Does The Population Planning Policy In China Affects Fertility As The Policymakers Expected?

By

Yebaihe Lu

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Applied Economic
College of Business and Economics

At

The University of Wisconsin-Whitewater

July, 2019

Graduate Studies

The members of the Committee approve the thesis of

Yebaihe Lu presented on (July 31st, 2019)

Dr. Nicholas Lovett, Chair

Dr. Russell Kashian

Dr. Yuhan Xue

Chinese Population Planning, Fertility and Human Capital Investment

By

Yebaihe Lu

The University of Wisconsin-Whitewater, Year 2019
Under the Supervision of Dr.Yuhan(Cathy) Xue

China’s implementation of the One Child Policy created a setting where households were compelled to reduce family size. Traditional economic theory clearly predicts reductions in fertility will drive households to pursue greater human capital investments in their children. I use Chinese census data to test if the One Child Policy induced reductions in child bearing and subsequently changed households’ human capital investment choices. I leverage the historical time-varying, province level roll out of the One Child Policy in a generalized difference-in-differences framework to confirm the causal impact on child bearing. To investigate the fertility reduction’s effect on human capital investment, I use an instrumental variable identification strategy is used in conjunction with tax penalties associated with the policy. Results are precisely estimated and supportive that the reduction in child bearing lead to greater human capital investments.

I. Introduction

An extensive theoretical literature has clearly and strongly claimed that households face a quantity-quality trade-off when making child bearing and fertility choices. Under this model, first developed in Becker (1960) and further extended in Becker and Lewis (1973) and Becker and Tomes (1976), reductions in number of children will explicitly cause households to choose greater levels of human capital investment for their children. This purported relationship rests at the very core of Macro growth models which link a nation's economic condition to productivity levels that are determined by human capital investments and fertility. It is not an understatement say that these conclusions have influenced many important policies including and family planning across the world with examples including forced sterilization in India, and vigorous campaigns for reduced family sizes in Mexico and Indonesia.

While the theory is clear and generally accepted, empirical evidence for a causal relationship has been far weaker. Importantly, the existence of a negative correlation between family size and educational attainment has been well documented (see for example Hanushek (1992)), but causal evidence is extremely limited. In particular, many strong studies with good identification strategies and strong data such as Black, Devereaux, and Salvanes (2005), and Angrist, Lavy, and Schlosser (2010) have found no evidence for a causal relationship, while Kugler and Kumar (2017) find some evidence that a relationship may indeed exist in India.

Within a setting of mixed empirical support I seek to contribute by providing new estimates with a new setting and source of variation. Specifically, I use China's One Child Policy (OCP) as a source of variation in family size. First I show that the policy causally lowered fertility and then I use an instrumental variables approach using a new instrument.

The OCP is attractive for both empirical and practical reasons. One reason is that it created pressure on families to reduce their fertility in a way that was unrelated to each family's specific characteristics, and is therefore probably a clean source of variation. This allows me to consider effects beyond simple correlations. A second reason is that within China implementation of the OCP varied across time, space, and intensity. As a result there are easy to identify treatment and control groups. A last reason is that the OCP exposed households to varying tax penalties that influenced fertility, but probably didn't directly affect education investment. Combined this means the tax penalty is an attractive instrument for an instrumental variable framework.

My IV results indicate that there is evidence for a quantity-quality trade-off and that when family size increases, educational attainment falls. As long as the exclusion restriction is satisfied, these results are an improvement over prior results from observational studies and are different than results using different instruments in different countries. Additionally, the results are precisely estimated and come from a large data source representing nearly one fifth of the world's population. From a big picture standpoint, the results suggest that the assumption of a trade-off that has informed both macro growth models, and global family planning policies is justified.

The remainder of my thesis is organized as follows: Section II presents the necessary background information and discusses the prior literature, Section III discusses the data used and presents descriptive analyses, Section IV outlines the empirical strategies and discusses the necessary assumptions, Section V presents my results, and Section VI concludes.

II. Background and Prior Literature

A. Background

The motivation for the OCP was the concern that the rate of population growth since the end of WWII was too high and would lead to a shortage of resources, low growth rates and possibly famine. During the 1970s, the phrase "Family Planning" first appeared in official government announcements along with some slogan suggested couples to have fewer children. Such as "Fewer and better births, happiness throughout your whole life." However, there no legal rules related with it at 1970s, which were intended to be a slogan or form of propaganda or a goal with official government backing. By 1980, 'only one child per family' has been added to the nation's constitution and with enforcement process. Limitations restricted the number of children each couple could produce in their lifetimes, and also delayed the minimum marriage.

Just after the change in the constitution fertility rates fell by approximately 5.5% but there existed lots of variance in how families responded. Specifically, regional differences were large and there were differences between the majority Han ethnic group and minority groups. Policymakers wanted to increase the policy's effect and peoples' compliance, but also wanted to avoid an angry public. This was done by clarifying enforcement of the policy and allowing for different requirements across ethnic groups, districts and occupations in 1982. These exceptions resulted in different policies across different districts, as well as different degrees of compliance and efficacy.

Since the introduction of tax penalties, and until 2015, the policy has remained fairly stable. Following the Fifth Plenary Session of the 18th Communist Party of China Central Committee held in 2015, all families were permitted to have a second child, with the old policy still in place for families that have more than two children.

Although the OCP has ended Chinese culture has been deeply affected and many believe that the Chinese economy and society has been harmed in important ways. These harms include abnormal gender ratios, the aging of the population, and problems with government pension programs.

While the negative impacts of the OCP are obvious, the policy's goal of reducing population growth has been successful. China's economic growth during this time is also the Becker Quantity-Quality Model, and Endogenous Growth Models. China has increased human capital, productivity and welfare at the same time it has reduced family size. It is widely accepted by Chinese that these goals would not have been met if population growth remained the same since the 1970s.

B. Prior Research

Prior studies can be broadly considered be from two types. One type determined the impacts of the OCP on important outcomes. The other type seeks to empirically test the Becker Quantity-Quality Model's prediction that smaller families will invest more in their children.

A common area of research looks at the effect of the OCP on birth gender ratios. The birth gender ratio is the ratio of males to females at birth. Around the world this ratio is about 105, meaning 105 boys are born for every 100 girls. In China the ratio has been much higher, around 120. Before the OCP, data show a ratio of 106.

There are three explanations for the increase in the birth gender ratio. One is an increase in unreported female births, another is an increase in female infant mortality, and a third is more sex-selective abortion of female fetuses. Ebenstein (2010) presents evidence that the third channel is the most important. A common thread across these three factors is a patriarchal society and son preference in China.

Ebenstein (2010) uses Chinese Census data and uses differences in OCP fertility strictness across four province groups to test the effect of the OCP on fertility. A key result is that households are more likely to violate the OCP following the birth of a daughter. Ebenstein (2010) notes that a family with two daughters is nearly three times as likely to have a third child compared to a family with two sons. Additionally, Ebenstein (2010) shows the differences in rules across the four province groups, and over time.

Ebenstein (2010) also calculates that the OCP resulted in the female population being 5.8 million less than if there had been no OCP between 1990 and 2000. Ebenstein (2010) also shows that maternal age is different when the first child is a boy and when it is girl. He finds sons arrive 0.34 years later than daughters. This suggests that some female fetuses are aborted because they are female, and this delays the arrival of the next child, a son. On the other hand when the first fetus is a boy it is born and not aborted in most cases. Additionally, these delays for boys are larger in provinces with stricter versions of the OCP.

Greenhalgh (2013) also considers the OCP and gender ratios using demographic methods, but also seeks to consider what future effects will be. Specifically, Greenhalgh (2013) considers the marriage market and shows that between 2005 and 2025 there will be an extra 22 million marriage-age males because of the sex ratio at birth. This means that up to 10.4 percent of marriage-age men will be unable to marry. Concerns are particularly strong for the group of people born between 1980 and 2000 when the OCP restrictions were at their most strict. Greenhalgh (2013) notes that public always have a prejudice which the extra marriage-age males that cannot marry could become a threat to social stability. Greenhalgh(2013) states that this threat narrative is overwrought, but this popular bias itself is also one form of negative influence result by OCP.

Xu and Pak (2015) also consider gender imbalance and family planning in with respect to the OCP. Specifically, they test if gender imbalances have improved or not following the end of the OCP. The authors build a model of optimal family planning before and after the constraint is shifted from one to two children. However, the result is that the optimal strategy under the OCP is simply shifted to the second birth. This results in limited changes in gender ratios. This is particularly true if the true cause of the optimal strategy are traditional social attitudes that are unlikely to suddenly change in because of changes to the OCP.

In addition to gender ratios and population levels, it is likely that the OCP has had many other effects. Li et al. (2017) studied how the OCP influenced the prices of Shanghai housing in different school districts by using Hedonic regression models. They find the OCP led to a premium increase in prices for housing in high-quality school districts, which few times higher than the price increase in medium-quality and low-quality school districts. These results are suggestive that families may be seeking to make additional human capital investments in their children by buying housing in better school districts.

In addition to considering the overall effect of the OCP, McElroy and Yang (2000) consider both the financial rewards and penalties associated with the OCP. They use a two-stage least squares methodology for a sample of rural households, while also considering different tax penalties from the OCP over time. They find evidence that the penalties are lead to lower fertility in areas with higher penalties. They also find evidence that more educated women are more responsive to the penalties.

Prior studies looking for causal evidence of a quantity-quality trade-off usually use twin births, the desire for mixed gender child cohorts, son preferences, and ethnic-level differences in family size preferences as sources of plausibly exogenous variation in child bearing. These are then used within instrumental variables approaches. These studies have examined data from both developed nations such as Norway and Israel, as well as developing nations such as India.

Rosenzweig and Wolpin (1980) were the first to empirically test the quantity-quality model using twins as a natural experiment to compare families of different sizes. The idea probably works well because families that did not make different fertility choices randomly end up with different family sizes since some randomly have twins. Their estimates suggest there is a trade-off, but the estimates are imprecise and the sample size is very small. Hanushek (1992) tries to fix the small sample and imprecision problems by using large datasets in an observational setting to show the existence of a relationship between family size and child quality. However, in order to use large samples he has to use observational methods that can only reveal correlations.

Black, Devareaux and Salvanes (2005) test the quantity-quality trade-off empirically using data from Norway while using an instrumental variables approach. They used the occurrence of twins as an instrument like Rosenzweig and Wolpin (1980), but have much better data that includes everyone in Norway. Results are precise with a very large sample size and indicate birth order is important, but family size is not. This is viewed as a very important paper by economists.

Angrist, Lavy, and Schlosserman (2010) also use twins as an instrument, but also consider other instruments. Besides twins they consider families preferences for a mixed gender child cohort, and ethnic differences in son preference among Ashkenazi and Sephardic Jews in Israel. They combine results across these instruments for precision and conclude that there is no evidence for a quantity-quality trade-off. This means there conclusion is similar to Black et al. (2005) but the evidence is coming from multiple instruments and a different setting.

Kugler and Kuman (2017) use differences in son preference across India to instrument for family size and arrive at the conclusion that there is some evidence that there is a trade-off between quantity and quality. Hence, across studies, evidence is mixed at best, and otherwise not supportive of a trade-off.

III. Data and Descriptive Statistics

In this section I outline the key data sources, the variables, and the final data base. Additionally, I provide descriptive analysis that is both informative and motivates the primary analysis.

A. Data

The key data source used is the Fifth National Population Census of the People's Republic of China for the year 2000. The census records observations at both the family unit level and at the individual level. There are 1,180,111 individuals that participated in the census in total. Additionally a sample of over one million people completed a long version of the basic survey with extended details. The extended version not only includes information such as age, gender, education, and family size, but also contains questions about the person's province of birth, their current residence and their residence five years ago. Importantly, there is information regarding the number of sons and daughters the person has ever had, and the ages and present locations of those sons and daughters.

The data allows for linking between children and their mothers. This allows me to determine dates of birth and location of birth. This means that I can identify the form of the OCP in place for each birth. An exception is that 5% of children are linked to two or more mothers. These observations are dropped.

In addition to Chinese Census data, I also use data from Ebenstein (2010) that measures the time and province varying nature of the OCP. It also shows the tax penalties that each household would face if they had more children than the number of children permitted under the OCP. I merge this data set with the census data set, so every household will have variables which represent the specific strictness of the OCP and the fine they face if they have more children than permitted.

B. Descriptive Statistics

After linking all mothers and their children, fertility is calculated and plotted in Figure 1, which shows the average number of births across the timespan, 1950 to 2000, covered by the census. The figure matches a pattern of high fertility rates from 1950 until the late 1970s, where there is a modest decline, followed by more declining rates through the 1980s and 1990s. Since then, fertility rates have stabilized at a low level.

Interestingly, there appears to be evidence for changes in the slope of the relationship between year and fertility at key points in time that match with changes in the OCP. These changes appear in 1980 when the policy was strongly put in place, again in 1982 when the enforcement become less restricted, and again in 1990 when tax penalties were put into place. Additionally the plots in Figure 3 clearly indicate that the pattern of fertility reductions by province group and time match the implementation and enforcement of the OCP across province groups and time.

While the plots are very informative, they are not sufficient to be taken as causal evidence of the policy's effect. This is because there is the potential for other time and province varying changes that could correlate with the implementation of the OCP. To provide stronger evidence, difference-in-differences estimates and IV estimates are provided in Section 5.

Given there isn't a causal link at this point, it may be useful to consider which variables might correlate with the total children each mother will have. Table 1 includes summary statistics for a minority indicator variable, a dummy variable if the individual lives in an urban area relative to a rural area, housing area, birth year, employment status, and recorded children's gender. Table 2 shows summary statistics for variables that are used in the regression analysis. The correlations are as expected with minority women, rural women, and low education women having more children.

IV. Methodology

A. Observational Methods

I first estimate a linear model via OLS with an observational framework to see which factors are significantly correlated with the total children per family. The model is:

$$\begin{aligned}\text{TotalKids} = {} & \beta_0 + \beta_1\text{ProvinceOneHalf} + \beta_2\text{ProvinceTwo} + \beta_3\text{ProvinceThr} + \beta_4\text{MomEdu} + \\ & \beta_5\text{HousingArea} + \beta_6\text{HousingArea}^2 + \beta_7\text{MomAge} + \beta_8\log(\text{Fine}+1) + \varepsilon\end{aligned}$$

The dependent variable is the total number of children a woman has had before November 2000. The most important variables are the three province group variables which measure the differing intensities of the OCP across province groups. I divide observations into four groups according to current address, minority, and the requirements of the OCP.

All residents who live in autonomous urban areas and two specific provinces (Jiangsu and Sichuan) were only permitted to have one child. This group is assigned a value of 1 for the indicator variable ProvinceOne, and is a reference group. Since it is a reference group it does not appear in the regression specification. Although there are some exceptions to the single child requirement, they are extremely infrequent and these rare observations will be captured by the error term.

The next group of provinces were subjected to a policy where parents can only have one child if the first birth is a boy, but may have another child if the first birth is a girl. This policy applied to households in several provinces and those that live in designated rural areas. These households are captured by the variable ProvinceOneHalf.

Similar to the two groups above, the third group, consists of families allowed to have two children who live in 6 provinces. These households are captured by the variable ProvinceTwo. The last group, captured by the variable ProvinceThr, pertains to non-Han, minority households living in 5 provinces, as well as a small share of families who live in very remote areas.

Additional variables are control variables that account for other important considerations. These include controls for housing space, the log of tax penalties, maternal education, a squared housing space measure, and maternal age at first birth.

B. Difference-in-Differences

Since observational studies are unlikely to report correct estimates because of omitted variables bias and reverse causality, I use a difference-in-differences approach. This has the potential to give more credible estimates. Specifically, I estimate two difference-in-difference models. The first is given by:

$$\text{TotalKids} = \beta_0 + \beta_1 Treatgroup + \beta_2 SecondPeriod + \beta_3 Treatpost + \varepsilon$$

where the *Treatgroup* variable is for each of the province groups discussed above. For each specification the control group is the province three group. Group one experiences the strictest limitations and it is expected that the difference with the control group will be significant. The second period variable is a post indicator that accounts for whether a child is born following the OCP restrictions of December, 1982.

The second set of difference in differences models are qualitatively similar but include control variables in order to assist with precision. These specifications take the form:

$$\text{TotalKids} = \beta_0 + \beta_1 Treatgroup + \beta_2 SecondPeriod + \beta_3 Treatpost + \beta_4 Edu + \beta_5 Age + \beta_6 HousingArea$$

$$+ \beta_7 Single + \beta_8 SecondMarriage + \beta_9 Divorced + \beta_{10} Widow + \beta_{11} PrivateBathroom + \varepsilon$$

In this regression, I use first marriage status as the reference group with indicators for other marital statuses. This seems reasonable as first marriage is the most common.

The key assumption in a difference-in-differences is the parallel trend assumption. If the parallel trend assumption is violated, then my estimates will not be an improvement over the observational approach. The parallel trend assumption says that in the absence of the treatment, the treatment group and the control group would have continued on a similar path or trend in the post period. Since, the treatment group gets the treatment, it is not possible to observe the case where the treatment group does not get the treatment and then compare the trend to the control group's trend. For this reason, it is never possible to fully answer if the parallel trend assumption is satisfied.

As a result, researchers check the trends in the pre-period for suggestive evidence of what would have happened in the post-period. Figure 3 allows me to visually compare the fertility rate trends prior to implementation of the OCP. Importantly, all four groups look very similar in terms of slopes in the pre-period. This provides strong evidence the four province groups are on the same trend. Since they are on the same trend, this provides strong evidence that the parallel trend assumption is met, given that it is not possible to see the treatment group trend in the post period under the case where the treatment group is not given the treatment.

C. Instrumental Variables

The last methodology I use is an instrumental variable. Although instrumental variables estimation can give imprecise results, there are two strong reasons to use an IV. One is that I may only care about the sign of the effect and accept imprecision. The second, and more important, reason is that the instrumental variable uses variation in child bearing that only comes from the size of the tax penalty. Since the tax penalty size is independent of families' human capital investment choices, this means I can estimate the effect of family size in an unbiased way. This is not possible with the observational approach.

The first stage regresses the fine on the number of children and is given by:

$$TotalNumberofKids = \beta_0 + \beta_1 TaxPenalty + \beta_2 Work + \beta_3 HousingArea + \beta_4 Single + \beta_5 SecondMarriage + \beta_6 Divorce + \beta_7 Widow + \beta_8 Bathroom + \sigma$$

The second stage then takes the number of children given by the first stage, and puts them on the right hand side of the second-stage regression equation. This non-endogenous variation in number of children is then used to estimate the true effect of family size on human capital investments.

$$EducYear = \beta_0 + \beta_1 \widehat{TotalKids} + \beta_2 Work + \beta_3 HousingArea + \beta_4 Single + \beta_5 SecondMarriage + \beta_6 Divorce + \beta_7 Widow + \beta_8 Bathroom + \varepsilon$$

In order for the instrument to give me an unbiased estimate two requirements need to be met. One is that the instrument must be predictive of the number of children in a household. Intuitively, this should be easily accepted. Economic theory is very clear that an increase in the price of a good should result in a reduction of the quantity of that good. In this case, as the tax penalty increases it is expected that child bearing will fall. Additionally, both the DID results and the observational study results are supportive of this. Another way of to evaluate the strength of the instrument is to consider the F-statistic of instrument exclusion in the first-stage regression. Stock and Yogo (2005) show that this F-statistic should be above 10. As shown in Table 5 the first stage f-statistic of the excluded instrument is always above 10. As a result, there is intuitive and empirical evidence that the instrument is not a weak instrument.

The second IV assumption is that the instrument is valid, which means the exclusion restriction is true. The exclusion restriction says that the tax penalty must not change human capital investments, *except* through the channel of changes in fertility. That means I can exclude the possibility that the tax penalty has a direct effect on human capital investments. This assumption is far more difficult to prove than the validity assumption because there are no empirical tests that can prove the instrument is valid.

Because I cannot test the validity assumption, instead I consider an intuitive approach. Specifically, the tax penalty is determined at the province level, but human capital investments are made at the household level. This means it is unlikely that the two decisions are related, and they are probably exogenous to each other. Although taxes will affect fertility, which in turn likely affects human capital investments, this is explicitly permitted because the exclusion restriction allows for the instrument to affect the outcome variable through changes in fertility. As a result, it seems intuitive that the validity assumption holds.

V. Results

A. Observational Approach

Table 3 reports the results of the observational regression. Every coefficient is significant. Holding other independent variables constant, we would expect a female to have 0.74 fewer children than others if she has ten more years of education. Housing area appears to have a non-linear effect. For small housing, an increase in the area has a positive effect on number of children. On the other hand, if housing space is above 18 square meters, the effect of an increase in housing area is negative, and the size of the effect grows as housing area increases.

The coefficient on tax penalties or fines imply that if the fine is 1% higher, the number of children would be expected to increase by 0.147 holding all other variables constant. This positive correlation provides evidence that higher fines were implemented where families were having larger size as a response.

Both maternal education years and housing area have coefficient sizes that suggest their influence on the dependent variable is large. The signs of all coefficients, positive or negative, are also consistent with my assumptions and expectations. The results would also suggest that an alternative government policy would be to require more education, which would also reduce the fertility rate, but maybe with less side effects.

B. DID Approach

The difference-in-differences approach is also supports the idea that the OCP has been effective in changing family size. Table 4 shows that the changes in fertility rates across province groups is closely matched by the strength of the family-planning requirements in each province group. It would appear to show that the enforcement of the policy is effective.

The treat*post estimate is the interaction term and the coefficient of the difference-in-differences estimator. Between the strongest and weakest requirements in province groups, the regression results show that on average, we would expect females will have 0.815 fewer children in the strictest province than females who give birth at the same time, but live in another weaker province group.

The primary advantage of the difference-in-differences estimator is that it gives me a better estimate than from the observational approach because I am less worried about omitted variable bias. Instead, the parallel trends assumption is the key and there is evidence this assumption is met. Finally, both the observational approach and the differences-in-differences approach indicate that the OCP did reduce family size.

With this fact in hand, I now turn to the IV results to see if reductions in family size result in increased human capital investments.

C. IV Approach

Table 5 reports the results of estimating the second stage specification shown in the Methodology section. Results for specifications both with and without controls are shown. OLS estimates are also shown so that it is easy to compare OLS and IV results. The central result is that the coefficient of family size is precisely estimated and shows that as family size grows investments in human capital decline. Importantly, this shows support for the quantity-quality trade-off as developed in the theory of Becker (1960), Becker and Lewis (1973), and Becker and Tomes (1976). This is true both with and without controls, and is significant at the 99% level for both. Finally, the effect size is similar both with and without controls.

Relative to the estimates given by OLS in the observational studies, the magnitudes of effects under IV estimation are larger. This suggests that the bias present in observational studies is likely to diminishing the magnitude of the estimate.

VI. Conclusion

While economic theory has strongly claimed that there is a quantity-quality trade-off in fertility, empirical evidence has been comparatively weak. A key reason for this is that observational studies can show a correlation between larger families and reduced human capital investment, but because of issues such as omitted variable bias, observational studies cannot offer strong support or reject the quantity-quality model.

I address this by instead using an instrumental variables approach with Chinese census data to offer empirical support for the model. Critically, I use China's OCP as a way of avoiding the endogeneity issues that are a problem with observational studies. China's OCP is very valuable for research on this question, because it forced households to reduce their family sizes in a way that was not related to their characteristics and preferences. As a result, the reduction in family size can be viewed as almost random. Specifically, I use differing versions of the policy that were implemented in different provinces and at different points in time to estimate the effect of the policy on child bearing and human capital investments.

First, I confirm that the OCP policy did reduce child bearing through both an observational approach and a difference-in-differences approach. Then I use the differences in tax penalties across time and place to instrument for family size in an IV regression framework to estimate the effect of family size on human capital investments. Results indicate that larger families do lead to per child reductions in education.

References

Angrist, J., Lavy, V. and Schlosser, A., 2010. Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, *28*(4), pp.773-824.

Becker, G.S., 1960. An Economic Analysis of Fertility. In *Demographic and Economic Change in Developed Countries* (pp. 209-240). Columbia University Press.

Becker, G.S. and Lewis, H.G., 1973. On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, *81*(2, Part 2), pp.S279-S288.

Becker, G.S. and Tomes, N., 1976. Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, *84*(4, Part 2), pp.S143-S162.

Black, S.E., Devereux, P.J. and Salvanes, K.G., 2005. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics*, *120*(2), pp.669-700.

China Daily. 2012. "Gender Imbalance Set to Ease," «npfpc.gov.cn/news/central», 31 March. Accessed 5/23/18. Retrieved from http://usa.chinadaily.com.cn/epaper/2012-03/30/content_14952480.htm

Chinese Department of Statistics, Bureau of Statistics of the People's Republic of China. (2010) Fifth National Population Census of the People's Republic of China. National Bureau of Statistics of the People's Republic of China.

Central Government of the People's Republic of China. (1984) *Report on Family Planning. No. 7*. Publication: Shanghai Municipal Health and Family Planning Commission.

Ebenstein, A., 2010. The "Missing Girls" of China and the Unintended Consequences of the One Child Policy. *Journal of Human Resources*, *45*(1), pp.87-115.

Greenhalgh, S., 2013. Patriarchal Demographics? China's Sex Ratio Reconsidered. *Population and Development Review*, *38*, pp.130-149.

Hanushek, E.A., 1992. The Trade-off between Child Quantity and Quality. *Journal of Political Economy*, *100*(1), pp.84-117.

Junhong, C., 2001. Prenatal Sex Determination and Sex-selective Abortion in Rural Central China. *Population and Development Review*, *27*(2), pp.259-281.

Kugler, A.D. and Kumar, S., 2017. Preference for Boys, Family Size, and Educational Attainment in India. *Demography*, *54*(3), pp.835-859.

Li, X.S., Chen, Y.M., Fang, F., and Zhang, Z., 2017. Two Child Policy and the Premium of School District Housing: Policy Evaluation Analysis Based on the Change of Population Policy. *Journal of Economics and Finance*, *43*(6), pp.93-104.

McElroy, M. and Yang, D.T., 2000. Carrots and Sticks: Fertility Effects of China's Population Policies. *American Economic Review*, *90*(2), pp.389-392.

National Health Commission of the People's Republic of China, (2015). The Report of the Fifth Plenary Session of the 18th CCP Central Committee. Publication: National Health Commission of the People's Republic of China.

Rosenzweig, M.R. and Wolpin, K.I., 1980. Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment. *Econometrica*, pp.227-240.

Stock, J.H. and Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. *Identification and Inference for Econometric Models*, p.80.

Xu, B. and Pak, M., 2015. Gender Ratio Under China's Two-Child Policy. *Journal of Economic Behavior & Organization*, *119*, pp.289-307.

Zhou, X.D., Wang, X.L., Li, L., and Hesketh, T., 2011. The Very High Sex Ratio in Rural China: Impact on the Psychosocial Wellbeing of Unmarried Men. *Social Science & Medicine*, *73*(9), pp.1422-1427.

Figure 1. Average fertility rate in China, 1960-2015

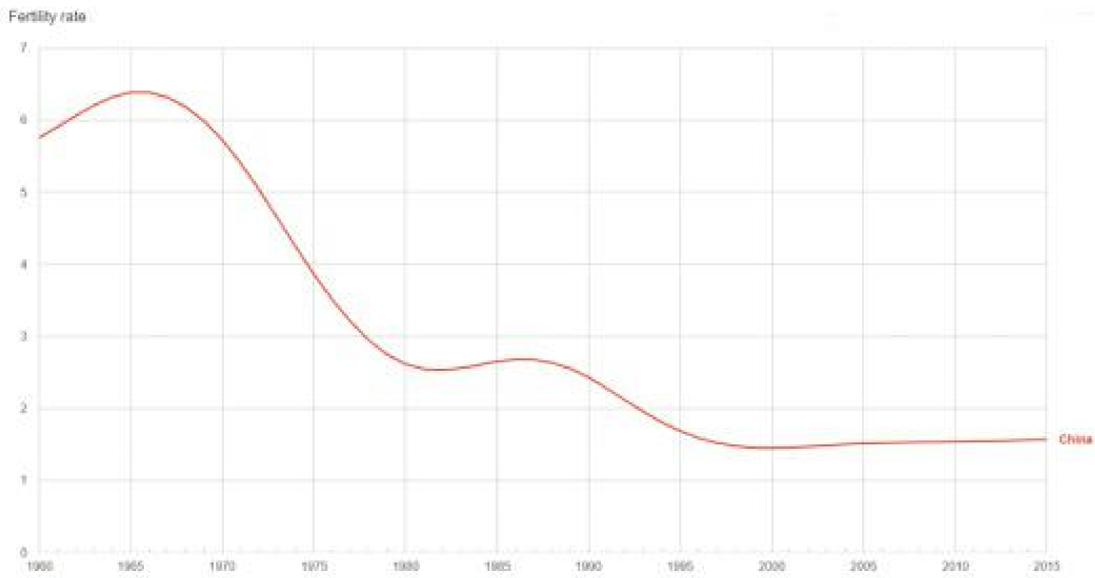


Figure 2. Fertility 1950-2000. (Vert. lines at 1980, 1982, and 1990)

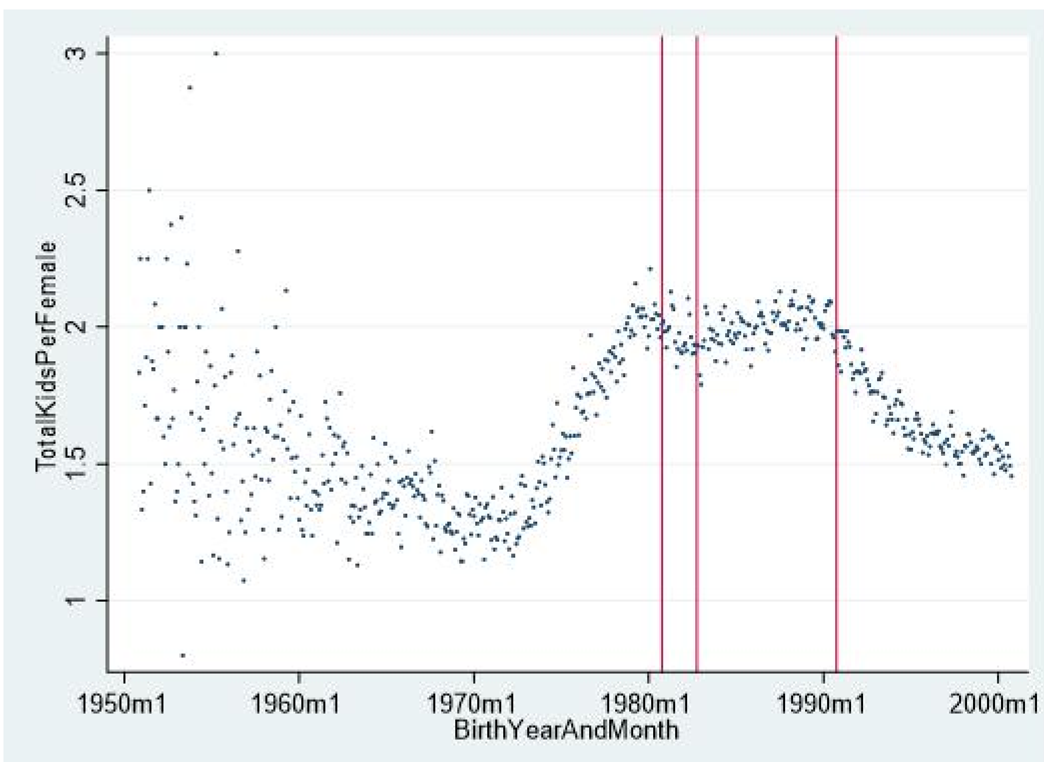
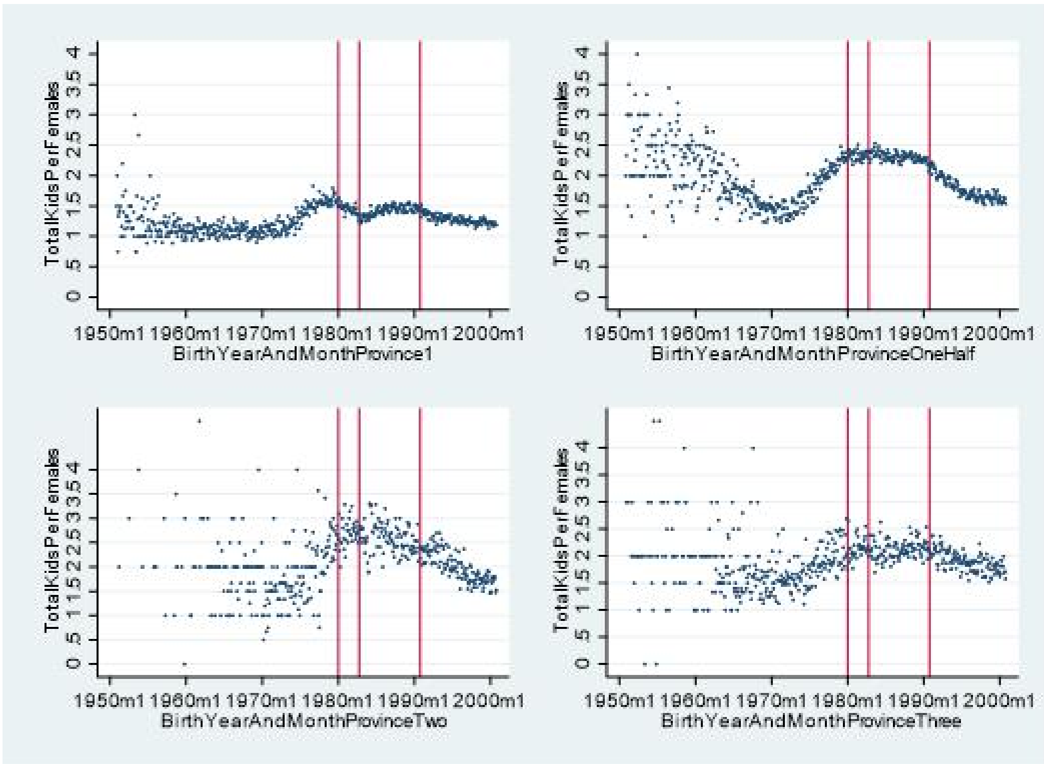


Figure 3. Fertility Rates by Province Group (Vert. lines at 1980, 1982, and 1990)



Notes: I group provinces into four province group categories by family planning policy requirements determined in 1982. Province1 means citizens can only have one child. ProvinceOneHalf means citizens could have a second child if the first birth is a girl. ProvinceTwo means citizens can have two kids and Province Three means citizens can have three or more kids.

Table 1. Summary Statistics for Fifth National Population Census.

Variable	Observation	Mean	Standard	Minimum	Maximum
Minority	246,084	0.9182	0.2741	0	1
Household Register Type	246,084	0.2572	0.4371	0	1
Housing Area Per Family	246,084	85.7683	55.7572	0	999
Mom Age	246,084	36.7621	7.7809	16.25	93.25
Work Status	246,084	0.8281	0.3773	0	1
Kids gender	384,354	0.5738	0.4945	0	1

Table 2. Descriptive Statistics for Observational Study Variables.

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
Total Children	179,820	1.8104	0.9701	0	12
Prov. Group One-Half	179,820	0.5418	0.4983	0	1
Prov. Group Two	179,820	0.03122	0.1739	0	1
Province Group Three	179,820	0.08365	0.2769	0	1
Maternal Education	179,820	7.7860	3.1535	0	18
Housing Area	179,820	84.1553	53.0611	0	999
Maternal Age	179,820	36.0967	6.7672	16	86.1667
Log(Fine + 1)	179,820	0.8692	0.3287	0.09531	1.7918

Table 3. Observational Regression Results

Variables	Total Children
Maternal Education	-0.0526***
	(0.000670)
Housing Area	0.00131***
	(0.0000645)
Housing Area Squared	-0.00000163***
	(0.000000169)
Province Group One-Half	0.561***
	(0.00445)
Province Group Two	0.713***
	(0.0108)
Provinces Group Three	0.180***
	(0.00740)
Maternal Age at First Birth	0.0537***
	(0.000307)
Log(Fine + 1)	0.155***
	(0.00591)
Constant	-0.229***
	(0.0166)
Observations	162,632
R-Squared	0.331
F-test	10056

Notes: In all models, *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 4. Difference-in-Differences Results

Dependent Variable	Total Children	
	No Controls	With Controls
Treatment	-0.692***	-0.551***
	0.00394	(0.00371)
Post Period	-1.540***	-0.609***
	(0.0298)	(0.0260)
Treat*Post	0.630***	0.0607***
	(0.0693)	(0.0589)
Maternal Education		-0.0542***
		(0.00570)
Work Status		0.161***
		(0.00436)
Maternal Age at First Birth		0.0543***
		(0.000214)
Housing Area		0.000792***
		(0.00003)
Single		-0.547***
		(0.0587)
Second Marriage		0.0399***
		(0.0105)
Divorced		-0.324***
		(0.0168)
Widow		-0.173***
		(0.0146)
Private Bathroom		-0.136***
		(0.00361)
Constant	1.984***	0.266***
	(0.00244)	(0.0106)
Observations	253,782	253,782
R-Squared	0.1152	0.3595
F-test	11013.67	11871.42

Notes: This table presents estimates using a difference-in-differences methodology. Treat group refers to the provinces that only allow one child. The first or pre period is 1979-1982, and the second or post period is from 1982 onward. Treat*Post represent the difference-in-differences estimator. The one time married group is the reference group for all marriage statuses. In all models, *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5. Instrumental Variable Results

	OLS		IV	
	With Controls	Without Controls	With Controls	Without Controls
Constant	9.884***	10.098***	10.121***	10.264***
	(0.0252)	(0.0155)	(0.0645)	(0.0662)
Predicted Total Kids	-1.162***	-1.197***	-1.316***	-1.291***
	(0.00796)	(0.00782)	(0.0395)	(0.0374)
Work Status	-0.189***		-0.159***	
	(0.0195)		(0.0209)	
Housing Area	-0.00119***		-0.000862***	
	(0.000140)		(0.000162)	
Single	-2.135		-2.394**	
	(1.168)		(1.171)	
Second Marriage	-0.501***		-0.467	
	(0.0481)		(0.0488)	
Divorce	1.416***		1.347***	
	(0.100)		(0.102)	
Widow	-0.255***		-0.208**	
	(0.0835)		(0.0844)	
Private Bath	0.590***		0.563***	
	(0.0159)		(0.0172)	
R-Squared	0.1363	0.1260	0.1343	0.1252
Observations	160,539	162,668	160,539	162,668
First Stage F-stat	3166.10	23444.80	638.57	1191.22

Notes: In all models, *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.