

NBER WORKING PAPER SERIES

QUANTITY-QUALITY AND THE ONE CHILD POLICY:  
THE ONLY-CHILD DISADVANTAGE IN SCHOOL ENROLLMENT IN RURAL CHINA

Nancy Qian

Working Paper 14973  
<http://www.nber.org/papers/w14973>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
May 2009

I thank Joshua Angrist, Daron Acemoglu, Abhijit Banerjee, Gary Becker, Esther Duflo, Oded Galor, Andrew Foster, Ashley Lester, Mark Rosenzweig and David Weil for their many insights; and participants at the MIT development lunch, BREAD and the Econometrics Society for useful comments. I am incredibly grateful to the China Health and Nutritional Survey at the University of North Carolina and the National Bureau of Statistics of China for sharing their data with me. All mistakes are my own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2009 by Nancy Qian. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrollment  
in Rural China

Nancy Qian

NBER Working Paper No. 14973

May 2009

JEL No. I20, J13, O1

**ABSTRACT**

Many believe that increasing the quantity of children will lead to a decrease in their quality. This paper exploits plausibly exogenous changes in family size caused by relaxations in China's One Child Policy to estimate the causal effect of family size on school enrollment of the first child. The results show that for one-child families, an additional child significantly increased school enrollment of first-born children by approximately 16 percentage-points. The effect is larger for households where the children are of the same sex.

Nancy Qian

Department of Economics

Brown University

64 Waterman Street

Providence, RI 02912

and NBER

nqian@brown.edu

# Quantity-Quality and the One Child Policy: The Only-Child Disadvantage in School Enrollment in Rural China

May 20, 2009

## Abstract

Many believe that increasing the quantity of children will lead to a decrease in their quality. This paper exploits plausibly exogenous changes in family size caused by relaxations in China's One Child Policy to estimate the causal effect of family size on school enrollment of the first child. The results show that for one-child families, an additional child *increased* school enrollment of first-born children by up to sixteen percentage-points. The benefit is larger for households where the children are of the same sex. (Keywords: Education, Development, Family Planning; JEL: J13, I2, O1)

## 1 Introduction

Since Malthus first theorized about population constraints on economic growth, the trade-off between quantity and quality has been a central question in economics. Contemporary economists, beginning with Becker and Lewis (1973) and Becker and Tomes (1976) have continued to theorize about the negative effects of family size on child quality (Galor and Weil, 2000; Hazan and Berdugo, 2002; Moav, 2005). For policy makers in developing countries today, understanding this trade-off is especially relevant as many governments have attempted to curb population growth as a way of increasing average human capital investment. Both China and India, the world's two most populous countries, have experimented with different family planning policies to limit family size.

The empirical evidence on the extent of the quantity-quality tradeoff or even the existence of it is conflicted. On the one hand, studies have found family size to have no effect or even a positive effect on child outcomes in Israel (Angrist, Lavy and Schlosser, 2005, 2006), Korea (Lee, 2009), the U.S. (Kessler, 1993), China (Guo and VanWey, 1991) and Africa (Gomes, 1984). On the other hand, the effect of family size on education has been found to be negative in the India (Rosenzweig and Wolpin, 1980), France (Goux and Maurin, 2004), the U.S. (Conley, 2004; Berhman et. al., 1989; and Stafford, 1987), and China (Rosenzweig and Zhang, 2009).<sup>1</sup> Theoretically, it is possible to motivate both models that predict negative family size effects and models that predict positive ones without hard-to-believe assumptions. The quantity-quality tradeoff can arise from the assumption that the cost of quality is increasing with the number of children. While this may be a reasonable assumption in many contexts, it seems equally reasonable to think that there exist contexts where the opposite is true: that there are economies of scale in raising children. For example, children may share clothes and text books which are easier to transfer between children of the same family than across children of different households. In this case, quantity can compliment quality.

Moreover, an important gap exists in the literature. Past studies have almost entirely focused on the effect of increasing the number of children from two to three. We know very little about the effects of increasing the number of children from one to two, the margin that is increasing in relevance as parents' preferred household size decreases over time. This margin is particularly important to examine since there are reasons to believe that the effect of family size on quality is not monotonic across the number of children. Most past studies which typically rely on continuous linear measures for family size have missed this point. Only recently, by using dummy variables for family size, have studies shown correlational evidence that children with no siblings do worse than children with one or two siblings even though the number of siblings is negatively correlated with child quality for households with three or more children (Black, Devereaux and Salanes, 2004; and Iacovou, 2001).<sup>2</sup>

Difficulty in establishing causality also hindered past studies from studying the effect of being an only child. Interpreting the cross-sectional correlation between family size and child outcomes faces two main difficulties. The source of endogeneity that concerned most previous studies is parental heterogeneity. For example, if parents who value education more may also prefer to have fewer children, then the correlation between quantity and quality will be

---

<sup>1</sup>The studies described here all focus on cross sectional evidence. Alternatively, Bleakely and Lange (2009) examines time-series evidence in the American South. They find that increased schooling caused a decrease in fertility. See Schultz (2005) for a detailed critique of the empirical literature.

<sup>2</sup>Iacovou (2001) argues that this could be because children benefit from social interactions with other children. Using detailed data on time use of children in the U.K., she finds that the one-child disadvantage decreases with the amount of time a child spends playing with other children after school.

driven by parental preferences rather than by family size. Endogeneity may also arise from the quality of the first child. For example, if parents are more likely to have a second child when the first child is of high quality, the ordinary least squares (OLS) estimate of the family size effect will be biased upwards. To address these issues, past studies have carefully constructed strategies that exploit plausibly exogenous variation in family size caused by multiple births or the sex composition of the first two children (Angrist, Lavy and Schlosser, 2005, 2006; Conley, 2004; Lee, 2003; Rosenzweig and Wolpin, 1980; Rosenzweig and Zhang, 2009).<sup>3</sup> However, these strategies work best in studying the effect of increasing the number of children from two to three. The sibling-sex strategy only applies to households that already have two children. The twins strategy can only estimate the effect of having two children instead of one under the assumption that there is nothing inherently different about twins.

The principal contribution of this paper is to establish causal identification for estimating the effect of increasing the number of children from one to two. First, I present a simple model that allows for economies of scale and illustrate how having a second child could increase school enrollment of the first. Then, I address the empirical issue of endogeneity by exploiting regional and time variation in relaxations of China's One Child Policy (OCP). I use the relaxation that allowed families to have a second child if the first child is a girl to instrument for the family size of first-born children. Three facts are exploited: first, an individual is only affected by the relaxation if she is born in a relaxed area; second, amongst first-born children born in relaxed areas, only girls are affected; and third, a girl is more likely to gain a sibling due to the relaxation if she is younger at the time of the policy announcement. The instrument for family size is the triple interaction term of an individual's sex, date of birth and region of birth. The interaction between whether a girl was born in a relaxed area and her birth year estimates the effect of the relaxation on family size. The additional comparison with boys controls for region-specific changes in school provision (and other cohort changes) that affected boys and girls similarly. This strategy differs from previous methods in that it essentially compares one-child households with two-child households. To exclude parents who may have chosen to keep girls in order to have a second child, the sample is restricted to first born children born before the relaxation began.

The first stage results show that the OCP was binding for children born four or less years before the OCP was announced. The relaxation increased the number of siblings for these

---

<sup>3</sup>The sibling sex composition method argues that parents prefer children of mixed sex. Therefore, they are more likely to have a third child if the first two are of the same sex. The twins method argues that the occurrence of twins (before the introduction of fertility treatments) is uncorrelated to individual characteristics. Hence, twinning is a plausibly exogenous source of variation for family size. Both methodologies examine the effect of an additional sibling for families with at least two children. Angrist, Lavy and Schlosser (2005, 2006) used both techniques and found that the results are similar.

children. On average, approximately one out of every four children gained a sibling because of the relaxation. The second stage results show that an additional child significantly increased the school enrollment of the first child by up to sixteen percentage points. The positive sign of the effect of a second child is consistent with the OLS evidence but the 2SLS estimates are larger in magnitude. To investigate whether this could be due to economies of scale in schooling costs, I estimate the differential effect for when the first two children are of the same sex versus when they are of different sex. If one assumes that economies are greater when children are of the same sex, then the presence of economies of scale would predict that the benefits are larger for such households. The results show that the benefits are larger when the two children are of the same sex. However, since sex can be endogenously chosen by parents, this should be interpreted cautiously as suggestive evidence.

The results from this study provide novel casual evidence for the benefit of an additional sibling for first born children. Furthermore, they suggest that there may be economies of scale in the costs of schooling. The main caveat for interpretation comes from the fact that the second child is still too young to attend school at the time the data is collected so that I cannot examine their outcomes. Therefore, the results cannot directly “test” the standard quantity-quality model which predicts that average quality is decreasing with quantity. That said, most existing empirical studies of the effects of family size focus on outcomes of the first born children. Hence, the results from this study are easily comparable to the existing evidence.

This paper has several advantages over past studies. First, China’s family planning policies provide a unique source of exogenous variation in family size. Second, the empirical strategy can evaluate the causal effect of an additional sibling for one-child households whereas past methodologies have mostly been limited to households with two or more children. Finally, this study evaluates the effects of the OCP, one of the most restrictive and large scale family planning policies ever undertaken. While demographers and sociologists have conducted many descriptive studies of the policy’s impact on fertility and sex ratios, the lack of local enforcement data has heretofore prevented an examination of the causal effect of the OCP on child outcomes. The findings indicate that the OCP decreased female survival by up to ten percentage-points, and the relaxation was successful in reducing the sex selection to pre-OCP levels. Interestingly, the results also show that the previous rule on four-year birth spacing was well enforced.

The paper is organized as follows. Section 2 discusses family planning policies and education in rural China. Section 3 presents the conceptual framework. Section 4 describes the data. Section 5 presents the empirical strategy. Section 6 presents the empirical results. Section 7 offers concluding remarks.

## 2 Background

### 2.1 Family Planning Policies

In the 1970s, after two decades of explicitly encouraging population growth, policy makers in China enacted a series of measures to curb population growth. The policies applied to individuals of Han ethnicity, who comprise 92% of China’s population. Beginning around 1972, the policy “Later [age], longer [the spacing of births], fewer [number of children]” offered economic incentives to parents who spaced the birth of their children at least four years apart. A more restrictive policy, commonly known as the One Child Policy (OCP), was announced in 1979. Actual implementation began in certain regions as early as 1978, and enforcement gradually tightened across the country until it was firmly in place in 1980 (Croll et. al., 1985; Banister, 1987). Second births became forbidden except under extenuating circumstances. Local cadres were given economic incentives to suppress fertility rates. In the early 1980s, parts of the country were swept by campaigns of forced abortion and sterilization and reports of female infanticide became widespread (Greenlaugh, 1986; Banister, 1987). Past studies have generally considered the OCP to have only affected the family size of cohorts born after 1979/80. However, if the previous four year birth spacing law was well enforced, then the OCP should be binding for those born in 1976 and after.

Local governments began issuing permits for a second child as early as 1982. This practice was made official when the Central Party Committee issued “Document 7” on April 13, 1984. The two main purposes of the document were to: 1) curb female infanticide, forced abortion and forced sterilization; and 2) devolve responsibility from the central government to the local and provincial government so that local conditions could be better addressed. In other words, it allowed for regional variation in family planning policies. The document allowed for second births for rural couples with “practical” difficulties, and strictly prohibited coercive methods. The main relaxation following Document 7 is called the “1-son-2-child” rule. It allowed rural couples to have a second child if the first child was a girl (Greenlaugh, 1986).<sup>4</sup> The explicit purpose of this relaxation was to decrease female infanticide of the first-born child. White (1992) found that 5% of rural households were allotted second child permits in 1982. These permits were generally granted to regions with extremely high levels of infanticide. After Document 7, the permits expanded to 10% of the rural population in 1984, 20% in 1985 and 50% by 1986.

Document 7 made provincial governments responsible for both maintaining low fertility rates and decreasing infanticide. While the exact process of granting permits is unclear, I

---

<sup>4</sup>Practical difficulties included households where a parent or first-born child was handicapped, or if a parent was engaged in a dangerous industry (e.g. mining).

use county level data on family planning policy to show in the next section that the probability for a county to obtain the 1-son-2-child relaxation is positively correlated with the rate of pre-relaxation sex-selection, and both are positively correlated with distance from the provincial capital. These facts most likely reflect that in order to maintain low aggregate fertility rates and decrease excess female mortality (EFM), provincial governments granted relaxations to regions that were distant to the administrative capital, and where EFM was more prevalent. The higher prevalence of sex-selection in rural areas can be due to both more boy-preference in distant rural areas and the fact that geographic distance increases the provincial government’s logistical difficulty of preventing EFM.<sup>5</sup> Issues of identification that arise from the correlation of obtaining a relaxation and sex-selection will be addressed explicitly in the section on robustness.

Note that ethnic minorities were always exempt from family planning policies. This will not have a large effect on this study since only a very small proportion of the Chinese population is comprised of ethnic minorities. To be cautious, the regression analysis will control for ethnicity.

## 2.2 Rural Education

Inequality in education provision greatly increased during the 1980s both across provinces and across counties within a province. Inequality between school finance increased as changes in the fiscal system reduced subsidies from rich regions to poor regions. The system of “eating from separate pots” (*fen zou chi fang*) devolved expenditure responsibilities from the central and provincial governments onto local governments in order to give the latter stronger incentives to generate revenue. The ratio of the per capita schooling expenditure in the highest spending province to the lowest spending province doubled in one decade.

Many rural schools were closed; rural enrollment rates dropped dramatically and did not recover until the mid to late 1990s (Hannum and Park, 2007). Using spending data from Gansu province, they found that per capita school expenditure was positively correlated with income and that significant variation in school quality existed across counties. They found little variation within counties, suggesting that studies examining education outcomes should focus on variation at the county level.

Hannum (1992) showed that the difference in school provision between rich and poor areas was much greater for middle schools and high schools than for primary schools. This is consistent with the CHNS data used in this study, where primary school enrollment remained

---

<sup>5</sup>Levels of income between counties with some relaxation and counties with no relaxation are comparable in the CHNS data.



stable while middle school and high school dropout rates increased for poor areas (Hannum and Park, 2007).

The CHNS data show that counties with some relaxation and counties with no relaxation had similar geographic access to schooling in 1989. However, the data does not reveal quality of schooling or the changes in school availability during the early 1980s. Because relaxed areas tend to be more rural, it is likely that the quality of schools declined in relaxed areas during the same time that the 1-son-2-child relaxation took effect. To control for this, I will compare outcomes for girls to boys within counties. The strategy is robust as long as the changes in school quality and the economic conditions that determine school quality in relaxed areas have the same impact on both boys and girls.

### 3 Conceptual Framework

This section presents a simple model of the household that illustrates the quantity-quality tradeoff.<sup>6</sup> The empirical analysis will only be able to examine the effects on the first child. For simplicity, the following model assumes that the effects are similar across birth-parities. Let the utility of a household be a function of the number of children it has,  $n$ , and the set of education choices for each child  $i$ ,  $\{e_i\}_0^n$ .  $e_i$  is the level of education chosen for child  $i$ .

$$U(n, \{e_i\}_0^n) \tag{1}$$

The cost of education is the sum of a fixed cost of educating each child,  $a$ , and a function of the level of education and the number of children in the household.

$$c(e_i) = f(e_i, n) \tag{2}$$

---

<sup>6</sup>It draws from the intuition proposed by the two models in the economics and sociology literature that predict an interaction between the quantity and quality of children. The classic quantity-quality model dates back to Becker (1960), Becker and Lewis (1973) and Becker and Tomes (1979). To explain the decrease in fertility, they theorized that when female wages increase, the substitution effect (e.g. between child care and working) will outweigh the income effect. This model also predicts that quality monotonically decreases with quantity. Similarly, the “resource dilution” model in sociology theorizes that when the number of children increase, the average resource available to each child decrease, thereby creating a quantity-quality tradeoff.

Outside of economics, in disciplines such as social learning and child psychology, studies focus on the determinants of child production functions. For example, the “confluence model”, which to date has not been explored in the economics literature argued that children benefit from interacting with adults and teaching younger children (Zajonc and Gregory, 1985). The quantity and quality of children are inversely related, because increasing the number of children decreases the adult-to-children ratio within a household. At the same time, children from one-child families and the youngest child from a multi-child family are worse off because they cannot take advantage of the learning which comes from teaching younger children. This model, therefore, predicts an “inverse-U” shape for the relationship between quantity and quality of children.

The budget constraint faced by the household is then

$$\sum_{i=1}^n (a + c(e_i)) \leq y \quad (3)$$

I assume that the cost of education is increasing in years of education,  $f_{e_i} > 0$  and the rate of increase is positive,  $f_{e_i e_i} \geq 0$ . I also assume that the cost of education is decreasing in number of children,  $f_n < 0$ . This departure from the standard quantity-quality framework is motivated by the observation that there are often returns to scale in education. For example, children can share books and other school supplies. School clothes can be handed down from older to younger children. Having a second child will lower the average and marginal cost of school attendance for the first child as long as the secondary market for these goods function such that transferring the goods to children from other families is more costly than transferring them to children within the household. Moreover, children who are close enough in age can also share transportation costs to school.

The constrained maximization faced by the household is

$$\mathcal{L} = \max_{n, e_i} U(n, \{e_i\}_0^n) - \lambda \left[ \sum_{i=1}^n (a + c(e_i)) - y \right] \quad (4)$$

The first order conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial n} &= \frac{\partial U}{\partial n} - \lambda [a + f(e_i, n) + n f_n(e_i, n)] \equiv 0 \\ \frac{\partial \mathcal{L}}{\partial e_i} &= \frac{\partial U}{\partial e_i} - \lambda [n f_{e_i}(e_i, n)] \equiv 0 \end{aligned} \quad (5)$$

Denote the equilibrium education investment and number of children  $e^*$ ,  $n^*$ . Note that in equilibrium,  $e^*$ , will be similar for all  $i$ .

$$\begin{aligned} \frac{\partial U}{\partial n^*} &= \lambda [a + f(e^*, n^*) + n^* f_{n^*}(e^*, n^*)] \\ \frac{\partial U}{\partial e^*} &= \lambda [n^* f_{e^*}(e^*, n^*)] \end{aligned} \quad (6)$$

The ratio of the marginal utility from quantity to the marginal utility from quality can be written as

$$\frac{\frac{\partial U}{\partial n^*}}{\frac{\partial U}{\partial e^*}} = \frac{a + f(e^*, n^*)}{n^* f_{e^*}(e^*, n^*)} + \frac{f_{n^*}(e^*, n^*)}{f_{e^*}(e^*, n^*)} \quad (7)$$

The last term is the ratio of the marginal cost of educating one additional child holding quality constant to the marginal cost of educating each child for one additional year holding quantity constant.

In order for there to be a negative quantity-quality tradeoff, the marginal utility of education with respect to years of education must be decreasing with the number of children.

$$\begin{aligned} \frac{\partial U_e^*}{\partial n^*} &= \frac{\partial U / \partial e^*}{\partial n^*} < 0 \\ &\rightarrow \frac{f_{e^*}(e^*, n^*)}{n^* f_{e^* n^*}(e^*, n^*)} > 1 \end{aligned} \tag{8}$$

The cross partial derivative,  $f_{e^* n^*}(e^*, n^*)$ , reveals how quickly the cost of an additional year of education falls with the number of children. The right hand side expression,  $n^* f_{e^* n^*}(e^*, n^*)$ , is the amount you save on an additional year of schooling when you have an additional child multiplied by the number of children you have. An additional child lowers the marginal cost of an additional year of schooling for all children. Hence, if the marginal cost of an additional year of education is greater than the total savings parents gain by having an additional child, then there will be a quantity-quality tradeoff. On the other hand, if the cost of an additional year of education is less than the savings parents gain from having an additional child, then there will be no tradeoff and quantity will compliment quality.

## 4 Data

This paper matches the 1% sample of the 1990 *Population Census* with the 1989 *China Health and Nutritional Survey* (CHNS) at the county level. The 1990 *Population Census* contains 52 variables including birth year, region of residence, whether an individual currently lives in his/her region of birth, sex, relationship to the head of the household and whether an individual is enrolled in school at the time of the survey. The data allows children to be linked to parents. Thus, family size and birth order of children within a household can be calculated. Because the identification is partially derived from the region of birth, the sample is restricted to individuals who reported living in their birth place in 1990. The CHNS uses a random cluster process to draw a sample of approximately 3,800 households with a total of 16,000 individuals in eight provinces that vary substantially in geography, economic development, public resources, and health indicators. Most importantly, the survey provides detailed village and township level information on family planning policy enforcement. Since ethnic minorities were exempt from all family planning policies, I restrict the analysis to four provinces which are mostly composed of individuals of Han ethnicity. The matched dataset

contains 21 counties in four provinces.<sup>7</sup> These provinces are in the middle range of GDP and GDP growth during this period. They do not include rich coastal provinces or poor interior provinces. The CHNS data is aggregated to the county level in order to be matched to the 1990 Census. Since the policy data is at the village and township level, the aggregated data set reports the percentage of the population in each county that is exposed to the relaxation.

For the analysis of family size and education, the sample is restricted to first-born children in cohorts born during 1962-1981. The reference group in the regression analysis is comprised of individuals born during 1962-1972. Those born after 1981 are excluded because after the relaxation was first introduced, parents who preferred larger families may have chosen to keep girls in order to have a second child. In that case, the 2SLS estimate without excluding those born after 1981 will be biased by parental preferences and show that girls with larger family sizes are better off.

Panels A and B in Table 1A show that amongst first-born children, girls on average have more siblings, more educated parents and higher school enrollment. Panels C and D show that only children are more likely to be male, have more educated parents and are more likely to be enrolled in school. This is consistent with the identification concern that parents with more education may prefer to have fewer children and value education more.

To use individuals in counties without relaxations as a control group for individuals in counties with relaxations, we would like the two groups to have similar characteristics in every respect other than the relaxation. Table 1B compares first-born children born in counties with no relaxation and first-born children born in counties with some relaxation. It shows that the two types of counties have similar demographic characteristics. Each has approximately 55% males amongst first-born children. Family size and sex composition of siblings are similar. Children born in relaxed areas have slightly more educated parents. School enrollment in both counties with and counties without relaxations are approximately 50%. Panel B shows that in counties with some relaxations, approximately 38% of first-born children are born in villages or townships with the relaxation. This is comparable to the survey evidence from White (1992). Counties with relaxations are further away from urban municipalities.

The treatment group comprises of children who are 9 to 14 years old in 1990. In principle, they should be enrolled in primary school or junior high school. The control group comprise of children who in principle should be in high school. The descriptive statistics show that children in counties with the relaxation must on average travel further to attend primary school. This biases against my finding a positive effect of the relaxation on school enrollment. The distances to middle schools and high schools are very similar between counties with and without the relaxation.

---

<sup>7</sup>Liaoning, Jiangsu, Shandong and Henan.

## 5 Empirical Strategy

Figure 1 plots the total number of children against the birth year of the first-born child. It shows that children born in more recent years have smaller family sizes. This reflects both the fact that parents of young children may not have finished having children and a decrease in family size over time. To reveal the commonly seen OLS evidence for the quantity-quality trade-off, I regress a dummy variable for school enrollment on dummy variables for the number of children in a household. Children from one-child households are the reference group. Figure 2A plots the coefficients. It shows that family size is negatively correlated with school enrollment regardless of whether county fixed effects are controlled for. However, this confounds the family size effect with several factors: 1) younger children are more likely to be in school; 2) younger children will have fewer siblings because their parents may not have finished having children; and 3) quantity and quality may be jointly determined by parental preferences. Controlling for birth years addresses the first two problems and causes the relationship between family size and school enrollment to become non-monotonic. Figure 2B plots the coefficients for family size when controlling for birth year fixed effects.<sup>8</sup> Relative to the reference group of children from one-child families, children from two-child families have higher school enrollment. However, the correlation between enrollment and family size is negative for households with two to five children.

The main second stage equation will control for birth county and birth year fixed effects. It can be written as the following.

$$enroll_{itc} = sibs_{itc}b + X'_{ict}\kappa + \sum_{l=1973}^{1981} (urban_c \times d_{il})\delta_l + \alpha + \gamma_t + \psi_c + \varepsilon_{itc} \quad (9)$$

School enrollment for individual  $i$ , born in county  $c$ , birth year  $t$ , is a function of:  $sibs_{itc}$ , the number siblings he or she has;  $X_{ict}$ , individual characteristics; the interaction term between  $urban_c$ , distance to urban area, and  $d_l$ , a variable indicating whether an individual was born in year  $l$ ;  $\gamma_t$ , birth year fixed effects; and  $\psi_c$ , county fixed effects.

This faces the problem that the number of children and investment in these children are jointly determined by parents. Hence, if parents who value education also prefer smaller households, then OLS will over-estimate the negative effect of an additional sibling on schooling. I address this by exploiting plausibly exogenous variation in family size caused by relaxations in the One Child Policy. Sex, date and region of birth jointly determine an individual's exposure to the 1-son-2-child relaxation. The relaxation allowed parents to have a second child only

---

<sup>8</sup>Estimates for the coefficients plotted in Figures 2A and 2B are shown in Appendix Table A1.

if the first-born child was a girl. Therefore, family size should be positively correlated with being a girl. Since parents are more likely to have a second child if the first girl was younger when the relaxation was announced, family size should be negatively correlated with the age of the first girl. The interaction between whether a girl was born in a relaxed area and her age estimates the effect of the relaxation on family size. The additional comparison with boys controls for region and year specific changes in policies such as education provision that affected both boys and girls similarly. The instrument for family size is therefore the triple interaction of an individual's sex, year of birth and region of birth. Only the combination of the three is exogenous. The exclusion restriction for the instrument is that it must be correlated with family size and have no direct effect on school enrollment or other right-hand-side variables.

To illustrate the identification strategy, I first estimate the effect of the policy on family size for boys and girls separately. If the policy was fully enforced, it should increase the number of siblings for first-born girls for whom the One Child Policy prevented from having younger siblings. The relaxation should have no effect on the family size of boys. I estimate the following equation separately for samples of first born boys and girls born during 1962-1981.

$$sibs_{itc} = \sum_{l=1973}^{1981} (relax_c \times d_{il})\beta_l + \gamma_t + \alpha + \psi_c + v_{itc} \quad (10)$$

The number of siblings for individual  $i$ , born in county  $c$ , birth year  $t$ , is a function of: the interaction term of  $relax_c$ , the extent of relaxation in county  $c$  and  $d_{il}$ , a dummy indicating whether the individual was born in year  $l$ ;  $\gamma_t$ , birth year fixed effects and  $\psi_c$ , county fixed effects. The reference group is comprised of individuals born during 1962-1972. It and all of its interaction terms are dropped. For all regressions, standard errors are clustered at the county level. If the four year birth spacing rule was well enforced, then  $\hat{\beta}_l > 0$  for  $l \geq 1976$ . If it was not well enforced, then  $\hat{\beta}_l > 0$  for  $l \geq 1980$ .

Then, to assess the statistical difference of the effect on boys and girls, I pool the data to estimate the first stage equation with the triple interaction terms on the right hand side.

$$\begin{aligned} sibs_{itc} = & \sum_{l=1973}^{1981} (relax_c \times girl_{itc} \times d_{il})\beta_l + \sum_{l=1973}^{1981} (relax_c \times d_{il})\delta_l \\ & + \sum_{l=1973}^{1981} (girl_{itc} \times d_{il})\zeta_l + (relax_c \times girl_{itc})\lambda + girl_{itc}\kappa \\ & + \alpha + \gamma_t + \psi_c + v_{itc} \end{aligned} \quad (11)$$

The number of siblings for individual  $i$ , born in county  $c$ , birth year  $t$ , is a function of:

the triple interaction term of  $relax_c$ , the extent of relaxation in county  $c$ ,  $girl_{itc}$ , a variable indicating whether a child is a girl and  $d_{il}$ , a dummy variable indicating whether the individual was born in year  $l$ ; the interaction term of  $relax_c$  and  $d_{il}$ ; the interaction term between  $girl_{itc}$ , and  $d_{il}$ ; the interaction term between  $relax_c$  and  $girl_{itc}$ ;  $girl_{itc}$ ;  $\gamma_t$ , birth year fixed effects; and  $\psi_c$ , county fixed effects. As before, the reference group of cohorts born 1962-1972 and all its interactions are dropped.  $\beta_l$  is the difference in the effect of being born in a relaxed area on family size between girls and boys born in year  $l$ . The estimates should be zero for earlier cohorts who were not affected by the One Child Policy and relaxation and positive for later affected cohorts.  $\beta_l$  is the effect of being born in a relaxed county on family size for an individual born in year  $l$ .

Like simple differences-in-differences estimators, cohort-invariant differences across regions are differenced out by the comparison across cohorts. Changes across cohorts which affect different regions similarly are differenced out by the comparison across regions. The triple difference adds the advantage that cohort-varying differences that affect boys and girls similarly across regions are also differenced out by the comparison between girls and boys within each cohort and region. The exclusion restriction is only violated if a change with differential impacts on areas with and without the relaxation *and* on boys and girls occurs at the same time the relaxation took effect. In other words, the 2SLS estimate will be biased only if there is a sex-specific change at the time of the relaxation in relaxed regions. For example, if local governments of relaxed regions implemented a program encouraging girls to attend school when the relaxation was enacted, then the 2SLS will confound the effects of this program with the effects of family size. There is little reason to think that such a change occurred.

The main concern with this strategy arises from the fact that the relaxations were introduced to curb sex selection. If the relaxation is strongly correlated with the extent of sex-selection for One Child Policy cohorts, two potential problems will arise. First, unobserved factors correlated with sex-selection may affect education investment differentially for boys and girls. This will bias the estimates if the factors driving sex-selection are time varying. For example, the larger boy-preference in regions that received the relaxation may be driven by higher relative adult male wages (Qian, 2008). Since education most likely has positive returns in the labor market, this would probably attenuate the 2SLS estimates of this study.<sup>9</sup> Second, there might be selection bias regarding the parents who choose to keep girls in relaxed regions. The main concern is that parents of girls in relaxed regions could have different unobservable characteristics from parents of girls in regions without the relaxation in such a way that would bias the 2SLS estimates upwards. For example, parents of girls in relaxed

---

<sup>9</sup>The CHNS does not have accurate data on individual income within the household since much of rural production is conducted at the household level and income cannot be accurately assigned to individual members. Consequently, I cannot directly examine the role of relative earnings in this study.

regions may on average have a higher consumption value for all things related to children such as education relative to parents of girls in non-relaxed regions. Then, the 2SLS estimate will overestimate the true effect of family size on school enrollment.

I address this concern by first investigating the extent to which the One Child Policy increased sex selection in areas that received the relaxation relative to other areas. The following equation estimates the effect of the relaxation on the fraction of males by birth year.

$$male_{itc} = \sum_{l=1969}^{1989} (relax_c \times d_{il})\beta_l + \gamma_t + \alpha + \psi_c + v_{itc} \quad (12)$$

The probability of being male for individual  $i$ , born in county  $c$ , birth year  $t$  is a function of: the interaction terms between  $relax_c$ , and birth year dummy variables,  $d_{il}$ ; birth county fixed effects,  $\psi_c$ ; and birth year fixed effects,  $\gamma_t$ .  $\beta_l$  is the correlation between being born in a relaxed county and the sex ratios of your cohort for each birth year  $l$ . To estimate the magnitude of the effect of the relaxation on sex ratios, a simpler specification is estimated

$$male_{itc} = \sum_{l=2}^3 (relax_c \times post_{il})\delta_l + \alpha + \gamma_t + \psi_c + \varepsilon_{itc} \quad (13)$$

The probability of being male for individual  $i$ , born in county  $c$ , birth year  $t$  is a function of: the interaction term between  $relax_c$ , and  $post_{il}$ , a variable indicating the individual's cohort group;  $\psi_c$ , county fixed effects and  $\gamma_t$ , cohort group fixed effect. The children are divided into three groups according to birth cohort. The reference group is comprised of individuals not affected by the One Child policy and the relaxation (born before 1978). The second group is comprised of children born after the One Child Policy but before the relaxation (1978-1981). The third group is comprised of children born after the relaxation (1982-1989).  $\hat{\delta}_l$  is the effect of the One Child Policy on sex selection in relaxed areas relative to areas without the relaxation.

Then, using the estimate of  $\hat{\delta}_l$ , I can calculate bounds for the effect of family size on school enrollment under the assumption that parents who kept daughters in areas that received the relaxation also value education more. See section on robustness for details.



## 6 Empirical Results

### 6.1 OLS Estimates of the Effect of a Sibling on Schooling

Panels A and B of Table 2 show the estimates from equation (9). All regressions control for the full set of double interaction terms from the first stage equation.<sup>10</sup> Columns (1)-(6) of Panel A show that amongst households with three or fewer children, an additional sibling decreases school enrollment of the first child by 1.1 percentage points. However, since the 2SLS will reveal the effect of increasing the number of children from one to two, the relevant OLS comparison should be on a sample of individuals with one or no sibling. Columns (1)-(6) of Panel B shows that in this sample, an additional sibling *increases* school enrollment by approximately 1.5 percentage-points. Estimates in both cases are statistically significant at the 1% significance level and robust to controls. In comparing the results of this study to studies that find a negative family size effect, the estimates in Panels A and B shows clearly that the sign switch is induced by restricting the comparison to be between households with one child to households with more than one.

### 6.2 The Effect of the 1-Son-2-Child Relaxation on Family Size

I first estimate equation (10) on separate samples for boys and girls. The estimates are shown in Table 3, columns (1) and (2). The estimates for girls are statistically significant at the 1% level for individuals born 1976 and later. This is consistent with the fact that before the One Child Policy was introduced in 1979/1980, there was a four year birth spacing law. Hence, the One Child Policy was binding for cohorts born four years previous to its introduction. The estimates for boys are statistically insignificant. The coefficients are plotted in Figure 3A. It shows that family size for boys and girls were similar for cohorts born 1973-1976, after which the family size for girls increased and the family size for boys remained the same.

The estimated coefficients for the triple interaction terms from equation (11) are shown in Table 3, column (5). They are statistically significant at the 5% level for the individuals born 1977-1981. Figure 3B plots the coefficients for the triple interaction term. It shows that the boy-girl difference in the effect of being born in a relaxed area on family size is zero for unaffected cohorts and positive for the affected cohorts. On average, the relaxation increased family size of first-born girls by approximately 0.25 children. In other words, approximately one out of every four children gained a sibling due to the relaxation. The discrete change in

---

<sup>10</sup>The double interactions include the interaction term of  $relax_c$  and  $d_{it}$ ; the interaction term between  $girl_{itc}$  and  $d_{it}$ ; the interaction term between  $relax_c$  and  $girl_i$ ; and  $girl_{itc}$ . The reference group is comprised of cohorts born during 1962-1972. The dummy variable for the reference cohort and all its interactions are dropped.

the effect of the triple interaction term between individuals born before 1976 and those born afterwards suggests that the previous four-year birth spacing rule was well enforced.

### 6.3 The Effect of the 1-Son-2-Child Relaxation on Enrollment

I first estimate the effect of the relaxation on enrollment separately for boys and girls using an equation identical to equation (10) but replacing the dependent variable with enrollment, the outcome of interest. The reference group is comprised of individuals born during 1962-1972. The dummy variable for the reference group and all its interactions are dropped. The coefficients for girls and boys are shown in Table 3, columns (3) and (4). The estimates are statistically significant for girls. Figure 4A plots the estimates for boys and girls. The plot of the reduced form shows that girls affected by the relaxation (born 1976 and after) had higher education enrollment than boys, whereas girls unaffected by the relaxation (born before 1976) had lower school enrollment rates than boys.

The estimates in Figure 4A show that relative to areas without the relaxation, enrollment for both boys and girls decreased after primary school. This is consistent with the hypothesis that school provision and quality in relaxed regions relative to regions without the relaxation declined during this period. I control for this by comparing the effect of the relaxation on enrollment for boys with the effect of the relaxation on enrollment for girls. I estimate an equation similar to equation (11) with school enrollment as the dependent variable. The reference group is comprised of individuals born during 1962-1972. The dummy variable for the reference group and all its interaction terms are dropped. The coefficients are shown in Table 3, column (6). The estimates show that for older cohorts not affected by the relaxation, individuals born in relaxed areas had on average 1% to 17% less school enrollment than areas without the relaxation. However, for cohorts affected by the relaxation, individuals born in relaxed areas were on average enrolled in school 5% more than individuals born in areas without the relaxation. The estimates are statistically significant at the 1% level. Figure 4B plots the triple difference reduced form estimates. It shows that school enrollment in relaxed areas was higher for girls of the affected cohort than for boys.<sup>11</sup>

### 6.4 2SLS Estimates of the Effect of a Sibling on Schooling

Columns (1)-(6) of Table 2 Panel C show the 2SLS estimates for households with three or fewer children. It shows that for a sample where 49% of individuals were enrolled in school, an

---

<sup>11</sup>Reduced from estimates for sample restricted to two or fewer children are presented in Appendix Table A2 column 6.

additional sibling increased enrollment of the first child by approximately fourteen to sixteen percentage-points. The estimates are statistically significant at the 1% level. Panel D restricts the sample to households with two or fewer children. The estimates show that for a sample where 54% of individuals were enrolled in school, an additional sibling increased enrollment of the first child by approximately 12 percentage-points. The estimates are mostly statistically significant at the 10% level.

## 6.5 Robustness

### 6.5.1 Selection

If parents of relaxed regions who chose to keep girls born under the One Child Policy valued education more than parents of other regions who kept girls, then the 2SLS strategy will overestimate the true effect of family size on school enrollment. This problem is partially addressed in the main estimation by excluding cohorts born after the relaxation first began. This excludes households which kept girls in order to have a second child.

To further address the problem of sample selection, I construct an alternative sample to estimate the lower bound of the absolute value of the family size effect. I remove the “extra” boys from relaxed counties and replace them with girls that I construct so that for each cohort, the sex ratio is equivalent between counties with some relaxation and counties without any relaxation. Only boys who are not enrolled in school are removed. Added girls are assumed to be enrolled in school. This increases the average enrollment rate for boys born 1976-1982 in counties with the relaxation, and decreases average enrollment rate for girls in counties with the relaxation. 2SLS using this “stacked” sample will be biased against finding a positive effect of family size on school enrollment and allow me to estimate the lower bounds of the positive family size effect and investigate the extent to which the main results are driven by selection.

To estimate the number of “extra” boys, I first need to examine the extent of sex-selection in regions that received the relaxation before the relaxation was enacted.<sup>12</sup> Any sex-selection caused by the One-Child Policy should be observed for cohorts born closely before and after 1980. I estimate the equation (12) using a sample of cohorts born between 1962 and 1989 for each birth parity.<sup>13</sup> The dependent variable indicates whether an individual is male. Table 4

---

<sup>12</sup>Past studies comparing hospital birth records and population census data, or by comparing sex ratios for the same cohort at different ages have found that sex-selection mostly occurs at very young ages, which is consistent with the lack of prenatal gender revealing technology and tough government enforcement against infanticide (Qian, 2008; Zeng et. al., 1993). Hence, any sex-selection caused by the One-Child Policy should be observed for cohorts born closely before and after 1980.

<sup>13</sup>Because of widespread under reporting of children under one year of age, I exclude the 1990 cohort (Zeng,

column (1) shows the estimates of  $\beta_l$  for first-born children. They are statistically significant. Column (2) shows that the estimates are robust to the addition of a control for whether individuals are ethnically Han. Figure 5A plots the coefficients and 95% confidence intervals. It shows that in areas that received the relaxation, the fraction of males increased after the One Child Policy relative to other areas. It also shows that the relaxation decreased the fraction of males to the pre-One Child Policy level.<sup>14</sup> To assess the magnitude, I estimate  $\delta_l$  from equation (13). Table 5 column 1 shows that amongst first-born children, there were 10.6 percentage-point more males in relaxed regions due to the initial One Child Policy. The estimate is statistically significant at the 1% level. The relaxation caused the sex ratios to be similar between regions.

Using the estimate  $\hat{\delta}_l$ , I calculate the number of extra boys due to the One Child Policy and construct the stacked sample to repeat the estimate of the main effect of family size on school enrollment. They are similar to the results from using the actual data. They are not reported in the paper for brevity. These results suggest that the main results are not driven by selection.

### 6.5.2 Family Composition

Using the triple difference as an instrument for family size requires that the instrument does not affect any right-hand-side variable other than family size. One concern is that the relaxation also changed the sex composition of children in families of the affected cohort. Then the 2SLS estimate will be confounded. I can explore this possibility by estimating equation (12) for second and higher parity births. The coefficients and standard errors for second born children are shown in Table 4 columns (3) and (4). For third and higher parity births, they are shown in columns (5) and (6). They and their 95% confidence intervals are plotted in Figures 5B-5C. The figures show that the One Child Policy and subsequent relaxations did not affect sex ratios of higher order births in relaxed counties relative to counties without relaxations. These results support the identification assumption.

---

1992). The reference cohort is composed of individuals born during 1962-1968.

<sup>14</sup>Note that although the One Child Policy constrained the family size of individuals born as early as 1976, sex-selection from the One Child Policy appears only in cohorts born after 1978. A plausible explanation is that sex-selection in China mostly occurs for very young children. In other words, once the policy was announced in 1979/80, parents were unwilling (or unable) to sex-select for children that were more than 1 or 2 years of age.

### 6.5.3 Marriage Market

One possible concern for the empirical strategy is that the instrument affected enrollment through channels apart from family size. In particular, the results may be driven by marriage market channels. The intensive boy-biased sex selection in regions that received the relaxation may have increased the value of girls in these regions. This may have a direct positive effect on the investment in girls' education apart from family size effects if there are positive returns to education on the marriage market. However, the increase in the value of girls in relaxed areas should also have resulted in an increase in female survival. Figures 5A-5C show that female survival in relaxed regions did not improve for any birth parity after the relaxation. This is inconsistent with the marriage market explanation.

## 6.6 Mechanisms

### 6.6.1 Economies of Scale

If the positive effect of the second child is driven by economies of scale in education, then one would reasonably expect the positive effects to be larger in magnitude when both children are of the same sex, i.e. the younger child is a girl. Economies of scale for same sex children would occur if children of the same sex are more likely to share similar interests or needs. For example, a higher percentage of a daughter's clothes can be re-used by a younger daughter than a younger son. To investigate whether such economies of scale exist, I separately estimate the 2SLS effect of family size on a sample excluding those where the first two children are the same sex, and a sample excluding those where the first two children are of different sex.<sup>15</sup> The estimates for the two samples are shown in columns (1) and (2) of Table 6. The positive main effects are driven by households where the children are of the same sex. This should be interpreted cautiously since sex can be endogenously chosen by parents.<sup>16</sup>

### 6.6.2 Permanent Income

Here, I explore the alternative hypothesis that the main results are driven by an increase in permanent income. In rural China, parents rely on children for income during old age.

---

<sup>15</sup>Approximatey 24% of the sample have siblings of the same sex.

<sup>16</sup>I also estimated the differential effect of family size across different age gaps between the first two children. The results showed that the benefits are larger when larger age gaps exist between children. Interpretation of this result is made difficult that age gaps can be a result of sex selection. Parents who want a son for a second child and who are constrained to have no more than two children will on average have further spacing between their children than parents who have weaker son preference or parents who have limited ability to sex select. If the ability to select is correlated with factors that also determine education such as income, then the estimated interaction effect will reflect the influence of those factors.

Therefore, an additional child can be seen as an increase to permanent income. This will lead to an increase in expenditures on children's schooling (and hence enrollment) if parents are not credit constrained.

It is very unlikely that households in rural Chinese communities during the early and mid 1980s were not credit constrained and could borrow against the future earnings of their children. That said, I can test the hypothesis that the main results are driven by an income effect under the assumption that parents are not credit constrained and believe that sons will earn more than daughters. In this case, any income effects should cause the positive effect on the enrollment of the first child to be larger when the second child is a boy. Hence, I divide the sample into those that do not have a younger sister and those that do not have a younger brother and examine whether the main effect differs by the sex of the younger child. The results are shown in Table 6 columns (3) and (4). They show that the effects are *larger* for those with a younger sister. This is inconsistent with the permanent income explanation. These results should be interpreted cautiously since the sex of the younger sibling can be endogenously chosen by parents.

### 6.6.3 Mother's Labor Supply

An additional channel through which having a younger sibling could affect the first child is the mother's labor supply. This can easily coexist with any economies of scale in raising children. If having a second child increases household needs for monetary income and public school is a source of inexpensive child care, then an additional child may cause the mother to enter the labor force and send the older child to school. As long as households are credit constrained, this will be true even if having more children increases parents' future (permanent) income. To explore this possibility, I estimate the effect of the relaxation on school delay of the first child and mother's work status.

Estimating the effect on school delay does not provide a clean test. It reflects two potentially offsetting forces. On the one hand, children with younger siblings may attend school earlier if parents wish to decrease the amount of at-home child care needed during the day. On the other hand, parents may hold back the first child with the belief that there are economies of scale to having two children in school at the same time. Using a restricted sample of individuals who are currently enrolled in school, I repeat the main estimation with school delay as the dependent variable. It is measured as the difference between an individual's years of education and the years of education she should have had assuming that she began at age seven.<sup>17</sup> Table 7 Panel A shows that on average, first-born children of households with three and fewer

---

<sup>17</sup>The means are shown in Appendix Table A1 columns (5)-(8).

children are 0.5 years and 1.6 years ahead in schooling relative to the legal requirement. The sample means are similar for boys and girls. The OLS estimates in Panel A show that an additional sibling is correlated with being behind in school relative to the mean. But the estimates are not statistically significant. In contrast, the 2SLS estimates show that having a younger sibling causes the first child to attend school earlier. However, these estimates are also not statistically significant and are at best suggestive.

Next, I estimate the effect of having an additional child on mother's labor supply. The dependent variable is a dummy variable that equals one if the mother does not work outside of the home. The results are presented in Table 8. The estimates are negative and almost statistically significant at the 10% level. They suggest that an additional child causes the mother to be *less* likely to stay at home and *more* likely to participate in the labor market. Columns (5) and (6) show that the effect is statistically similar between those with a younger son and those with a younger daughter. These results are consistent with the hypothesis that parents view schools as an alternative source of child care for the first child and send her to school while the mother enters the labor force.

## 7 Conclusion

This paper estimates the effect of family size on school enrollment for first-born children. It resolves the problem of joint determination by exploiting the plausibly exogenous variation in family size caused by relaxations in the One Child policy. The results show that both the One Child Policy and the previous four-year birth spacing policy were well enforced; and that the 1-son-2-child relaxation increased family size for girls born in relaxed areas. Then, it uses the variation in family size caused by this relaxation to provide evidence that a second child increased school enrollment of the first child.

The empirical results do not directly speak to effect of family size in average school enrollment across all children within a household because the second child is too young to be enrolled in school when the data was collected. Hence, these results, taken literally, do not reject the predictions of the standard models of quantity-quality tradeoff. That said, the findings do provide empirical evidence for a novel insight for first born children, who have thus far been the focus of most existing empirical studies of quantity-quality. They show clearly that first born children benefit from having a younger sibling and suggest that the effect of family size on the first born child may be non-monotonic across the number of siblings.

It is beyond the scope of this paper to provide conclusive evidence on the mechanisms driving the main effects. The evidence suggests that there may exist economies of scale in schooling. This may be complimented by parents use of public schools as alternative child

care. Interpreting these results outside of the context of rural China requires caution. This is especially true if parents in these other contexts do not have access to inexpensive public schooling or good labor market opportunities.<sup>18</sup>

---

<sup>18</sup>To verify that the Chinese context is not inherently different from other contexts where the quantity-quality tradeoff has been studied, I replicate past studies that used the occurrence of twins at the second birth to instrument for having three children instead two children using the data of this study. The estimates are similar to other studies using the twins strategy. They show that increasing the number of siblings from two to three decreases school enrollment of the first child. The results are reported in previous versions of the paper.



## References

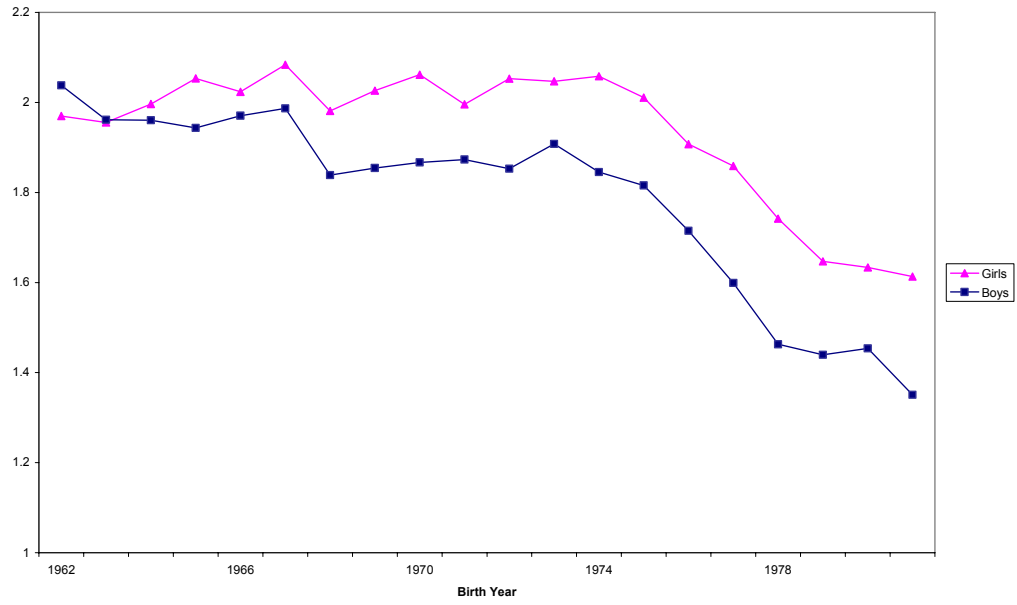
- [1] Angrist, Joshua “How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America’s Second Generation.” *The Quarterly Journal of Economics*, 117(3), 2002, pp. 997-1038.
- [2] Angrist, Joshua D. and Evans, Williams N. “Children and their parents’ labor supply: Evidence from exogenous variation in family size.” *The American Economic Review*, 88(3), Jun. 1998, pp. 450-477.
- [3] Angrist, Joshua D., Victor Lavy and Analia Schlosser “New evidence on the causal link between quantity and quality of children.” NBER Working Paper 11835, 2005.
- [4] Angrist, Joshua D., Victor Lavy and Analia Schlosser “Multiple Experiments for the Causal Link Between the Quantity and Quality of Children” MIT Working Paper, 2006.
- [5] Almond, Douglas, Kenneth Chay and David Lee “The Costs of Low Birth Weight”, *The Quarterly Journal of Economics*, August 2005.
- [6] Banister, J. *China’s changing population*. Stanford, California., Stanford University Press, 1987.
- [7] Becker, Gary S. and Lewis, H. Gregg “On the interaction between the quantity and quality of children.” *The Journal of Political Economy*, 81(2) Part 2, Mar.-Apr. 1973, pp. S279-S288.
- [8] Becker, Gary S. and Tomes, Nigel “Child endowments and the quantity and quality of children.” *The Journal of Political Economy*, 84(4) Part 2, Aug. 1976, pp. S143-S162.
- [9] Behrman, Jeremy; Pollak, Robert A. and Paul Taubman "Family resources, family size and access to financing for education." *The Journal of Political Economy*, 97(2), Apr. 1989, pp. 389-419.
- [10] Behrman, Jeremy and Taubman, Paul “Birth Order, schooling, and earnings.” *Journal of Labor Economics*, 4(3) Part 2, Jul. 1986, pp. S121-S145.
- [11] Black, Sandra E.; Devereux, Paul J. and Kjell G. Salvanes “The more the merrier? The effect of family composition on children’s education.” NBER, Working Paper 10720.
- [12] Black, Sandra E.; Devereux, Paul J. and Kjell G. Salvanes “From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes”, *Quarterly Journal of Economics*, March 2007.

- [13] Blake, Judith “Family size and the quality of children.” *Demography*, 18, pp. 421-442.
- [14] Bleakely, Hoyt and Lange, Fabian “Chronic Disease Burden and the Interaction of Education, Fertility and Growth.” *Review of Economics and Statistics*, Forthcoming 2009.
- [15] Brown, Phil and Park, Albert “Education and poverty in rural China.” *The Economics of Education Review*, Forthcoming.
- [16] Conley, Dalton “What is the ‘true’ effect of sibship size and birth order on education? Instrumental variable estimates from exogenous variation in fertility.” New York University, Working Paper.
- [17] Croll, E., D. Davin, et. al., Eds. *China’s One Child Family Policy*. London, Macmillan, 1985.
- [18] Duflo, Esther “Evaluating the effect of birth-spacing on child mortality.” MIT, Mimeo.
- [19] Galor, Oded and Weil, David “Population, Technology, and Growth: From Malthusian stagnation to demographic transition and beyond.” *The American Economic Review*, 90(4), 2000, pp. 806-828.
- [20] Gomes, Melba “Family size and education attainment in China.” *Population and Development Review*, 10(4), pp. 647-660.
- [21] Goux, Dominique, and Maurin, Eric “The effects of overcrowded housing on children’s performance at school.” CEPR Discussion Paper No. 3818.
- [22] Greenlaugh, S. “Shifts in China’s population policy, 1984-1986: Views from the central, provincial, and local Levels.” *Population and Development Review*, 12(3), 1986, pp. 493-515.
- [23] Guo, Guang and VanWey, Leah K. “Sibship size and intellectual development: Is the relationship causal?” *American Sociological Review*, 64(2), Apr. 1999, pp. 169-187.
- [24] Hannum, Emily “Education stratification by ethnicity in China: Enrollment and attainment in the early reform years.” *Demography*, 39(1), 2002, pp. 95-117.
- [25] Hannum, Emily and Park, Albert. *Education and reform in China*. New York, NY : Routledge, 2007.
- [26] Hanushek, Eric “The trade-off between child quantity and quality.” *The Journal of Political Economy*, 100(1), Feb. 1992, pp. 84-117.

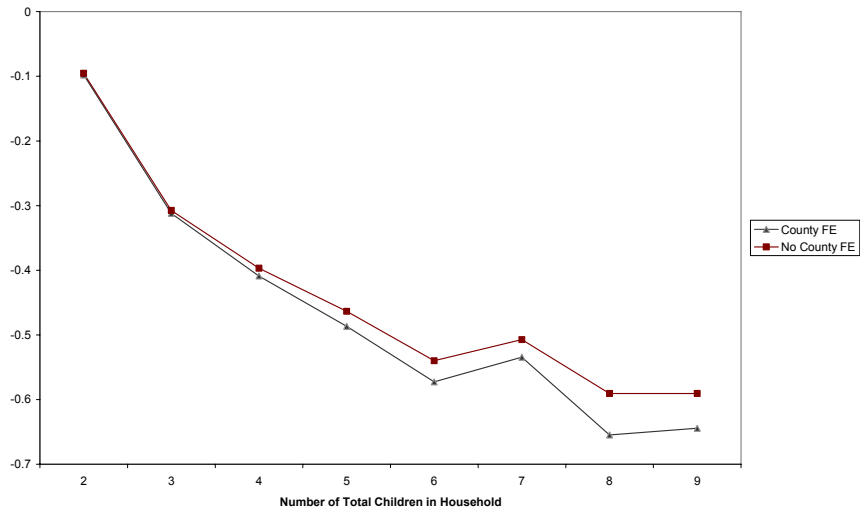
- [27] Hazan, M. and Berdugo, B. "Child Labor, Fertility and Economic Growth," *The Economic Journal*, 112(Oct.), 2002, pp. 467-475.
- [28] Iacovou, Maria "Family composition and children's educational outcomes." Institute for Social and Economic Research, Working Paper, 2001.
- [29] Kaufman, J., Z. Shang, et. al. "Family planning policy and practice in China: A Study of four rural counties." *Population and Development Review*. 15(4), 1989, pp. 707-729.
- [30] Kessler, Daniel "Birth order, family size and achievement: Family structure and wage determination." *Journal of Labor Economics*, 9(4), Oct. 1991, pp. 413-426.
- [31] Lee, Jungmin "Sibling size and investment in children's education: An Asian instrument." *Journal of Population Economics*, Forthcoming 2009.
- [32] Moav, Omer "Cheap Children and the Persistence of Poverty." *The Economic Journal*, Vol. 115, 2005, pp. 88-110.
- [33] Pettersson-Lidbom, Per and Peter Skogman Thoursie, "Does child spacing affect children's outcomes? Evidence from a Swedish reform," Stockholm University Working Paper, 2007.
- [34] Qian, Nancy "Missing women and the price of tea in China: The effect of income on sex imbalance." *Quarterly Journal of Economics*, November, 2008.
- [35] Rosenzweig, Mark R. and Wolpin, Kenneth I. "Testing the quantity-quality fertility model: The use of twins as a natural experiment." *Econometrica*, January 1980, 48(1), pp.227-240.
- [36] Rosenzweig, Mark R. and Schultz, T. Paul "Fertility and investments in human capital: Estimates of the consequence of imperfect fertility control in Malaysia." *Journal of Econometrics*, 1987, 36, pp. 163-184.
- [37] Rosenzweig, Mark R. and Junsen, Zhang "Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's One Child Policy", *Review of Economic Studies*, Forthcoming 2009..
- [38] Samuelson, Paul "Models of Thought in Economics and Biology." *American Economic Review Papers and Proceedings*, 75, 1985, pp. 166-172.
- [39] Schultz, T.P., "Effects of Fertility Decline on Family Well-Being: Evaluation of Population Programs." Draft for MacArthur Foundation Consultation Meeting, 2005.

- [40] Stafford, Frank P. "Women's work, sibling competition and children's school performance." *The American Economic Review*, 77(5), Dec. 1987, pp. 972-980.
- [41] White, T. "Birth planning between plan and market: The impact of reform on China's One-Child Policy", *China's Economic Dilemmas in the 1990's: The Problems of Reforms, Modernization, and Interdependence. Studies in Contemporary China*. Armonk, N.Y.U.S.C.J.E. Committee and London, Sharpe, 1992, pp. 252-69.
- [42] Zajonc, Robert B. and Markus, Gregory B. "Birth order and intellectual development" *Psychological Review*, 1982, pp. 74-43.
- [43] Zeng, Yi. "Is the Chinese family planning program 'tightening up?'" *Population and Development Review*, 15(2), 1989, pp. 333-37.
- [44] Zeng, Y., Tu, P. et. al. "Causes and implications of the recent increase in the reported sex ratio at birth in China." *Population and Development Review*, 19(2), 1993, pp. 283-302.

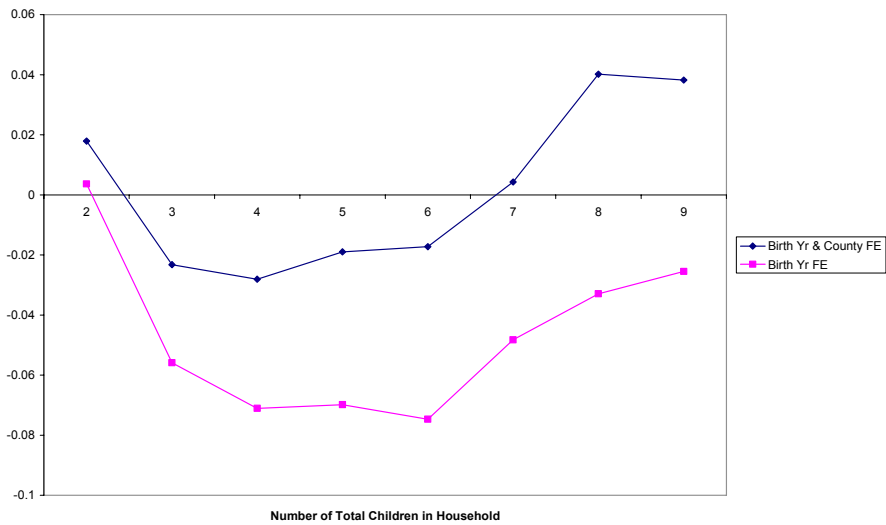
**Figure 1: The Number of Total Children in Household by Birth Year for Households with 3 or fewer children**  
Total Number of Total Children in a Household



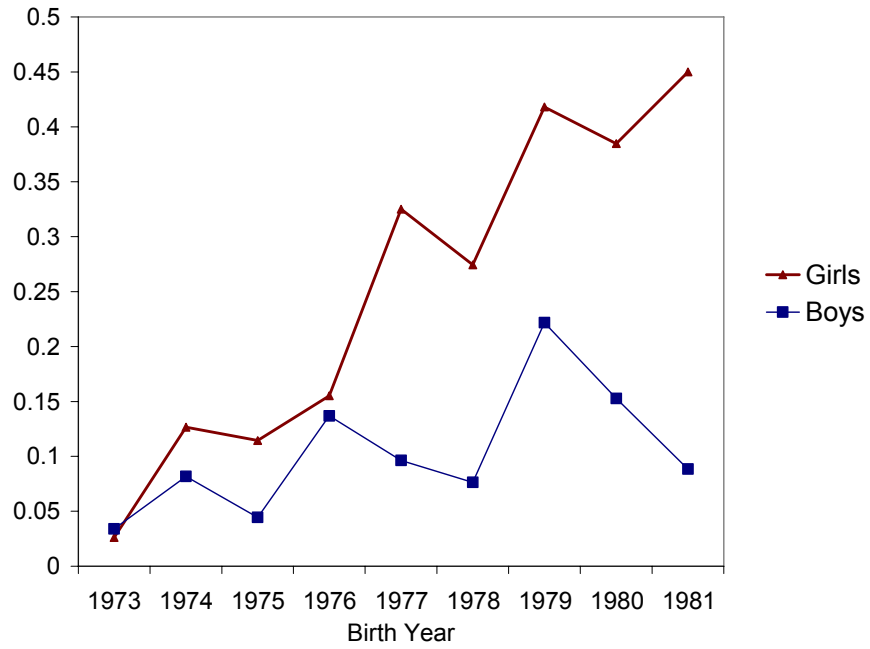
**Figure 2A: Correlation between Family Size and School Enrollment  
by Family Size with No Birth Year Controls**  
Coefficients of the number of total children in the household



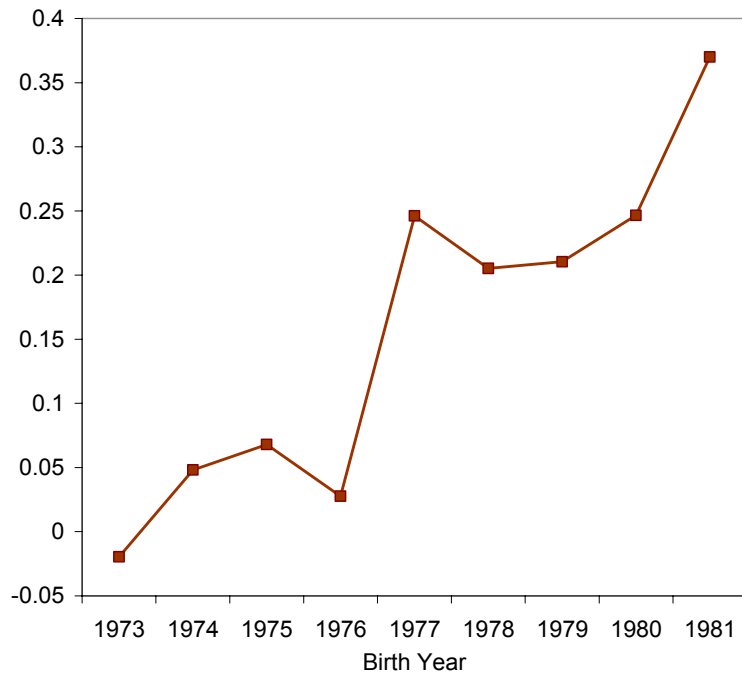
**Figure 2B: Correlation between Family size and School Enrollment  
By Family Size with Birth Year Controls**  
Coefficients of the number of total children in the household



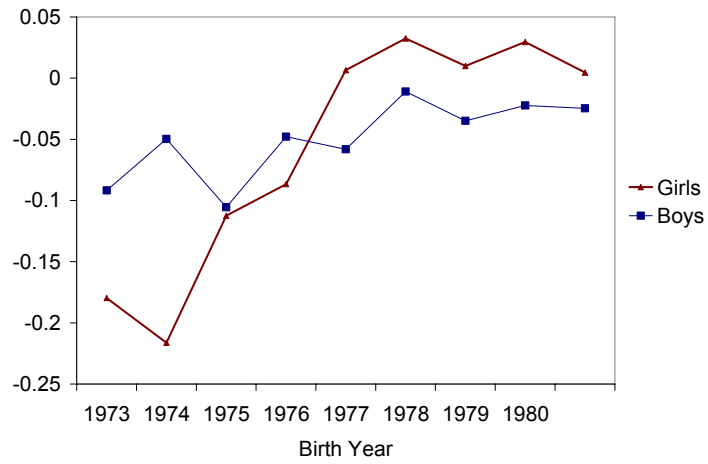
**Figure 3A: The Effect of Relaxation on Family Size**  
 Coefficients of the Interactions between  
 Born in a Relaxed Area \* Birth Regions



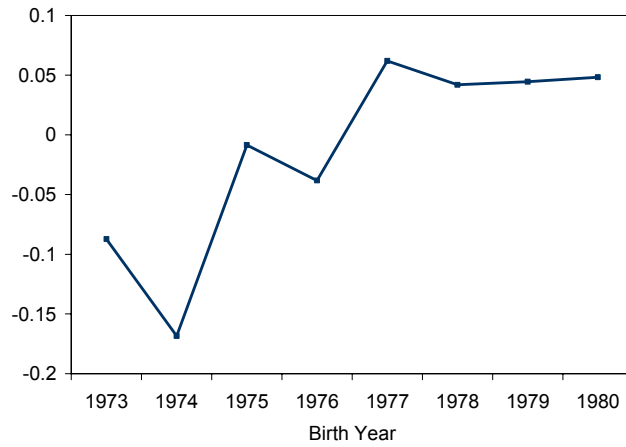
**Figure 3B: The Effect of Relaxation on Family Size**  
 Coefficients of the Interactions between  
 Dummy for Girl \* Born in a Relaxed Region \* Birth Year



**Figure 4A: The Effect of Relaxation on School Enrollment**  
 Coefficients of Interactions between  
 Born in Relaxed Region \* Birth Year

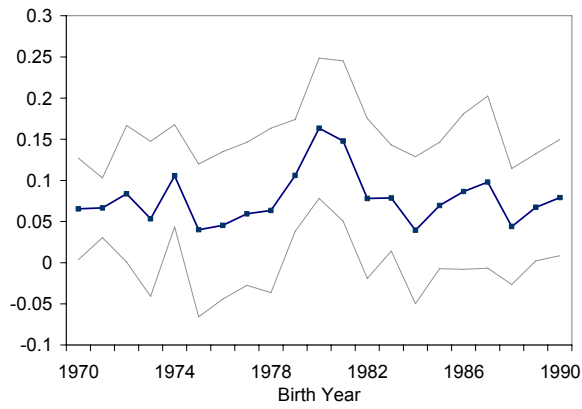


**Figure 4B: The Effect of Relaxation on School Enrollment**  
 Coefficients of Interactions between  
 Dummy for Girl \* Born in Relaxed Region \* Birth Year

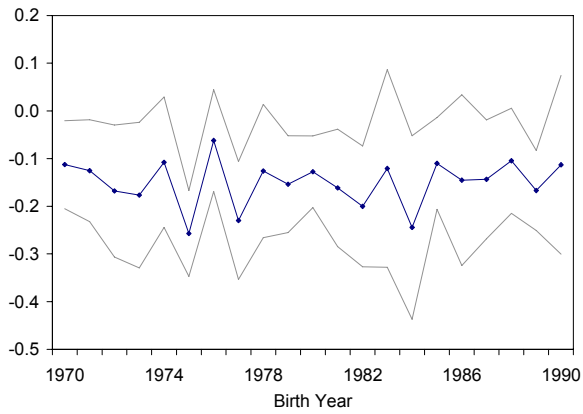




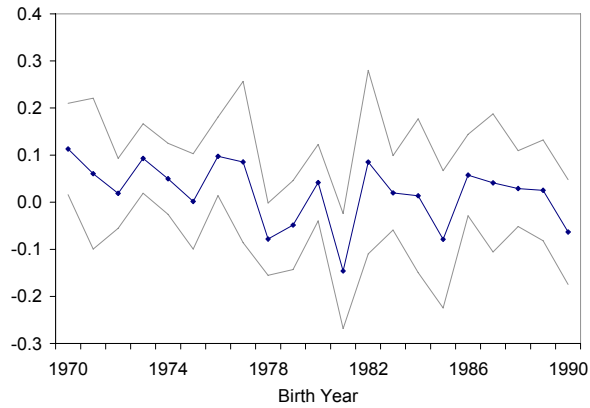
**Figure 5A: The Effect of Relaxation on Sex Ratios of First Born Children and 95% Confidence Intervals**  
 Coefficients of the Interactions between  
 Born in Relaxed Region \* Birth Year



**Figure 5B: The Effect of Relaxation on Sex Ratios of Second Born Children and 95% Confidence Intervals**  
 Coefficients of the Interactions between  
 Born in Relaxed Region \* Birth Year



**Figure 5C: The Effect of Relaxation on Sex Ratios of Later Born Children and 95% Confidence Intervals**  
 Coefficients of the Interactions between  
 Born in Relaxed Region \* Birth Year



**Table 1A: Descriptive Statistics**  
CHNS 1989 and 1% Sample of China Population Census

Variable	Obs	Mean	Std. Err.	Obs	Mean	Std. Err.
	<b>A. Girls</b>			<b>B. Boys</b>		
Han	11938	0.943	(0.002)	14518	0.949	(0.002)
# Siblings	11938	0.908	(0.007)	14518	0.759	(0.006)
Sisters	11938	0.359	(0.005)	14518	0.413	(0.005)
Brothers	11938	0.550	(0.006)	14518	0.345	(0.004)
Enrolled	11938	0.504	(0.005)	14518	0.477	(0.004)
Mother's Education	11551	6.252	(0.040)	13944	5.805	(0.036)
Father's Education	10872	8.191	(0.036)	13305	7.729	(0.034)
Mother at Home	11938	0.116	(0.003)	14518	0.135	(0.003)
Relaxation	11938	0.243	(0.003)	14518	0.238	(0.003)
Agricultural	11927	0.597	(0.004)	14481	0.608	(0.004)
	<b>C. &lt;=3 Children</b>			<b>D. Only Child</b>		
Sex	16723	0.512	(0.004)	9733	0.611	(0.005)
Han	16723	0.939	(0.002)	9733	0.958	(0.002)
Enrolled	16723	0.430	(0.004)	9733	0.591	(0.005)
Mother's Education	16231	5.469	(0.031)	9264	6.952	(0.048)
Father's Education	15427	7.600	(0.029)	8750	8.530	(0.044)
Mother at Home	16723	0.129	(0.003)	9733	0.121	(0.003)
Relaxation	16723	0.272	(0.003)	9733	0.186	(0.003)
Dist to Prov. Capital	16723	169.968	(1.147)	9733	134.971	(1.607)
Dist to Big City	15806	7.711	(0.081)	9369	10.667	(0.116)
Agricultural	16694	0.709	(0.004)	9714	0.423	(0.005)

Sample of cohorts born 1962-1981

**Table 1B: Descriptive Statistics for Counties with Relaxation**

Variable	No Relaxation			Some Relaxation		
	Obs	Mean	Std. Err.	Obs	Mean	Std. Err.
<b>A. Demographic</b>						
Sex	9915	0.555	(0.005)	16541	0.545	(0.004)
Han	9915	0.967	(0.002)	16541	0.934	(0.002)
# Siblings	9915	0.833	(0.007)	16541	0.822	(0.006)
Sisters	9915	0.385	(0.006)	16541	0.391	(0.004)
Brothers	9915	0.448	(0.006)	16541	0.431	(0.004)
Enrolled	9915	0.459	(0.005)	16541	0.507	(0.004)
Mother's Education	9558	5.153	(0.042)	15937	6.520	(0.034)
Father's Education	9055	7.518	(0.038)	15122	8.187	(0.032)
Mother at Home	9915	0.110	(0.003)	16541	0.136	(0.003)
Agricultural	9905	0.700	(0.005)	16503	0.546	(0.004)
<b>B. Infrastructural</b>						
Relaxation	9915	0.000	0.000	16541	0.384	(0.003)
Dist to Prov Capital	9915	178.135	(1.340)	16541	144.480	(1.264)
Dist to Big City	8634	2.106	(0.017)	16541	12.311	(0.091)
Dist to Primary School	9915	0.245	(0.007)	15281	0.420	(0.004)
Dist to Middle School	9914	1.014	(0.010)	15281	1.607	(0.011)
Dist to High School	9914	4.989	(0.089)	15281	4.470	(0.069)

Sample of households with <=3 children amongst cohorts born 1962-1981.

**Table 2: OLS and 2SLS Estimates of the Effect of Family Size on School Enrollment**  
Coefficients of the number of siblings

	<b>Dependent Variable: School Enrollment</b>					
	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	All	All	All	All
<b>OLS</b>						
<b>A. Households with &lt;=3 children</b>						
<b>Sample Mean of Dep Var</b>			<b>0.489</b>			
# Siblings	-0.011 (0.004)	-0.011 (0.004)	-0.012 (0.005)	-0.010 (0.003)	-0.011 (0.004)	-0.011 (0.004)
Observations	26456	26456	26456	25495	26456	25495
R-squared	0.69	0.69	0.69	0.70	0.69	0.70
<b>B. Households with &lt;=2 children</b>						
<b>Sample Mean of Dep Var</b>			<b>0.539</b>			
# Siblings	0.016 (0.005)	0.016 (0.005)	0.013 (0.004)	0.018 (0.004)	0.016 (0.005)	0.015 (0.005)
Observations	21321	21321	21321	20497	21321	20497
R-squared	0.70	0.70	0.70	0.71	0.70	0.71
<b>2SLS</b>						
<b>C. Households with &lt;=3 children</b>						
# Siblings	0.161 (0.058)	0.161 (0.059)	0.167 (0.062)	0.146 (0.043)	0.159 (0.062)	0.140 (0.046)
Observations	26456	26456	26456	25495	26456	25495
R-squared	0.64	0.64	0.64	0.66	0.64	0.66
<b>D. Households with &lt;=2 children</b>						
# Siblings	0.122 (0.065)	0.122 (0.065)	0.134 (0.064)	0.097 (0.063)	0.125 (0.067)	0.104 (0.062)
Observations	21321	21321	21321	20497	21321	20497
R-squared	0.69	0.69	0.69	0.70	0.69	0.70
<b>Controls</b>						
Han	N	Y	Y	N	N	Y
Distance to Urban	N	N	Y	N	N	Y
Mother's Education	N	N	N	Y	N	Y
Household Income	N	N	N	N	Y	Y

All regressions control for the full set of interaction terms and birth year and county fixed effects. Standard errors are clustered at the county level.

**Table 3: The Effect of Relaxation on Family Size and School Enrollment by Size**  
Coefficients for columns (1)-(4) are the interaction terms between born in a relaxed region \* year of birth.  
Coefficients in columns (5)-(6) are the triple interactions between dummy for girl \* born in a relaxed region \* year of birth.

	Dependent Variables						
	# Sibs		Enrollment		# Siblings	Enrollment	
	(1)	(2)	(3)	(4)	(5)	(6)	
	Girls	Boys	Girls	Boys	All	All	
<b>Sample Mean of Dependent Variable</b>	<b>1.153</b>	<b>0.922</b>	<b>0.473</b>	<b>0.456</b>	<b>Sample Mean of Dependent Variable</b>	<b>1.028</b>	<b>0.464</b>
relax*born 1973	0.026 (0.110)	0.034 (0.116)	-0.180 (0.082)	-0.092 (0.071)	relax*girl*born 1973	-0.020 (0.099)	-0.087 (0.037)
relax*born 1974	0.127 (0.115)	0.082 (0.107)	-0.216 (0.098)	-0.050 (0.078)	relax*girl*born 1974	0.048 (0.073)	-0.168 (0.070)
relax*born 1975	0.115 (0.071)	0.045 (0.139)	-0.112 (0.078)	-0.106 (0.048)	relax*girl*born 1975	0.068 (0.132)	-0.009 (0.056)
relax*born 1976	0.155 (0.128)	0.137 (0.157)	-0.087 (0.062)	-0.048 (0.030)	relax*girl*born 1976	0.028 (0.170)	-0.038 (0.074)
relax*born 1977	0.325 (0.136)	0.096 (0.101)	0.007 (0.055)	-0.058 (0.037)	relax*girl*born 1977	0.246 (0.116)	0.062 (0.061)
relax*born 1978	0.274 (0.152)	0.076 (0.161)	0.032 (0.028)	-0.011 (0.027)	relax*girl*born 1978	0.205 (0.171)	0.042 (0.022)
relax*born 1979	0.418 (0.158)	0.222 (0.159)	0.010 (0.034)	-0.035 (0.025)	relax*girl*born 1979	0.210 (0.183)	0.044 (0.022)
relax*born 1980	0.385 (0.180)	0.153 (0.128)	0.030 (0.033)	-0.022 (0.028)	relax*girl*born 1980	0.247 (0.168)	0.048 (0.018)
relax*born 1981	0.450 (0.186)	0.088 (0.154)	0.005 (0.035)	-0.025 (0.030)	relax*girl*born 1981	0.370 (0.194)	0.029 (0.014)
Observations	13271	15500	13271	15500	Observations	28771	28771
R-squared	0.26	0.23	0.70	0.69	R-squared	0.25	0.69

Regressions in columns (1)-(4) include county and birth year fixed effects.

Regressions in columns (5)-(6) include controls for relax\*girl, relax\*birthyear, girl\*birthyear, girl, birthyear fixed effects and county fixed effects.

Standard errors clustered at county level.

**Table 4: The Effect of Relaxation on Sex Ratios by Birth Parity**  
Coefficients of the interaction terms between born in a relaxed region\*birth year.

	Dependent Variable: Dummy for Male											
	1st Born				2nd Born				3+ Born			
	(1)		(2)		(3)		(4)		(5)		(6)	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
<b>Sample Mean of Dep Var</b>	<b>0.530</b>				<b>0.509</b>				<b>0.495</b>			
Relax * Born 1970	0.065	(0.032)	0.064	(0.031)	0.113	(0.047)	0.110	(0.046)	0.113	(0.049)	0.112	(0.049)
Relax * Born 1971	0.067	(0.018)	0.066	(0.018)	0.126	(0.055)	0.124	(0.054)	0.061	(0.082)	0.055	(0.081)
Relax * Born 1972	0.084	(0.042)	0.086	(0.043)	0.168	(0.071)	0.170	(0.071)	0.019	(0.038)	0.018	(0.037)
Relax * Born 1973	0.053	(0.048)	0.056	(0.049)	0.177	(0.078)	0.178	(0.077)	0.093	(0.038)	0.091	(0.037)
Relax * Born 1974	0.106	(0.032)	0.103	(0.033)	0.108	(0.070)	0.111	(0.069)	0.050	(0.039)	0.047	(0.039)
Relax * Born 1975	0.027	(0.047)	0.025	(0.049)	0.257	(0.046)	0.256	(0.047)	0.002	(0.052)	0.002	(0.052)
Relax * Born 1976	0.045	(0.046)	0.050	(0.047)	0.062	(0.054)	0.066	(0.055)	0.097	(0.043)	0.094	(0.044)
Relax * Born 1977	0.059	(0.044)	0.057	(0.045)	0.230	(0.063)	0.225	(0.063)	0.085	(0.087)	0.081	(0.087)
Relax * Born 1978	0.063	(0.051)	0.056	(0.052)	0.126	(0.071)	0.122	(0.071)	0.078	(0.039)	0.080	(0.039)
Relax * Born 1979	0.106	(0.035)	0.101	(0.036)	0.154	(0.052)	0.154	(0.051)	0.048	(0.048)	0.050	(0.049)
Relax * Born 1980	0.163	(0.044)	0.162	(0.044)	0.128	(0.038)	0.126	(0.039)	0.042	(0.041)	0.041	(0.041)
Relax * Born 1981	0.148	(0.050)	0.145	(0.051)	0.161	(0.063)	0.169	(0.064)	0.146	(0.062)	0.144	(0.063)
Relax * Born 1982	0.078	(0.050)	0.077	(0.050)	0.200	(0.065)	0.200	(0.065)	0.085	(0.099)	0.073	(0.098)
Relax * Born 1983	0.079	(0.033)	0.081	(0.034)	0.121	(0.106)	0.120	(0.105)	0.020	(0.040)	0.007	(0.037)
Relax * Born 1984	0.040	(0.046)	0.038	(0.046)	0.244	(0.098)	0.241	(0.098)	0.014	(0.083)	0.003	(0.083)
Relax * Born 1985	0.070	(0.039)	0.066	(0.039)	0.110	(0.049)	0.114	(0.048)	0.079	(0.074)	0.079	(0.076)
Relax * Born 1986	0.086	(0.048)	0.089	(0.048)	0.145	(0.091)	0.158	(0.092)	0.058	(0.044)	0.047	(0.049)
Relax * Born 1987	0.098	(0.053)	0.097	(0.053)	0.144	(0.064)	0.145	(0.062)	0.041	(0.075)	0.038	(0.075)
Relax * Born 1988	0.044	(0.036)	0.038	(0.036)	0.105	(0.056)	0.108	(0.057)	0.029	(0.041)	0.028	(0.041)
Relax * Born 1989	0.067	(0.033)	0.075	(0.033)	0.167	(0.043)	0.167	(0.044)	0.025	(0.055)	0.021	(0.055)
Relax * Born 1990	0.079	(0.036)	0.077	(0.037)	0.113	(0.096)	0.118	(0.095)	0.063	(0.057)	0.064	(0.057)
Han	N		Y		N		Y		N		Y	
Observations	44754		44754		23306		23306		14495		14495	
R-squared	0.01		0.01		0.01		0.01		0.01		0.01	

Regressions include county and birthyear fixed effects.

Standard errors clustered at the county level.

**Table 5: The Effect of Relaxation on Sex Ratios for First Born Children**

Coefficients are the interaction terms between born in a relaxed region \* born in the One Child Policy Cohort and between born in a relaxed region \* born in the relaxation cohort

<b>The Effect of Relaxation on Sex Ratios for First Borns</b>		
<b>Dependent Variable: Dummy for Male</b>		
	(1)	(2)
Born in relaxed region * Born 1976-1981	0.106 (0.024)	0.098 (0.025)
Born in relaxed region * Born 1982-1989	0.037 (0.021)	0.035 (0.021)
Han, Han * Birth Cohort	N	Y
Observations	44234	44234
R-squared	0.00	0.00

Regressions include county and birth cohort fixed effects.

Standard errors clustered at the county level.

**Table 6: OLS and 2SLS Estimates of the Effect of Family Size on School Enrollment by Sex Composition of Siblings**

	<b>Dependent Variable: Fraction enrolled in School</b>			
	(1)	(2)	(3)	(4)
	Different Sex	Same Sex	Younger Sister	Younger Brother
<b>OLS</b>				
<b>A. Households with &lt;=3 children</b>				
<b>Sample Mean of Dep Var</b>	<b>0.511</b>	<b>0.419</b>	<b>0.534</b>	<b>0.536</b>
# Siblings	-0.053 (0.011)	-0.011 (0.005)	-0.009 (0.004)	-0.005 (0.004)
Observations	6218	19277	17785	18601
R-squared	.55	0.6	0.73	0.72
<b>B. Households with &lt;=2 children</b>				
<b>Sample Mean of Dep Var</b>	<b>.483</b>	<b>0.554</b>	<b>0.378</b>	<b>0.423</b>
# Siblings	0.000 (0.000)	0.005 (0.006)	0.018 (0.005)	0.015 (0.006)
Observations	3907	16590	14979	16213
R-squared	0.55	0.6	0.73	0.72
<b>2SLS</b>				
<b>C. Households with &lt;=3 children</b>				
# Siblings	0.058 (0.225)	0.176 (0.082)	0.190 (0.049)	0.049 (0.048)
Observations	6383	20073	18514	19316
R-squared	0.57	0.67	0.65	0.71
<b>D. Households with &lt;=2 children</b>				
# Siblings	0.000 (0.000)	0.112 (0.058)	0.125 (0.063)	0.078 (0.052)
Observations	4018	17303	15641	16857
R-squared	0.58	0.72	0.72	0.71
<b>Controls</b>				
Han	Y	Y	Y	Y
Distance to Urban	Y	Y	Y	Y
Mother's Education	Y	Y	Y	Y
Household Income	Y	Y	Y	Y

All regressions control for the full set of interaction terms and birth year and county fixed effects. Standard errors are clustered at the county level. Column (1) excludes individuals whose next youngest sibling is of the same sex. Column (2) excludes individuals whose next youngest sibling is of different sex. Column (3) is restricted to the sample individuals who do not have a younger brother. Column (4) is restricted to the sample individuals who do not have a younger sister.



**Table 7: OLS and 2SLS Estimates of the Effect of Family Size on Education Delay**  
Coefficients for the number of siblings a first born child has

	<b>Dependent Variable: Actual Years of Education - Supposed Years of Education</b>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	All	All	All	All	Younger Sister	Younger Brother
<b>Households with &lt;=3 children</b>								
<b>A. OLS</b>								
Sample Mean of Dep Var			<b>0.489</b>				<b>0.534</b>	<b>0.536</b>
# Siblings	-0.121 (0.026)	-0.121 (0.026)	-0.127 (0.029)	-0.076 (0.021)	-0.125 (0.021)	-0.095 (0.024)	-0.121 (0.027)	-0.115 (0.026)
Observations	12940	12940	12940	12715	12940	12715	9817	10306
R-squared	0.63	0.63	0.63	0.64	0.63	0.65	0.68	0.68
<b>2SLS</b>								
<b>B. 2SLS</b>								
# Siblings	0.297 (0.440)	0.299 (0.439)	0.274 (0.438)	0.303 (0.469)	0.274 (0.433)	0.213 (0.454)	0.179 (0.436)	0.504 (0.606)
Observations	12940	12940	12940	12715	12940	12715	10008	10499
R-squared	0.61	0.61	0.62	0.63	0.62	0.65	0.65	0.63
Controls								
Han	N	Y	Y	N	N	Y	Y	Y
Distance to Urban	N	N	Y	N	N	Y	Y	Y
Mother's Education	N	N	N	Y	N	Y	Y	Y
Household Income	N	N	N	N	Y	Y	Y	Y

Sample restricted to individuals enrolled in school.

All regressions control for the full set of interaction terms and birth year and county fixed effects.

Standard errors are clustered at the county level.

Column (7) is restricted to the sample individuals who do not have a younger brother.

Column (8) is restricted to the sample individuals who do not have a younger sister.

**Table 8: The Effect of Family Size on Female Labor Supply**

	<b>Dependent Variable: Dummy Variable for Mother who Stays at Home</b>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	All	All	All	All	All	Younger Sister	Younger Brother
<b>Households with &lt;=3 Children</b>								
<b>Sample Mean of Dep. Var</b>			<b>0.126</b>				<b>0.124</b>	<b>0.123</b>
<b>A. OLS</b>								
<b># Siblings</b>	-0.029 (0.012)	-0.029 (0.012)	-0.029 (0.012)	-0.036 (0.013)	-0.028 (0.012)	-0.036 (0.013)	-0.036 (0.013)	-0.038 (0.013)
<b>Observations</b>	26456	26456	26456	25495	26456	25495	17785	18601
<b>R-squared</b>	0.13	0.13	0.13	0.19	0.13	0.19	0.20	0.20
<b>B. 2SLS</b>								
<b># Siblings</b>	-0.142 (0.098)	-0.144 (0.098)	-0.144 (0.098)	-0.122 (0.088)	-0.139 (0.098)	-0.119 (0.088)	-0.137 (0.055)	-0.141 (0.079)
<b>Observations</b>	26456	26456	26456	25495	26456	25495	18514	19316
<b>R-squared</b>	0.08	0.08	0.08	0.16	0.09	0.17	0.09	0.10
Controls								
Han	N	Y	Y	N	N	Y	Y	Y
Distance to Urban	N	N	Y	N	N	Y	Y	Y
Mother's Education	N	N	N	Y	N	Y	Y	Y
Household Income	N	N	N	N	Y	Y	Y	Y

All regressions control for the full set of interaction terms and birth year and county fixed effects.

Standard errors are clustered at the county level.

Column (7) is restricted to the sample individuals who do not have a younger brother.

Column (8) is restricted to the sample individuals who do not have a younger sister.

## Appendix

**Table A1: The Correlation between Family Size and School Enrollment**  
Coefficients of dummy variables for the total number of children in each household

	Dependent Variables							
	School Enrollment				Actual Years of Edu - Supposed Years of Edu			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Sample Mean for Dependent Variable</b>		<b>0.464</b>				<b>-2.470</b>		
Total Kids = 2	-0.095 (0.051)	-0.098 (0.050)	0.004 (0.004)	0.018 (0.006)	-0.700 (0.115)	-0.657 (0.140)	-0.156 (0.033)	0.010 (0.042)
Total Kids = 3	-0.308 (0.057)	-0.312 (0.057)	-0.056 (0.011)	-0.023 (0.009)	-1.312 (0.168)	-1.297 (0.172)	-0.581 (0.099)	-0.247 (0.062)
Total Kids = 4	-0.397 (0.058)	-0.409 (0.059)	-0.071 (0.011)	-0.028 (0.009)	-1.462 (0.175)	-1.449 (0.203)	-0.794 (0.126)	-0.334 (0.074)
Total Kids = 5	-0.464 (0.055)	-0.487 (0.056)	-0.070 (0.018)	-0.019 (0.015)	-2.192 (0.290)	-2.219 (0.362)	-1.254 (0.198)	-0.778 (0.163)
Total Kids = 6	-0.540 (0.051)	-0.573 (0.056)	-0.075 (0.024)	-0.017 (0.020)	-2.015 (0.310)	-1.668 (0.359)	-0.472 (0.285)	0.328 (0.197)
Total Kids = 7	-0.507 (0.061)	-0.534 (0.063)	-0.048 (0.047)	0.004 (0.046)	-2.515 (0.364)	-2.784 (0.325)	-0.881 (0.450)	-0.823 (0.202)
Total Kids = 8	-0.591 (0.039)	-0.655 (0.048)	-0.033 (0.006)	0.040 (0.016)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Total Kids = 9	-0.591 (0.039)	-0.644 (0.048)	-0.025 (0.006)	0.038 (0.020)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Constant	0.591 (0.039)	0.594 (0.034)	0.478 (0.015)	-0.001 (0.016)	2.015 (0.075)	1.995 (0.080)	1.680 (0.066)	-5.029 (0.724)
Controls								
County Fixed Effect	N	Y	N	Y	N	Y	N	Y
Birth Year Fixed Effect	N	N	Y	Y	N	N	Y	Y
Observations	28771	28771	28771	28771	13338	13338	13338	13338
R-squared	0.08	0.08	0.71	0.72	0.07	0.08	0.60	0.62

All standard errors clustered at county level.

Columns (5)-(8) restricted to sample of children enrolled in school.

**Table A2: The Effect of Relaxation on Family Size and School Enrollment by Size for Households with 3 or fewer Children**

Coefficients for columns (1)-(4) are the interaction terms between born in a relaxed region \* year of birth.

Coefficients in columns (5)-(6) are the triple interactions between dummy for girl \* born in a relaxed region \* year of birth.

	Dependent Variables					
	# Siblings		School Enrollment		# Siblings	School Enrollment
	(1) Girls	(2) Boys	(3) Girls	(4) Boys		
<b>Sample Mean of Dependent Variable</b>	<b>0.908</b>	<b>0.759</b>	<b>0.500</b>	<b>0.480</b>	<b>0.826</b>	<b>0.489</b>
Relax* Born 1973	0.006 (0.080)	0.074 (0.099)	0.209 (0.079)	0.106 (0.074)	Relax*Girl*Born 1973 0.083 (0.086)	0.101 (0.041)
Relax* Born 1974	0.225 (0.070)	0.105 (0.060)	0.230 (0.103)	0.035 (0.078)	Relax*Girl*Born 1974 0.129 (0.067)	0.194 (0.076)
Relax* Born 1975	0.211 (0.087)	0.148 (0.111)	0.100 (0.063)	0.124 (0.054)	Relax*Girl*Born 1975 0.063 (0.059)	0.023 (0.037)
Relax* Born 1976	0.317 (0.070)	0.158 (0.133)	0.074 (0.060)	0.040 (0.033)	Relax*Girl*Born 1976 0.179 (0.104)	0.033 (0.079)
Relax* Born 1977	0.509 (0.088)	0.242 (0.119)	0.015 (0.043)	0.060 (0.038)	Relax*Girl*Born 1977 0.287 (0.105)	0.074 (0.050)
Relax* Born 1978	0.493 (0.117)	0.145 (0.151)	0.032 (0.029)	0.015 (0.027)	Relax*Girl*Born 1978 0.361 (0.133)	0.045 (0.023)
Relax* Born 1979	0.571 (0.117)	0.275 (0.162)	0.016 (0.034)	0.040 (0.024)	Relax*Girl*Born 1979 0.315 (0.157)	0.056 (0.019)
Relax* Born 1980	0.525 (0.153)	0.260 (0.154)	0.026 (0.034)	0.026 (0.026)	Relax*Girl*Born 1980 0.268 (0.176)	0.049 (0.019)
Relax* Born 1981	0.551 (0.135)	0.168 (0.174)	0.003 (0.035)	0.028 (0.028)	Relax*Girl*Born 1981 0.411 (0.174)	0.030 (0.016)
					Relax*Girl*Born 1982	
Observations	11938	14518	11938	14518	26456	26456
R-squared	0.18	0.20	0.70	0.68	0.19	0.69

Regressions in columns (1)-(4) include county and birth year fixed effects.

Regressions in columns (5)-(6) include controls for relax\*girl, relax\*birthyear, girl\*birthyear, girl, birthyear fixed effects and county fixed effects.

Standard errors clustered at county level.

