savings (above τ) is always negative. Indeed, simple algebra shows that the marginal utility expressed above is only positive if

$$\frac{p}{\theta + S - \tau} + \frac{1 - p}{S} > \frac{1}{Y - S}.$$

Consider, for example, the highest value the left hand side of the inequality can take (i.e. $p = 0$). For the inequality to hold, it must be that $Y > 2S$. But $Y > 2S$ contradicts $Y < 2\tau$ when $S > \tau$. Thus, no interior solution exists (utility is decreasing in S over this range), and the optimal choice is the 'corner solution' $S = \tau$. Utility is given by $U2^*$. Since $p = 0$ is the highest possible value, any $p > 0$ requires an even higher income for the inequality to hold, creating the same contradiction, and again leading to an optimal choice of $S = \tau$.[2]

Step 3: equating $U1^*$ (A.3) and $U2^*$ (A.4), and solving for p gives

$$\ln\left(\frac{1}{4} Y^2\right) = \ln(Y - \tau) + p \cdot \ln \theta + (1 - p) \cdot \ln \tau.$$

[2]Interestingly, with a high enough p, even some high-income households will be at their corner solution. For example, with $p = 1$, the inequality above becomes $Y > \theta + \tau$. So, households with income $2\tau < Y < \theta + \tau$ will also find it optimal to choose $S = \tau$.

Collecting the p terms gives

$$p(\ln \tau - \ln \theta) = \ln(Y - \tau) + \ln \tau - \ln\left(\frac{1}{4}Y^2\right)$$

$$p \ln \frac{\tau}{\theta} = \ln\left(\frac{(Y-\tau)\tau}{\frac{1}{4}Y^2}\right)$$

Hence, the college probability that leaves the household indifferent between saving $S = \tau$ and $S = \frac{1}{2}Y < \tau$ is

$$p^t = \frac{\ln\left[\frac{Y^2}{4(Y-\tau)\tau}\right]}{\ln \frac{\theta}{\tau}}. \tag{A.5}$$

Again, this result relies on the condition that $Y < 2\tau$. Note, a household facing $p > p^t$ will still choose $S = \tau$ because any p larger than p^t only increases the utility from saving at least τ, and, as the derivation implicitly shows, the household will remain at its constrained solution.

Following Proposition 1, we claim that p^t is decreasing in the benefits (θ) from college, increasing in the costs (τ), and decreasing in household income (Y) over the relevant set of parameter values. We show each of these in turn.

First, it is straightforward to see that $\partial p^t/\partial\theta < 0$ (as long as p^t is well defined, with $\tau < Y$). Thus, a higher return from college lowers p^t.

Second, applying a simple quotient rule, we find $\partial p^t/\partial\tau$

$$\frac{\partial p^t}{\partial \tau} = \frac{\frac{\partial \ln\left[\frac{Y^2}{4(Y-\tau)\tau}\right]}{\partial \tau} \cdot \ln\left(\frac{\theta}{\tau}\right) - \frac{\partial \ln\left(\frac{\theta}{\tau}\right)}{\partial \tau} \cdot \ln\left[\frac{Y^2}{4(Y-\tau)\tau}\right]}{\left(\ln \frac{\theta}{\tau}\right)^2}.$$

Clearly, $\left(\ln \frac{\theta}{\tau}\right)^2 > 0$ in the denominator. The first term in the numerator is

$$\begin{aligned}
\frac{\partial \ln\left[\frac{Y^2}{4(Y-\tau)\tau}\right]}{\partial \tau} &= \frac{1}{4}Y^2\left[-(Y-\tau)^{-2} \times (-1) \times \tau^{-1} + (Y-\tau)^{-1}(-\tau^{-2})\right] \\
&= \frac{1}{4}Y^2 \frac{2\tau - Y}{(Y-\tau)^2\tau^2}.
\end{aligned}$$

This again is positive, as long as $\tau < Y$. The second term in the numerator is also positive because

$$\begin{aligned}
\frac{\partial \ln\left(\frac{\theta}{\tau}\right)}{\partial \tau} &= \frac{\tau}{\theta} \cdot \left(-\frac{\theta}{\tau^2}\right) \\
&= -\frac{1}{\tau} < 0.
\end{aligned}$$

Hence,

$$\frac{\partial p^t}{\partial \tau} > 0.$$

Finally, it is straightforward to show that $\partial p^t/\partial Y < 0$.

$$\frac{\partial p^t}{\partial Y} = \frac{1}{\ln \frac{\theta}{\tau}} \cdot \frac{4(Y-\tau)\tau}{Y^2} \cdot \frac{2Y[4(Y-\tau)\tau] - Y^2(4\tau)}{16(Y-\tau)^2\tau^2}$$
$$= \ln \frac{\tau}{\theta} \cdot \frac{1}{Y} \cdot \frac{(Y-2\tau)}{(Y-\tau)} < 0.$$

The last inequality comes from the assumption that $\frac{1}{2}Y < \tau$. Therefore, p^t is decreasing in household income.

A.2 Saving Response for High-Saving Households

Consider a household saving more than τ, $S > \tau$. For example, a high-income household may have an optimal saving choice that exceeds the cost of college at any value of p, even $p = 0$. Let D be the saving in excess of tuition costs ($D = S - \tau$). Then, the utility function can be written

$$U = \ln(Y - \tau - D) + p\ln(\theta + D) + (1-p)\ln(\tau + D).$$

The first order condition with respect to excess savings D is

$$\frac{\partial U}{\partial D} = \frac{-1}{Y - \tau - D} + \frac{p}{\theta + D} + \frac{1-p}{\tau + D} = 0,$$

We are interested in the response of excess savings D to increasing college opportunities, i.e. $\partial D/\partial p$. Using the implicit function theorem gives

$$\frac{\partial D}{\partial p} = \frac{\frac{1}{\theta+D} - \frac{1}{\tau+D}}{\frac{1}{(Y-\tau-D)^2} + \frac{p}{(\theta+D)^2} + \frac{1-p}{(\tau+D)^2}}$$
$$= \frac{\frac{\tau-\theta}{(\theta+D)(\tau+D)}}{\frac{1}{(Y-\tau-D)^2} + \frac{p}{(\theta+D)^2} + \frac{1-p}{(\tau+D)^2}} < 0.$$

The inequality holds because $\tau < \theta$. So, households with $Y > 2\tau$ and $D > 0$ decrease their savings in response to an increase in p (but only to $D = 0$, as noted, no maximizing solution exists for $S < \tau$).

B Data

Our primary data source is the China Household Income Project (CHIP) 1995 and 2002 urban modules. The data was downloaded from http://www.ciidbnu.org/chip/index.asp. We merge the individual data (that contains earnings for all family members and detailed information on the household head) and household-level data and conduct our analysis at the household level. We also utilize China Yearly Statistical Books to obtain provincial characteristics. Table B.1 lists the definitions for the key variables.

Table B.1: Variable Definitions

Variable	Definition
Saving rate	Household income minus consumption, divided by income - Equation (1)
Child college	Dummy variable indicating a child aged 18 to 23 having attended college
Enrollment ratio	New college enrollees divided by high school graduates, by province
No. kids	Self-reported number of children
No. elderly people	Self-reported number of parents or grandparents of the household head
Private house	Dummy variable indicating private house ownerhsip
Housing accumulation fund	Dummy variable indicating the head has a housing accumulation fund
Head age	Self-reported age of head
Head male	Dummy variable indicating whether the head is a male
Head working	Dummy variable indicating the head's current employment
Head college degree	Dummy variable indicating whether head attended college
Head SOE job	Dummy variable indicating employment at a State-Owned Enterprise for head
Head public health	Dummy variable indicating whether head has health care provided by the state or work unit, or has compulsory medical insurance for serious diseases
Head tenure	Self-reported years of work in current job by head
Head years of schooling	Self-reported years of education for head
Spouse years of schooling	Self-reported years of education for spouse
Annual income	Household income, totaled across family members
Annual expenses	Self-reported household consumption expenditures
Total assets	Self-reported household total assets
Single	Dummy variable, equals 1 if a household has only one dependent child
Population	Provincial population in millions
Urban employment	Urban employment in millions
GDP Growth rate	Growth rate of provincial GDP

Notes: Income, expenses and assets are measured in current (unadjusted) prices.

C Additional Tables

This section provides supporting evidence for our main analysis.

Table C.1. presents the share of college students from home province. We use the Census 2000 data and keep college-age students (18-22) and calculate the share by dividing number of college students from the province by total number of college students in the province.

Table C.1: Percentage of Local College Students

Province	Share
Beijing	38.65%
Tianjin	59.72%
Hebei	81.59%
Shanxi	83.18%
Inner Mongolia	93.04%
Liaoning	79.55%
Jilin	73.70%
Heilongjiang	78.28%
Shanghai	64.68%
Jiangsu	83.04%
Zhejiang	88.87%
Anhui	85.16%
Fujian	92.11%
Jiangxi	80.37%
Shangdong	90.69%
Henan	85.54%
Hubei	72.26%
Hunan	80.82%
Guangdong	87.31%
Guangxi	89.15%
Hainan	66.96%
Chongqing	56.12%
Sichuan	75.96%
Guizhou	93.81%
Yunnan	86.49%
Tibet	78.38%
Shaanxi	58.51%
Gansu	74.52%
Qinghai	92.54%
Ningxia	83.79%
Xinjiang	97.15%

Notes: This table shows share of college students that are from the same province using the Census 2000 data.

Table C.2 presents the supplemental analysis that examines whether enrollment rates are good indicators for college probabilities. We use the college-age individuals and run the following probit regression:

$$\Pr(College_{ij}) = \Phi(\beta_0 + \beta_1 f(ER_{ij}) + \delta X_j + \gamma_j + u_{ij})$$

where $College_{ij} = 1$ if individual i from province j is a college student and 0 otherwise. X_j are provincial characteristics and γ_j is province fixed effect. ER_{ij} is the enrollment rate.

Table C.2: Probit Regressions of College Attendance Using Census 2000

	(1)	(2)	(3)	(4)	(5)
Enrollment rate	1.068***	4.539***	4.539***	1.385***	2.218*
	(0.391)	(1.613)	(1.613)	(0.448)	(1.218)
Enrollment rate square					-0.323
					(0.338)
Population				-0.002***	-0.002***
				(0.001)	(0.001)
GDP growth rate				2.754***	2.658***
				(0.670)	(0.666)
Urban employment				0.002***	0.002***
				(0.001)	(0.001)
Constant	0.153	-6.355**	-6.355**	0.099	-0.455
	(0.169)	(2.624)	(2.624)	(1.329)	(1.719)
Province FE		✓	✓	✓	✓
Observations	52,790	52,790	52,790	52,790	52,790

Notes: This table reports probit regression results of college attendance using census 2000 dataset.

This next section provides additional details for the probit regressions (Equation 11) used to estimate the changes in college probability.

Table C.3: Probit Estimates for the Probability of Attending College

	(1) 1995	(2) 2002
Enrollment rate	3.529***	-3.743**
	(1.320)	(1.562)
Enrollment rate square	-1.230***	0.393
	(0.404)	(0.337)
No. kids	0.115	0.041
	(0.110)	(0.125)
No. elderly people	0.051	0.114
	(0.185)	(0.153)
Private house	-0.262***	0.199
	(0.080)	(0.145)
Housing accumulation fund	-0.073	0.035
	(0.135)	(0.123)
Head gender	-0.061	-0.122
	(0.141)	(0.135)
Currently working	-0.149	0.002
	(0.188)	(0.126)
Head tenure	-0.008	0.002
	(0.006)	(0.005)
Head age	0.040***	0.035*
	(0.014)	(0.020)
Head with college degree	0.432***	0.569*
	(0.166)	(0.295)
Head years of schooling	0.054***	0.031
	(0.020)	(0.025)
Spouse years of schooling	0.064**	0.014
	(0.028)	(0.017)
SOE job	0.179	-0.024
	(0.281)	(0.099)
Public health	0.170	0.093
	(0.155)	(0.131)
Annual income	-0.023	-0.003
	(0.263)	(0.093)
Annual income square	-0.007	-0.009
	(0.035)	(0.007)
Annual expense	0.106	0.231***
	(0.081)	(0.070)
Total asset	-0.002	0.019*
	(0.013)	(0.011)
Population, millions	-0.075	1.278*
	(0.572)	(0.699)
Urban employment, millions	0.856	-0.686***
	(1.477)	(0.228)
GDP growth rate	-2.101*	-15.742**
	(1.153)	(7.149)
Observations	1,218	973

Notes: This table reports coefficient estimates from the probit regressions. The dependent variable is a dummy variable for whether the college-age child has attended college. Income, expenses, and assets are in 10,000s of yuan at current prices. All specifications include province fixed effects. Robust standard errors clustered by province are reported in parentheses.

After obtaining the first stage probit results, we predict each household's probability of having a college child and then average the probabilities across households within each province to obtain the mean probability before and after the expansion. Table C.4 shows the change in probability by province for both the 1995 and 2002 survey samples.

Table C.4: Change in College Probabilities by Province

	1995	2002
Beijing	0.021	0.150
Shanxi	-0.016	0.002
Liaoning	-0.019	0.018
Jiangsu	-0.011	0.034
Anhui	0.004	0.039
Henan	0.004	0.034
Hubei	0.038	0.102
Guangdong	-0.064	-0.014
Sichuan	-0.006	0.031
Yunnan	-0.045	-0.021
Gansu	0.023	0.081
Chongqing		0.154
Observations	2,900	2,357

Notes: This table shows the average change in college probability by province after the first step probit estimation for households surveyed in 1995 and 2002. Provincial variables are not used as we include province fixed effects.

D Triple Difference-in-Difference Estimates

This section reports triple difference-in-difference (DDD) estimates, exploiting the variation due to the other policy reforms and the evolving demographics. The DDD estimates help to further control for confounding trends coming from a third, possibly omitted, dimension. The dependent variable remains the household saving rate, and the below tables report the estimated coefficients for the change in college probability.

Table D.1 reports the estimates when dividing the sample (of households with school age children) based on the variables related to other policy reforms. Looking at panel A, column (3) indicates that the DD estimates for both SOE and non-SOE households are large. The DDD estimate is not tiny, but it is not statistically different from zero. Panel B shows that the DD estimate is large for both households with and without access to public health. Although, we note that the impact on the saving rates of households without public health is markedly larger. The overall pattern remains the same in panel C. As in panels A and B, the coefficient estimate on the change in p is relatively small in 1995 and much larger in 2002. The estimated impact is similar for households with and without privately owned homes.

Table D.1: Triple Difference-in-Difference using Other Reforms

	(1) 1995		(2) 2002		(3) DD	
	Coeff.	Std. Err	Coeff.	Std. Err	Mean	Std. Err
A. SOE job						
No	-0.252	(0.314)	0.359*	(0.195)	0.611	(0.370)
Yes	-0.089	(0.080)	0.867**	(0.408)	0.956	(0.416)
DD	0.163	(0.324)	0.508	(0.452)		
DDD					0.345	
					(0.556)	
B. Public health						
No	-0.341	(0.230)	0.612*	(0.332)	0.953	(0.404)
Yes	-0.049	(0.079)	0.421	(0.273)	0.470	(0.284)
DD	0.292	(0.243)	-0.191	(0.430)		
DDD					-0.483	
					(0.494)	
C. Private house						
No	-0.123	(0.082)	0.670*	(0.405)	0.793	(0.413)
Yes	-0.217	(0.178)	0.451*	(0.262)	0.668	(0.317)
DD	-0.094	(0.196)	-0.219	(0.482)		
DDD					-0.125	
					(0.520)	

Notes: Each cell uses the same specification as Equation (8). The regressions include province fixed effects, and the parentheses report bootstrapped robust standard errors clustered at province level with 50 repetitions.

Table D.2 presents the DDD estimates related to household demographics. Panel A breaks households into two groups based on whether there is a single-dependent child or multiple kids. Panel B leverages the variation from households with differing gender compositions. Panel C is based on the household head's age. For this exercise, we classify households into young households (age 36 and below) and old households (over age 43), based on the bottom and top quartiles of the age distribution.

Table D.2: Triple Difference-in-Difference using Demographics

	(1) 1995		(2) 2002		(3) DD	
	Coeff.	Std. Err	Coeff.	Std. Err	Mean	Std. Err
A. Single-dependent						
No	-0.160	(0.285)	0.708	(0.490)	0.868	(0.567)
Yes	-0.110	(0.084)	0.487*	(0.249)	0.75	(0.263)
DD	0.05	(0.297)	-0.221	(0.550)		
DDD					-0.118	
					(0.625)	
B. Gender						
Female	-0.143	(0.122)	0.538**	(0.215)	0.681	(0.247)
Male	-0.069	(0.102)	0.429	(0.336)	0.498	(0.351)
DD	0.074	(0.159)	-0.109	(0.399)		
DDD					-0.183	
					(0.430)	
C. Age of head						
Young	-0.250	(0.289)	0.432	(0.314)	0.682	(0.533)
Old	-0.188*	(0.097)	0.779***	(0.196)	0.967	(0.219)
DD	0.062	(0.305)	0.347	(0.370)		
DDD					0.285	
					(0.480)	

Notes: Each cell uses the same specification in Equation (8). The regressions include province fixed effects, and the parentheses report bootstrapped robust standard errors clustered at province level with 500 repetitions.

To summarize the DDD section, across the many specifications, the general patterns remain unchanged from our baseline regression results. The estimated change in college probability has only a small correlation with household saving rates within the 1995 sample, and this is true across many different sub-groups. In contrast, the correlation in the 2002 sample is very large for all sub-groups, indicating that the expansion of education opportunities increased household saving rates.

E Additional Figures

In addition, we find evidence that households without dependent children do not alter their saving behavior after the expansion. The age-saving profiles show that there is a parallel shift of saving rates for this type of households while households with children changed dramatically.

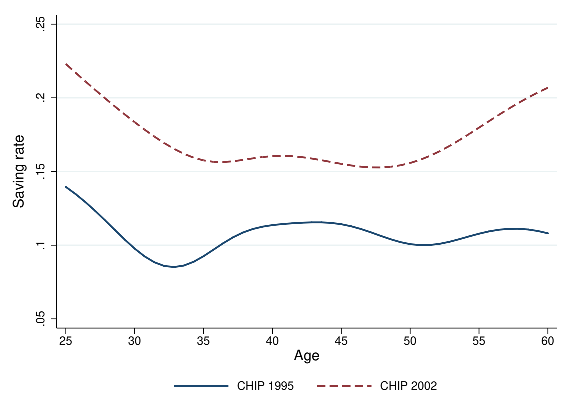

Figure E.1: Age-Saving Profiles for Households without Children

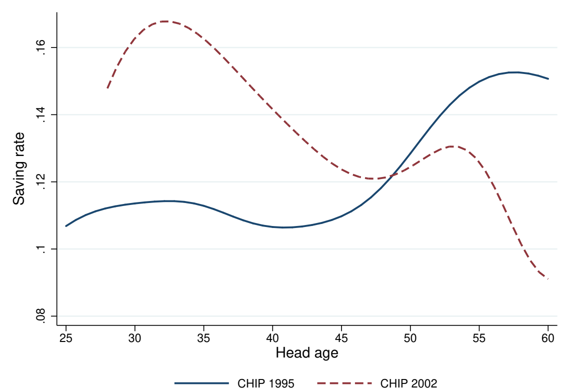

Figure E.2: Age-Saving Profiles for Households with Children