Empirically Probing the Quantity-Quality Model

Emla Fitzsimons and Bansi Malde

June 2013

Abstract This paper estimates the causal effects of family size on girls' education in Mexico, exploiting prenatal son preference as a source of random variation in the propensity to have more children within an Instrumental Variables framework. It finds no evidence of family size having an adverse effect on education. The paper then weakens the identification assumption and allows for the possibility that the instrument is invalid. It finds that the effects of family size on girls' schooling remain extremely modest at most. Families that are relatively large compensate for reduced per child resources by increasing maternal labour supply.

Keywords Fertility · Education · Instrumental Variables · Latin America

JEL classification I20 · J13· J16

E. Fitzsimons (✉) · B. Malde
Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE, United Kingdom
email: emla_f@ifs.org.uk   tel +44 (0)20 7291 4800   fax +44 (0)20 7323 4785

B. Malde
University College London, Gower Street, London WC1E 6BT

1 Introduction

This paper investigates the effect of family size on schooling in a Latin American country - Mexico. Policymakers in developing countries, including a number in Latin America, have often advocated policies promoting smaller families as a way of improving human capital accumulation and economic development. Though the quantity-quality model suggests that this type of policy is likely to be effective - since as quantity (number of children) rises, the total cost of quality (investment into children) also rises, thus decreasing the demand for quality (Becker 1960; Becker and Lewis 1973; Becker and Tomes 1976) - other fields such as psychology suggest that large families may be advantageous for children's human capital due to the potentially beneficial effects of children on each other's development (Zajonc 1976). Further, in developing countries, some siblings may bring resources and thus contribute to the household budget to the benefit of other siblings, or households may adjust on margins such as mother's labour supply, leading to an ambiguous effect of family size on children's schooling. The issue is largely an empirical one, with important implications for policymakers in deciding whether policies to reduce family size are likely to be an effective way of increasing parental investment in children's human capital thus improving their long-run productivity, facilitating economic growth, and reducing the intergenerational transmission of poverty and economic inequality - issues which are arguably even more acute in a developing country context. This paper estimates the causal impacts of fertility on children's education in a developing country, thus providing important new evidence for policymakers, and in a context where evidence remains scarce.[1]

The most widely used approaches to identify the causal effects of family size on children's education use same sex composition and/or twin births as instruments for family size and so require very large samples, which until recently have been scarce in developing countries. Further, with the exception of Lee (2008) for Korea, Ponczek and Souza (2012) for Brazil, the existing work on developing countries pertains to China, and findings are contradictory and difficult to extrapolate to other contexts given China's one child policy (Li et al. 2008; Rosenzweig and Zhang 2009; Qian 2009).[2] Our paper contributes to this gap in the literature

[1] There is an abundant literature showing that parents with large families invest less in children's education than parents with small families, but much of this evidence is non-causal. Schultz (2008) provides a review.
[2] Li et al. (2008) and Rosenzweig and Zhang (2009) find evidence consistent with the quantity-quality model, whilst Qian (2009) finds a positive effect of an additional child on school enrolment. Other than these studies, work that estimates the effects of family size on children's education generally relates to developed countries

by providing evidence on how family size affects girls' education in the rural population of a large Latin American country, where fertility remains high. The source of exogenous variation in family size exploited is parental preferences for having at least one son. We find no evidence that family size has an adverse effect on girls' accumulated stock of education: the observed negative correlation between family size and education disappears when we allow for the endogeneity of family size. This is a robust finding, which is true across different family size margins and different measures of the stock of education. We find evidence however that families are adjusting on another margin, with mothers increasing labour supply in response to having more children.

What remains contentious throughout this literature is the extent to which findings are an artefact of instrument invalidity. This is evident from two recent papers: Rosenzweig and Zhang (2009) find that differential birth endowments of twins are important for education choices; they also find evidence of economies of scale with respect to gender sameness, and suggest that these could be driving the findings commonly found in the literature. Angrist et al. (2010) on the other hand find no evidence invalidating the identifying restrictions in an Israeli context.[3] Very few other studies directly examine the extent to which concerns about instrument validity underlie findings. In this paper on the other hand, we investigate the extent to which our findings are driven by instrument invalidity. We first show that the particular concerns about validity (son preferences and economies of scale) are not important from an empirical viewpoint in our context. Thereafter, the paper allows for the possibility that the instrument is indeed imperfect, using the methods recently developed by Nevo and Rosen (2012). It shows that even if the instrument is invalid, the qualitative findings are not affected much: the effects of family size on children's outcomes remain extremely modest.

The data used in this paper span over half a million relatively poor households in marginalized communities in rural Mexico, allowing us to test the effect at different margins of increase in family size, and for children of different birth orders. Indeed, this is one of the few studies to consider family size increases above 2 to 3. These higher margins are arguably the more important ones to consider for developing countries: the average family size in the

---

and generally shows no or only very weak evidence of a quantity-quality trade-off (Black et al. 2005; Cáceres-Delpiano 2006; Conley and Glauber 2006 – all for the US; Angrist et al. 2010 for Israel).

[3] Angrist and Evans (1998) also defend the validity of the same-sex instrument for the US; Rosenzweig and Wolpin (2000) on the other hand find evidence of economies of scale in India.

Mexican sample used here is just under 4. Moreover to the best of our knowledge, this is the first study to test the quantity-quality model in Mexico, thus providing evidence from a new country to add to the growing body of studies. Such replication of IV estimates on new data sets has indeed been stressed by Angrist (2004) as a crucial component in establishing the external validity of IV estimates.

The paper proceeds as follows. Section 2 sets out the methodology for estimating the effects of family size on children's school outcomes. In section 3, the data used in the analysis are described, alongside some descriptive statistics. The main body of the paper is contained in section 4, where the results are shown. Section 5 contains robustness tests and a discussion of findings, and the paper concludes in section 6.

2 Methodology

The basic model to be estimated is the following

$$Y_{ij}=\beta_{0j}+\beta_{1j}X+\beta_{2j}F_{ij}+u_{ij} \tag{1}$$

where the outcome variables, Y_{ij}, pertain to the education of child i at birth parity j and include a 0-1 indicator of current enrolment in school, years of schooling, a 0-1 indicator for completed primary schooling, and a 0-1 indicator for completed lower secondary schooling; X is a vector of covariates including child age (dummy variables), a quadratic in maternal age, maternal years of schooling, household characteristics including asset ownership as captured by an asset index[4], a land ownership dummy, an agricultural household dummy, a dummy indicating the presence of children other than siblings in the household, year and state dummies and a range of village characteristics measuring available infrastructure and public services; F_{ij} is family size of child i; and the error term u_{ij} denotes unobserved factors that affect Y_{ij} and that may be correlated with F_{ij}.[5] This equation is estimated separately for parities j=1,...,3, using pooled cross-sectional data from 1996 through 1999, covering the entire population of rural indigent communities in Mexico (detailed in section 3).

[4] The asset index is computed by aggregating indicators for whether or not a household owns 10 assets including blender, fridge, gas stove and radio, among others.
[5] We use a linear specification in this paper, given that the instrumental variables are binary.

Estimating equation (1) by ordinary least squares (OLS) would render the coefficient of interest, β_{2j}, biased and inconsistent if omitted variables, such as parental preferences, influence both children's outcomes and family size. To obtain a consistent estimate of β_{2j}, an instrumental variable method is used, which requires the existence of a variable, Z that is correlated with F_{ij} but uncorrelated with u_{ij}. In a first-stage regression, we estimate

$$F_{ij} = \alpha_{0j} + \alpha_{1j}Z + \alpha_{3j}X + \xi_{ij} \quad (2)$$

The main source of exogenous variation in family size used is all-female births.[6] Our population exhibits strong prenatal son preferences: that the first n births are female is highly correlated with further childbearing[7]; the relationship between all-male births and further childbearing is considerably weaker however, as we show later on in section 3.2.3.[8]

The instrument is effectively the sex of the n^{th} child in households in which the first n-1 births are female[9]: we expect, and later show, family size to be higher in households where the n^{th} birth is also female. We consider the outcomes of the first n-1 children, all female by definition.[10] We do this for n=2...4. As n increases we can consider outcomes of higher birth parities, so when n=2 we consider the outcomes of first-borns; for n=3, first- and second-borns; for n=4, first-, second- and third-borns. We first consider effects separately by birth parity and instrument. So for instance for female first-borns, we construct three analysis samples: those in families with n≥2 (instrument=female at 2^{nd} birth; 'ff'); those in families

---

[6] Sex composition was first used as an instrument for family size by Angrist and Evans (1998) and has since been applied by others such as Angrist et al. (2010) and Conley and Glauber (2006). These studies use same-sex births as the instrument, whether all-male or all-female; Lee (2008) on the other hand uses all-female births. Another commonly used instrument is twin births. Rosenzweig and Zhang (2009) highlight a number of concerns underlying the validity of this instrument (including differential birth endowments and birth intervals of twins versus singletons). The most likely direction of ensuing bias of the IV estimates is positive (Behrman et al. 1994; Rosenzweig and Zhang 2009): we conducted an analysis using twin births as an instrument and found some positive IV estimates, leaving us with concerns that these issues may indeed be relevant in our context but are unfortunately not possible to investigate further given the available data. Finally, a third type of instrument exploits variation in implementation and enforcement of fertility policies as an instrument. Qian (2009) and Wu and Li (2012) use variation in the implementation and enforcement of the Chinese One-Child Policy to identify causal effects of family size on children's education and maternal health respectively.

[7] Similar correlations have been found in contexts with son-biased fertility preferences. See for example Rosenblum (2013) for India.

[8] In using only all-female births as our instrument, the reader may be concerned that we are using just one part of the variation induced by sex composition preferences. For completeness, Table A1 in the Appendix reports results from the analysis using sex composition (either same sex births or all-male and all-female births) as an instrument for family size. As can be seen from the Table, the all-female instrument has considerably more power in the first stage and results are primarily driven by it.

[9] We condition on the first n-1 births being female as the instrument is preference for at least one son.

[10] One reason for this is that children of the n^{th} birth may be of different sexes; another reason is to avoid any selection bias arising from families that have children after a male birth being different from those that do not.

with n≥3 where the first two are female (instrument=female at 3rd birth, 'fff'), and those in families with n≥4 where the first three are female (instrument=female at 4th birth, 'ffff'). Creating these subsamples allows us to estimate effects along different margins of increase in family size. We next pool all the parity- (and instrument-) specific subsamples in order to improve precision.[11]

One common criticism of this methodology is the issue of instrument validity. We devote section 5 to this important issue. We first provide evidence relating to its validity in our context. This evidence is reassuring, but to address lingering concerns, we impose weaker assumptions on the instrument and allow for the exclusion restriction to be violated (Nevo and Rosen 2012). This allows us, for the first time in this literature, to derive informative bounds of the effect of family size on outcomes. Therefore, we can directly answer the question of how much the assumption of instrument exogeneity drives the results.

Finally, we note that in the presence of heterogeneous effects, the parameter identified is a local average treatment effect (LATE), the effect of increased family size on education for households whose treatment status is manipulated by the instrumental variable. Hence, for the all-female instrument, we identify the effect of increasing family size on education for the sub-population of households with n females that go on to have an additional child solely because they wish to have a boy. This sub-population is called the compliers (Angrist et al. 1996). In section 4, we first decompose the first stages to understand better the range of variation in family size induced by the instruments, before describing the characteristics of the compliers in order to understand better just how representative our findings are for the population in our survey as a whole.

3 Data and Descriptive Statistics

3.1 The Data

The data used in this paper are cross-sectional socio-economic data that were collected across marginalized rural areas throughout 31 states in Mexico between 1996 and 1999.[12] Our

[11] Though the importance of birth order for education choices has been highlighted in the literature (Black et al. 2010, 2011; Rosenzweig and Zhang 2009), as we will see, we find little evidence of heterogeneity in the effects of family size by birth order in the sample considered here.
[12] Most localities were chosen on the basis of having been graded with a high degree of marginalisation on the basis of the 1995 Census data.

sample comprises particularly poor households, as the descriptive statistics will show later on. The survey - the Survey of Household Socio-Economic Characteristics (*Encuesta de Caracteristicas Socioeconomicas de los Hogares, ENCASEH*) - was conducted in order to aid in the targeting of the PROGRESA (now *Oportunidades*) welfare program, introduced in selected marginalized rural villages across 7 states in 1998, and later expanded to cover the whole country. The survey collected data from all households in these communities and contains a rich cross-section of information on individual and household characteristics, along with village data. Moreover, being a census of the rural parts of all states in Mexico, the sample sizes are extremely large, which is very advantageous for the research here as it facilitates an analysis using different margins of increase in family size and different birth parities.

The analysis is restricted to 12-17 year olds, as school enrolment before age 12 - at just over 97% - is practically universal. We retain children living in the same household as their mother, regardless of the mother's marital/cohabiting status.[13] This is in order to avoid sample selection bias, as cohabitation status is likely to be a function of the instrument, as suggested by Dahl and Moretti (2008).[14] In 90% of cases the mother is married or cohabiting; in 4% she is divorced, in 4% she is widowed, and in the remaining 2% she is single or her status is unknown. We take a family to be the mother and all children born to her (in practically all cases (99.7%) there is one family per household). So family size is the number of children recorded as having the same mother. A potential concern is that we may miscode family size if older children have left the household permanently and are thus not part of the survey. Reassuringly, only 2.1% of households report that a household member left the household permanently in the past 5 years. Finally, note that we drop households in which the eldest sibling is 18 or above and thus beyond school age.[15] Our final sample is extremely large, containing just over 630,000 families across approximately 1,500 villages.

[13] This is the vast majority (92%) of those aged less than 18. For the remaining 8%, it is the case that the mother is deceased or in another household.

[14] To check whether this is the case, we follow Dahl and Moretti (2008) and test whether sex composition affects the probability of maternal divorce and parental cohabitation. We find a very small statistically significant positive (~0.1%) correlation between all-female composition and maternal divorce (relative to an all-male composition) and a small statistically significant negative correlation with parental cohabitation.

[15] Though we could potentially retain them in the sample when we consider the outcomes of second- and third-borns, a reason for not doing so is that we have some concerns about coding birth orders for households with children above age 18. Note that we also drop households that reported more than one household head (0.03%) and households (1.5%) with suspect data, mainly the reporting of implausible ages.

3.2 Descriptive Statistics

3.2.1 Our Sample

We first show some characteristics of the sample in Table 1. The average family size is just under 4. Around 50% of households have children of the same sex in the first two births: just under half of these have two females. The mothers in our sample are 38 years old on average and have just over 3 years of schooling. Less than 30% have completed primary schooling or above. As mentioned already, a majority of mothers (~90%) are married or cohabiting, 4% are widowed and divorce is low at 4%. Agricultural work is widespread, with three quarters of households engaged in it. Indicators of poverty such as the quality of the roof of the dwelling and the availability of a toilet and running water, confirm that the households are quite poor.

[Insert Table 1 here]

3.2.2 Measures of Schooling

The objective of this study is to estimate the causal impact of family size on the accumulation of one form of human capital: education. To measure this we use school enrolment at the time of the survey and three different measures of the stock of education: years of schooling, completion of primary schooling and completion of lower secondary schooling.[16] The stock variables are our preferred outcome measures, as they embody past investments in education and are thus a cleaner measure of educational attainment and accumulation: school enrolment on the other hand relates to a one-off decision and does not necessarily capture accumulation of education. Moreover, enrolment in school is relatively less costly, both in terms of time and other inputs, than is completion of schooling levels. As the stock variables more closely reflect investments in human capital (in terms of time and money), they are the more relevant outcomes for testing the quantity-quality model. They are also more relevant for policymakers in a context where most children complete primary schooling but just under one third complete lower secondary schooling. This is despite the fact that compulsory basic education (grades 1–9, covering 6 years of primary and 3 of lower secondary) in Mexico is free of charge and publicly provided.[17] Completion of levels is also of interest in the presence of

[16] These latter 2 levels are ones that children of our age range should have achieved (for instance, Mexican children would complete lower secondary school by age 14 if they started primary school at age 6 and progressed through without repeating any grades). Note also that all of these outcomes are measured at a particular point in time between ages 12 and 17 and are thus not necessarily indicators of completed schooling.
[17] At the basic education level, participation in private education in Mexico is low, at 10%, and is not relevant for the poor population considered here.

'sheepskin effects' in the returns to schooling - where there are returns from obtaining a qualification conditional on completed years of schooling.

The following two figures depict these measures for both males and females. They show that educational attainment is fairly equal between males and females: though school enrolment is slightly higher for males after the age of 12, these differences are very low (see Figure 1). Moreover by age 17 they have converged. Nor do any of the three measures of the stock of education display any stark differences between the sexes: if anything, females are engaged more in education according to these measures. Though the focus of the paper is on females, the male-female comparison highlights the similarity in their education, which, though not ruling out the possibility that son preferences affect intra-household allocation choices once a child is born, suggests that they do not.[18] As we will see in section 5, this is reassuring from the point of view of the validity of the instrument.

Figure 1 also shows a sharp drop in school enrolment at age 12, before which it is practically universal. We thus consider outcomes from age 12 onwards only. The figure also shows that years of schooling are increasing with age, though not one-to-one.

[Insert Figure 1 here]

Figure 2 displays primary school completion and lower secondary school completion for 12-17 year olds (both of which are free and publicly provided).[19] By age 12, the age at which a child should have completed primary schooling, less than 40% of children has done so, and less than 80% of males and females have completed primary schooling by age 17. For lower secondary schooling, less than 10% of those who should - those aged 14 and above - have completed lower secondary schooling, and this proportion stands at just under 40% by age 17.

[Insert Figure 2 here]

[18] Schultz (2004) and Behrman et al. (2005) document lower secondary school enrolment amongst girls than boys in the communities comprising the sample for PROGRESA, justifying the premium for girls in the subsidy. However it should be noted that there is a sizeable literature attributing any differences between the sexes to availability of schools/distance to schools/marriage markets rather than preferences for boys' schooling *per se*.

[19] Though there are no fees for public schools, direct costs of schooling include purchasing textbooks, stationary, school uniforms; and transportation to and from school. Note also that the opportunity cost of schooling is increasing with age, which may explain the observed patterns.

3.2.3 Son Preferences

As has been mentioned already, the main source of exogenous variation in family size exploited in this paper is parental preferences for having at least one son. Here we provide more concrete evidence of the presence of son preferences in our sample, by estimating how family size responds to a succession of all-male or all-female births (a modified version of equation (2), at the family level), separately for families with a succession of all-male or all-female births. Table 2 shows the coefficient estimates for a succession of all-female and all-male births (left and right hand panels respectively). It shows in particular that around 1 in 10 families go on to have an additional child following a succession of all-female births; the corresponding figure for males is between 1 in 20 and 1 in 100. This provides fairly stark evidence that the relationship between all-male births and family size is considerably weaker than that between all-female births and fertility, and is consistent with Dahl and Moretti (2008), who find a similar pattern for Mexico.[20]

[Insert Table 2 here]

3.2.4 Are the instruments randomly assigned?

The IV methodology, in the presence of heterogeneous effects of family size, requires that the instrument is random conditional on observed covariates. The randomisation assumption could be violated if parents choose the sex of their children (via sex-selective abortions). Sex selection must be a concern in areas where cultural norms value male children over female children. We believe that this issue is unlikely to arise in our sample however: Mexico is a predominantly Catholic country where abortion is legally restricted throughout the period of our data; moreover access to the technology for determining sex is likely to be very low for our population of very poor rural households, so aborting on the basis of gender is unlikely to be an issue. We are reassured that the sex ratio at birth for Mexico is at its usual norm of 1.05 male/female, as an imbalance in this would indicate gender-biased preferences (Bhaskar

[20] Whilst some of the existing literature pools males and females, we eschew from doing this (in line with Ponczek and Souza 2012). We have two reasons for this. First, we have just shown that there are strong son preferences in the population, where an all-female composition is more than twice as likely (and in one case almost 10 times as likely!) to induce families to continue their child-bearing compared to an all-male composition (Table 2). Second, pooling restricts the causal effect of family size on education to be the same for both genders. It is however well known that, particularly in developing country contexts, due to economic reasons such as higher costs of sending females to school and/or lower returns for females in the labour market (e.g. Airola and Juhn 2005; Attanasio and Binelli 2010), or social norms supporting preferences for sons (e.g. Deaton and Subramaniam 1996; Oster 2009; Chakravarty 2010), the gender of a child plays an important role in parents' human capital investment decisions and processes.

2011).[21] Furthermore, Table 3 compares characteristics of mothers whose first n-1 births are females, and who have either a female or a male at the n^{th} birth (separately by panel, for n=2,...,4). Though in a handful of cases we observe statistical differences, these are extremely small in magnitude. Overall, the Table supports the randomness of sibling sex.

[Insert Table 3 here]

However, the evidence presented in Table 3 is on the basis of observed characteristics only. To provide further evidence regarding prenatal sex selection, we test whether the birth interval preceding a male birth is longer than that preceding a female birth: if prenatal sex selection is an issue, we would expect the birth interval before a son to exceed the interval before a daughter, as the son's birth is more likely to be preceded by abortion of female foetuses, thereby increasing the time until his birth. To see if this is the case, we estimate the following regression following Ebenstein (2008), on the sample of families where the first n-1 births are female, for n=2,...,4.

$$E(B_{n,h} \mid f_{n,h}, X, f_{1,h},..,.f_{n-1,h}=1) = \gamma_1 f_{n,h} + \gamma_2 X + \varepsilon_h \quad (3)$$

where $B_{n,h}$ is interval of time in years preceding the n^{th} birth in family h, $f_{n,h}$ is an indicator for a female at the n^{th} birth, and X is a vector of family and village characteristics as per equation (1). As discussed above, in the presence of prenatal sex selection, we would expect γ_1 to be negative: the birth interval before a female would be lower than that before a male. Estimates from this regression are displayed in Table 4 and provide further evidence that prenatal sex selection is not occurring in our sample: females are not more likely to be preceded by shorter birth intervals than males.

[Insert Table 4 here]

4 Results

In this section we show the main results of the paper, first displaying estimates by birth parity and instrument for the four measures of education outlined in section 3.2.2, and thereafter

[21] The sex ratio at birth in Mexico has historically stood around this level: it was 1.01 between 1990 and 1994 (Parazzini et al. 1998), and the Mexican Demographic and Health Surveys suggest that the sex ratio of all children ever born stood at 1.03 in 1987 (Arnold 1992).

showing estimates when we pool birth parities. We show the first stage results, along with OLS (LPM) and IV estimates.[22]

4.1 By Birth Parity

Tables 5 and 7 display, respectively, the results for first-borns and for second- and third-borns, using different instruments depending on the parity being considered, as explained in section 2. The top panel of each table shows the first-stage coefficients for the different instruments, while the bottom panels display the OLS and IV estimates for 4 different measures of education: school enrolment, years of schooling, primary school completion and lower secondary school completion. Note that in all that follows, standard errors are clustered at the village level in order to account for spatial autocorrelation within the village.

[Insert Table 5 here]

Considering the results for first-borns, we see from the top panel of Table 5 that the all-female instruments are all very strong, as is evident from the F-tests. Their magnitude is such that they increase family size by an average of 0.1 children - that is 1 in 10 first-born females gain an additional sibling due to the instrument.

We decompose this overall proportion of compliers to obtain more insight into the ranges of variation in family size induced by each instrument. This is displayed graphically in Figure 3.[23] The horizontal axis displays completed family size[24]; the vertical axis shows the proportion of families that has that family size because the instrument is switched on, and that would not otherwise have continued its fertility. We see from the Figure that just over 2% of the sample is induced to go on to have 3 children because ff=1, around 3.5% of the sample is induced to go on to have 4 children, and so on, with statistically significant fertility increases occurring up to 7 children (beyond which increases are no longer statistically different from zero, as can be seen from the 95% confidence intervals around the estimates). More generally, the fertility increases induced by the instruments are high, reaching 8 children for the ffff instrument, implying that the all-female instruments capture the effects of a family size of up

[22] Three of our outcome variables, school enrolment, primary school completion and lower secondary completion, are binary: thus we use LPM (linear probability model) estimation in this case. For convenience we use the term OLS throughout the text.
[23] This follows Angrist and Imbens (1995).
[24] By ‘completed family size’, we mean completed as at the time of the survey.

to 8 children.[25] So the effects of family size that we go on to estimate are a weighted average over a wide range of family sizes, a range that contains margins relevant for the population we consider (where the average number of children per family is just under 4).

We also investigate, in Table 6, the characteristics of these compliers to understand the types of family for whom our findings will be applicable. Whilst compliers are not an identifiable subpopulation, we can, in the spirit of Angrist and Imbens (1995), describe them in relation to the general population in terms of observed characteristics. For instance, the relative likelihood that a complier household has a highly educated mother, compared to the overall sample, is given by the ratio of the first stage for highly educated mothers to the overall first stage. The characteristics considered include maternal age, education and marital status, household head occupation and measures of family wealth including dummy variables for asset ownership. Overall, compliers are relatively better off than the population in our survey at large: they include mothers from considerably more educated backgrounds compared to the general population, and are more likely to own the listed assets.[26]

[Insert Figure 3 here]

[Table 6 here]

Looking at the IV estimates of the effect of family size on the education of first-born females, we see that regardless of outcome or instrument, the OLS estimates are negative and significantly different from 0, with an additional child associated with a reduction of 2 percentage points in school enrolment, 0.1 years reduction in completed years of schooling, a 1.4 percentage point reduction in primary school completion, and an approximately 2 percentage point decline in the probability of completing lower secondary school. These magnitudes are in line with those of Angrist et al. (2010) for Israel. When we instrument for family size, we find that coefficients are generally small and statistically indistinguishable from 0, with mixed signs. Moreover differences between the OLS and IV estimates are

[25] Whilst we do not explicitly consider non-linear effects of family size in this paper (Mogstad and Wiswall 2010), our use of different instruments affecting different margins of increase in family size allows us to see whether there is any evidence of non-linearities in the effects of family size on children's education.

[26] The figures in the columns give the relative likelihood that compliers have the characteristic listed in column (A). For instance a figure of 0.75 means that the population of ff compliers is ¾ as likely to have a non-qualified father compared to the overall population.

typically not statistically significant, raising concerns that the IV estimates are not estimated precisely enough, despite the strong first stages.

Turning to the results for second- and third-borns, we see from the top panel of Table 7 that, as with first-borns, the first stages are very strong, with magnitudes ranging from 0.124 to 0.159 on average. The OLS estimates are, similarly, all negative with magnitudes similar to those for first-borns across the four outcomes. When we account for the endogeneity of family size, no consistent pattern emerges. All IV estimates are statistically indistinguishable from 0, with mixed signs across outcomes and instruments. However, as with first-borns, we lose substantial precision in the IV estimates (despite our very large samples) such that the majority of differences between the OLS and IV estimates are not statistically distinguishable from one another.

[Insert Table 7 here]

In order to boost precision in the estimation, we pool the different instrument and birth parity subsamples, as described in the following sub-section.

4.2 Pooling Parities

We follow Angrist et al. (2010) and pool the different instrument and parity specific samples and estimate the effects on this pooled sample. Specifically, we pool first-born females from households with at least 2 children, second-borns from households with at least 3 children (where the first 2 are female) and third-borns from households with at least 4 children (where the first 3 are female). The first stage equation is as follows:

$$F_i = \alpha_0 + \alpha_1 Z + \alpha_3 X + \xi_i \tag{4}$$

where F_i is family size of child i, and X variables are those listed in section 2 (with the addition of birth parity dummies), and where the instrument, Z, is effectively that the n^{th} born is female in households in which the first n-1 births are female: so an all-female sex composition. Whilst there are gains in precision from pooling birth parities, it imposes the assumption that the relationship between the birth of a subsequent female and family size is the same regardless of birth parity. However this is consistent with the first stage estimates shown in Tables 5 and 7, which are very similar across parities. In the second stage, we estimate

$$Y_i = \beta_0 + \beta_1 X + \beta_2 F_i + u_i \quad (5)$$

As can be seen, pooling parities also restricts the effects of family size to be the same for different birth parities. This is consistent with Tables 5 and 7, where the coefficient estimates for the three birth orders considered are statistically indistinguishable from one another. Further support for the plausibility of this assumption is provided in Figure 4, which plots the relationship between education measures and family size, separately by birth parity. As is evident from the Figure, the relationship between an additional child and schooling outcomes is similar across all three birth parities (with just minor differences at higher family sizes), which provides further justification for pooling the three birth parities.

[Insert Figure 4 here]

Table 8 shows estimates from the specification where the three birth parities are pooled: first-born females in households with at least 2 children, second-born females in households with at least 3 children of which the first 2 are female, and third-born females in households with at least 4 children of which the first 3 are female. For any particular birth parity, the instrument is the sex of the *subsequent* birth: so the instrument for first-born females is that the second-born is female; for second-born females it is that the third-born is female, and for third-born females it is that the fourth born is female. As discussed above, this method assumes that the relationship between family size and the birth of a subsequent female is homogeneous across birth parities.

[Insert Table 8 here]

We see from Table 8 that this method improves precision considerably. For all of the 3 stock measures of schooling considered, the IV estimates are statistically different from their OLS counterparts and generate no evidence of an adverse effect of larger family size on years of schooling, primary school completion or lower secondary school completion. For the flow measure, school enrolment, the IV estimate is still not precise enough to be able to reject that it is different from the OLS estimate. However, we reiterate that this is a weak proxy for parental investment into children's education, and for this reason not our preferred outcome measure.[27]

Interestingly, our findings are very much in line with those for developed countries, including Cáceres-Delpiano (2006) for the US, Black et al. (2005) for Norway and Angrist et al. (2010) for Israel. There is considerably less evidence on the quantity-quality tradeoff in developing countries to compare our findings to, with the exception of a recent study in Brazil (Ponczek and Souza 2012) which, using twin births as an instrument for family size, points towards a quantity-quality tradeoff. However, this finding is mainly driven by children with relatively low educated mothers, whereas our compliers are more likely to be children with relatively better educated mothers, and thus our findings are not all that comparable. There are also some studies on China (Li et al. 2008; Qian 2009; Rosenzweig and Zhang 2009) though this is a very different environment with strict fertility restrictions and not comparable to ours. It remains the case that more evidence is needed before general conclusions can be drawn.

5 Robustness

The previous section showed that when the endogeneity of family size is taken into account, there is no evidence of an adverse effect of larger families on the educational attainment and accumulation of females. We now go on to probe this conclusion further. First, we investigate to what extent the findings are an artefact of invalid instruments, rather than picking up the effects of family size *per se*. Second, we investigate whether families are adjusting on margins other than children's education, in particular mother's labour supply.

A key concern throughout this literature, and indeed throughout the literature relating to estimation using instrumental variables more generally, concerns the validity of instruments. It is posited in particular that sex composition may affect education directly through economies of scale, which are difficult to control for. Yet despite its importance for inference,

[27] We also experimented with pooling parities by instrument instead. So for instance, for the fff instrument, we estimated a specification in which we pooled first- and second-borns in families with at least three children in which the first two are female. Whilst an advantage of this alternative is that we need not impose the assumption that the causal impact of family size on education is the same across all instrument-specific subsamples (and so across all margins of increase in family size), we fail to boost precision sufficiently by pooling in this manner. These results are available on request.

more often than not, instrument validity is not directly addressed.[28] In this paper, we first provide direct evidence on the likely validity of the instrument in our context. Though the evidence we show is reassuring, instrument validity cannot of course ever be established with certainty. We take a new approach in this paper by testing directly the robustness of findings to weaker identification assumptions, allowing explicitly for the instruments to be correlated with the error term in the outcome equation, using methods developed by Nevo and Rosen (2012). With these weaker assumptions on the instrument, we can estimate bounds on the magnitude of the effects of family size. Thus for the first time in this literature, we can show to what extent instrument invalidity matters for inference. Finally, we consider alternative channels on which families may be adjusting in response to increased family size.

5.1.1 Evidence on instrument validity

As has been discussed, the exclusion restriction is that the sex of the n^{th} born has no direct effect on the education of the outcome child. There are at least two concerns with this. The first is that son preferences may directly affect the education of females in the household. The second is economies of scale in all-female households, arising from children of the same sex being able to share more items (such as clothes and shoes). In both cases, the direction of the resulting bias of the IV estimate is positive: if postnatal son preferences exist (and affect education decisions), then a sister is more beneficial for girls' schooling than a brother; if scale economies are important, savings may be higher in all-female households (relative to mixed gender households), which is also beneficial for schooling.[29]

Concerning son preferences, Lee (2008) points out that the instrument concerns prenatal and not postnatal son preferences, in other words that parents prefer to have sons rather than daughters *ex-ante*, and not that parents treat sons more favourably than daughters *ex-post*.

[28] Exceptions include Rosenzweig and Wolpin (2000) and Rosenzweig and Zhang (2009), who provide direct evidence on the likely validity of same-sex and twin instruments respectively. Angrist et al. (2010) address the issue mainly by comparing twins and sex-composition estimates, as the omitted variables bias associated with each type of instrument should act differently.

[29] Though these are the most commonly cited concerns in the literature, we acknowledge that other unobserved factors might result in a negative correlation between the instrument and the error term in the structural equation, thereby biasing downward the IV estimates. For instance, the quality of marriage might be lower in all-female households, which may be adverse for children's schooling (see for instance Brown and Flinn 2011); another example concerns the role of socialization at home, which differs depending on sibling composition - girls with brothers tend to have more 'masculine' traits perhaps because brothers encourage girls to be more assertive and outspoken (Koch 1955). So it is plausible to expect that, if assertiveness is associated with better success at school, girls with sisters will do less well in school (see Butcher and Case 1994).

However if postnatal son preferences exist, the sibship gender composition may affect intra-household schooling choices. The concern is whether the subgroup of families whose fertility decisions are affected by the all-female instrument exhibit postnatal son preferences, something which we cannot test. The fact that education outcomes for males and females are very similar (see Figures 1 and 2), which conforms to recent trends in Mexico showing convergence in education between the sexes, is at least encouraging in this regard.[30]

A potentially more important concern, and one that has received much attention in the literature, is economies of scale from same-sex births resulting in savings which may trickle through to education choices (Rosenzweig and Wolpin 2000; Rosenzweig and Zhang 2009). We argue here that cultural customs are so different from western industrialized countries that the scope for economies of scale is much more limited. Clothes tend to be unisex, especially at young ages, so traditional hand-me-downs that can generate economies of scale are unlikely to be gender-specific. School books are also likely to be common to both sexes given the predominance of mixed-sex schools in our setting. Another remark worth making is that the sharing of gender-specific goods is unlikely to be restricted to within the household, but to take place right across the extended family and social network.[31] It is difficult to think of items other than these that offer the potential for economies of scale arising from sex composition.

To provide more concrete evidence, we use detailed data on expenditures on children's clothes and shoes reported in the PROGRESA evaluation sample, which was drawn from our population (ENCASEH), to test more directly for evidence of scale economies in families with a same sex composition.[32] The data we use, which cover seven states in rural Mexico in 1998/99, suggest that such scale economies are unlikely to be important.[33] First, an extensive

[30] Recent UNESCO statistics for Mexico show that 98% of girls and 98% of boys are in primary school; 74% of girls and 71% of boys are in secondary school (UNESCO 2011); evidence from Parker and Pederzini (2000) shows that the gender gap in education in Mexico has fallen substantially over the last 30 years, to the extent that females and males below the age of 20 no longer display significant differences in educational attainment, as measured by years of schooling. Duryea et al. (2007) analyze the educational gender gap in Latin America and the Caribbean and find that the most striking differences are across income groups and not gender.

[31] Angelucci et al. (2010, 2012) document the importance of extended family networks for this population in making schooling choices (particularly in response to the PROGRESA grant) and in providing support following adverse events.

[32] The PROGRESA evaluation sample includes ~24,000 households in 506 villages, who were interviewed on 8 occasions over the period 1997-2007. We are unable to match households in the PROGRESA evaluation sample to those in our sample as different household identifiers are used in the two datasets.

[33] Data on clothing are not available for our main sample, ENCASEH. We pool post-program data from surveys in October 1998 and May 1999, from control villages only, to ensure the analysis is not contaminated by any

proportion, 80%, of the family budget is spent on food - the corresponding figure in western economies over the same period is less than 20% (UK and US National Accounts Data). More relevant still, the purchase of children's clothes and shoes is fairly infrequent: just 65% of families had purchased these items over the previous 6 month period; amongst those that had purchased some, expenditure accounts for just 4% of monthly non-durable consumption, leaving very little scope for scale economies in these goods. In what follows, we pool expenditures on children's clothes and shoes and refer to both together as clothing.

We estimate the following two equations to test whether there is evidence of the sex composition of children affecting the family's purchase of children's clothing:

$$D_h = \lambda_0 + \lambda_1 Z + \lambda_2 X + \xi_h \tag{6}$$

$$M_h = \vartheta_0 + \vartheta_1 Z + \vartheta_2 X + \upsilon_h \tag{7}$$

where D_h in equation (6) is a dummy variable equal to 1 if the family reports purchasing children's clothing in the previous 6 months and 0 otherwise, and M_h in equation (7) is family peso expenditure on children's clothing. X is a vector of control variables including household demographics, maternal age and education, family size, village size and distance to the nearest town. Note that controlling for family size is necessary as per-child expenditure on clothing is, unfortunately, not observed. To mitigate ensuing concerns that households that chosen to have the same family size despite having different gender compositions may differ in unobserved ways that also affect their purchases of children, we condition on families for whom the correlation between same-sex composition and family size is relatively low - those with a succession of male births - and so for whom such biases will be much less important. Therefore, the variable of interest, Z, is a dummy variable equal to 1 if the n^{th} child is male and 0 if it is female, in families where the first n-1 births are male.[34]

Equations (6) and (7) are estimated using probit and tobit models respectively. Estimates are shown in the upper panel of Table 9. We find no evidence that, in families with n-1 males, the n^{th}-born being male as opposed to female results in lower purchases of children's clothing,

potential program effects. We retain families where the first-born child is below 18 years old - not just 12-17 years of age as in main analysis - to boost sample sizes. Compared to our main sample, families here have fewer children on average; parents are also on average younger and more educated.

[34] The sample pools families with at least two children where the first is a male, families with at least three children where the first two are male, and families with at least four children where the first three are male.

either at the extensive or intensive margin. Taken together, these pieces of evidence suggest that scale economies arising from same sex compositions are unlikely to be driving our findings.

[Insert Table 9 here]

This evidence suggests that the threats to the validity of the all-female instrument are not very serious in this context.[35] Still, this evidence alone does not (and could not) establish validity of the instrument. We next allow for the instrument to be imperfect and under weaker identification assumptions, derive bounds on the effects of family size on education.

5.1.2 Bounds

Whilst this evidence is reassuring, doubts often linger as to whether the instrument satisfies validity, in other words, whether it is uncorrelated with the structural error term: $E[Zu_i] = 0$ (see equations (5) and (6)). In this section we ask the question: if the validity assumption fails, that is if $E[Zu_i] \neq 0$, can we learn anything about the parameters of interest? To answer this, we use the methods developed by Nevo and Rosen (2012), who relax the validity condition to allow for some correlation between the instrument and the structural error term (detailed below), and derive (set) identification results for the parameters of interest. This allows us to derive informative bounds for the causal effects of family size on education. This is a potentially very useful approach in this literature to directly investigate the extent to which weakening the identification assumption affects findings.

Rather than assuming that the instrument is uncorrelated with the error term, we assume, in line with Nevo and Rosen (2012) that

1. The correlations between the endogenous regressor and the structural error term, and between the instrument and the structural error term, have the same sign:

$$\rho_{Fu}\rho_{Zu} \geq 0 \tag{8}$$

[35] As further reassuring evidence, we re-emphasize that there is no relation between all-female births and any of the covariates in our model – see Table 3. Another salient point is that whilst we cannot control for savings in our data, results are robust to the inclusion or exclusion of proxies for resources (mother's education, household assets, home and land ownership).

As discussed in section 5.1.1, the most likely direction of correlation between the instrument and the error term in the structural equation is positive. We know, however, that the most likely correlation between the endogenous variable (F) and the error term is negative (Becker 1960; Becker and Lewis 1973). So to satisfy this assumption, we simply specify the endogenous variable as –F.

Nevo and Rosen (2012) show that if assumption 1 holds and, as in our case, the covariance between the endogenous variable (re-specified as –F) and the instrument is negative, then we have a two-sided bound given by

$$B^* = [\beta^{OLS}, \beta_Z^{IV}] \tag{9}$$

The important thing to note is that because the correlation between the instrument (all-females) and the endogenous regressor (-F) is negative, we obtain two-sided bounds on the parameter of interest.[36] So this assumption can provide finite, economically informative bounds on the parameter of interest.

If, as in Nevo and Rosen (2012), we make the additional assumption that

2. The correlation between the instruments and the structural error term is less strong in absolute terms than the correlation between the endogenous regressor and the structural error term

$$|\rho_{Zu}| \leq |\rho_{Fu}| \tag{10}$$

which considerably weakens the usual validity assumption for instrumental variables which requires that the correlation between the instrument and the structural error term is zero, then we obtain tighter bounds. We believe it is reasonable to expect the all-female instrument to be less correlated with the error term in the structural equation, than family size. Indeed, the whole premise of the quantity-quality model is that both the number of children (family size) and quality per child (here, education) are jointly chosen by parents, which means that they are both affected by unobservable parental preferences: these preferences are absorbed into the error term of the structural equation. Whilst the all-female instrument may matter for education in ways not controlled for (such as economies of scale), we believe that these

[36] If it is positive, we only obtain one-sided bounds.

effects are small relative to the unobserved factors associated with family size, such as preferences for education.

So if, as we believe is reasonable in our case, assumptions 1 and 2 hold, we obtain a two-sided bound given by:

$$B^{**} = [\beta_{Z*}^{IV}, \beta_{Z}^{IV}] \tag{11}$$

where $\boldsymbol{\beta_{Z*}^{IV}}$ is a TSLS estimator in which $\boldsymbol{Z^{*}} = \boldsymbol{\sigma_Z F - \sigma_F Z}$, which is constructed in such a manner as to be exogenous, serves as an instrument for the endogenous variable, and where $\boldsymbol{\sigma_Z}$ and $\boldsymbol{\sigma_F}$ are the standard deviations of F and Z respectively. Essentially, the key implication of assumption 2 is that $\boldsymbol{\beta_{Z*}^{IV}}$ improves on the lower bound given by $\boldsymbol{\beta^{OLS}}$.[37]

The confidence interval for the set of bounds $[\beta_{Z*}^{IV}, \beta_{Z}^{IV}]$ is formed as

$$CI_{\alpha} = \left[\hat{\beta}_{Z*}^{IV} - c_{\alpha}\frac{\widehat{\sigma}_{\hat{\beta}_{Z*}^{IV}}}{\sqrt{n}}, \hat{\beta}_{Z}^{IV} + c_{\alpha}\frac{\widehat{\sigma}_{\hat{\beta}_{Z}^{IV}}}{\sqrt{n}}\right] \tag{12}$$

where $\widehat{\sigma}_{\hat{\beta}_{Z*}^{IV}}$ $\left(\widehat{\sigma}_{\hat{\beta}_{Z}^{IV}}\right)$ is a standard error for β_{Z*}^{IV} (β_{Z}^{IV}) and where c_{α} is chosen as 1.96 for a 95% confidence interval (Stoye 2009).

[Insert Table 10 here]

Table 11 shows the bounds of the effect of family size on education (parity-pooled sample).[38] They yield some informative insights. Focusing on years of schooling and primary school completion, they suggest that even if the instrument is invalid, this does not affect findings much. For instance, the coefficient for years of schooling is between -0.098 and 0.021 (with a confidence interval of -0.113 and 0.141). This conclusion holds for lower secondary school completion as well, where the magnitudes of the effects remain very modest. We note from the Table that the upper bound in all cases is the IV estimate, the case where the instrument is uncorrelated with the error term. The lower bound is tighter than the OLS estimate, due to assumption 2 above (though the magnitude of the improvement is low). In the absence of assumption 2, the lower bound would simply coincide with the OLS estimate. We note from

[37] This is clear from Corollary 1 of Nevo and Rosen (2012), which implies that, in our case, $\beta_{Z*}^{IV} > \beta^{OLS}$. Moreover, the Corollary shows that the larger the correlation between the instrument and the endogenous regressor, the greater the improvement of β_{Z*}^{IV} over β^{OLS} and thus the tighter the lower bound.

[38] We use the parity-pooled sample given the considerable gains in precision as discussed in Section 4.2.

the confidence intervals that across all outcomes, we cannot reject that the point estimate is zero. These estimates are very useful for policy making: even if the identification strategy is flawed, inferences remain the same and we detect no evidence of important effects of family size on children’s education. This is a conclusion similar to the one reached by Rosenzweig and Zhang (2009).

5.1.3 Discussion

We have found little evidence in this paper that family size affects the stock of education of females: we investigate the extent to which families may be adjusting on margins other than children’s education. One that has been commonly looked at in the literature is female labour supply (for instance, Rosenzweig and Wolpin 1980; Angrist and Evans 1998; Agüero and Marks 2008). We here investigate the extent to which mothers increase labour supply if they have more children. The definition of labour supply we consider is wage work, the most reliable measure available in the survey. Around 10 per cent of mothers in our sample report working for a wage.

We see from the OLS estimates in Table 11 that, in line with previous work, mothers with large families work less than those with small families. However, the IV estimates show the opposite: mothers with large families are significantly more likely to work. This evidence, though limited, suggests that families may indeed be adjusting on other margins in an attempt to protect their children’s education. A more complete look at this would also consider other margins of adjustment such as health investments, found to be important by Millimet and Wang (2011), though beyond the scope of this present study.[39]

[Insert Table 11 here]

6 Conclusion

This paper considers the effect of family size on girls’ schooling across a population of relatively poor households in rural Mexico. It accounts for the endogeneity of family size

[39] Another possibility - and one raised by Angrist et al. (2010) - is that households use public funds to smooth the shock to fertility. The main candidate in our context is the PROGRESA program, providing mainly subsidies for school attendance. However the data used for the analysis in this paper relate to the period *before* PROGRESA was introduced (indeed our data were collected in order to identify households eligible for the subsidy - see section 3.1).

using a succession of female-only births as a source of exogenous variation in family size. The paper exploits extremely large samples and high fertility rates to consider the effects of family size on a range of different education outcomes. We find no evidence of family size having a detrimental effect on girls' educational accumulation, though there is evidence to suggest that families may be adjusting on other margins – mother's labour supply – to protect their children's education. Moreover, there remains the possibility that households use other mechanisms to smooth the shock to fertility, such as public funds.

A divisive issue in this literature relates to the validity of the instruments. Various threats to instrument validity have been raised by different authors and evidence on their empirical importance remains mixed. We have taken a new approach to tackling this issue, allowing for the instruments to be imperfect and have estimated bounds on the effects, along the lines of Nevo and Rosen (2012). This is a new and potentially very useful approach in this literature to directly answer the question of how much the assumption of instrument exogeneity drives the results. We find that the bounds on the effect identified by the instruments are informative. Moreover, OLS estimates, which are generally very modest in magnitude, are shown to provide a lower bound of the effect of family size on education. This indicates that the effect of family size on education is very modest at most.

One explanation behind these findings may be that households choose to adjust on margins other than children's education. We investigate one possible channel - mother's labour supply - and find evidence to suggest that this may indeed be happening, with mothers engaging more in work in large than in small families. Other margins could be health investments, investigation of which is unfortunately outside the scope of this study, though an important agenda for future work.

Acknowledgements We are grateful to conference participants at the European Economic Association Annual Conference, Milan 2008, the Latin American and Caribbean Economic Meetings, Medellin 2010, and the Royal Economic Society Conference, London 2011; also to Orazio Attanasio, Valérie Lechene, Marta Mier y Terán, David Phillips, Nancy Qian, Adam Rosen, Marta Rubio-Codina and Marcos Vera-Hernández for useful discussions, to two anonymous referees for useful suggestions and to Marta Mier y Terán for providing access to the data. All errors are the responsibility of the authors. We are grateful to the Economic and Social Research Council for funding this work through research grant RES-000-22-2877.

Table 1: Characteristics of Households

Variable	Mean	Std. Dev.
Family size	3.816	1.842
Proportion of households with 1st 2 births of the same sex	0.505	0.500
Proportion of households with 1st 2 births female (ff)	0.235	0.424
Proportion of households with 1st 3 births female (fff)	0.119	0.324
Proportion of households with 1st 4 births female (ffff)	0.062	0.242
Socio-Economic Variables		
Mother's age	38.082	7.465
Mother's years of schooling	3.262	2.891
Mother has no schooling	0.269	0.444
Mother has at least completed primary schooling	0.280	0.449
Mother is married	0.896	0.306
Mother is divorced	0.044	0.204
Mother is widowed	0.042	0.201
Mother is single	0.017	0.129
Indigenous language speakers	0.338	0.473
Household owns dwelling	0.916	0.277
Wall materials of dwelling (0 = poor quality)	0.873	0.333
Roof materials of dwelling (0 = poor quality)	0.409	0.492
Water supply in dwelling	0.243	0.429
Electricity in dwelling	0.794	0.404
Number of rooms in dwelling	1.922	1.224
Household has own toilet	0.628	0.483
Household has water in toilet	0.199	0.399
Household owns land	0.502	0.500
Household head works in agriculture	0.756	0.429
N	636,438	

Sample of families with at least one 12-17 year old, in which the eldest child is <age 18. 'Sibling' refers to children born to the same mother.

Table 2: Son Preferences

	Dependent variable = family size					
	s=female			s=male		
	n=2	n=3	n=4	n=2	n=3	n=4
	[1]	[2]	[3]	[4]	[5]	[6]
n^{th} birth = s	0.110**	0.105**	0.118**	0.016**	0.053**	0.043**
	[0.005]	[0.008]	[0.012]	[0.005]	[0.007]	[0.011]
Observations	259,131	107,928	39,838	289,562	122,271	44,717
F Test	417.30	155.60	101.80	9.58	56.96	14.68
Sample	2+, f=1	3+, ff=1	4+, fff=1	2+,m=1	3+,mm=1	4+,mmm=1

All regressions include controls for mother's age and education, family and village characteristics and state dummies. Note also that we condition implicitly on the sex composition of the first n-1 births by restricting the sample to the first n-1 births being all-female. Standard errors clustered at the village level in parentheses. Sample in column 1(4) includes the families with at least 2 children, with a first born female(male) aged < 18 years, sample in columns 2(5) includes families with at least 3 children and where the eldest 2 are females(males) aged < 18 years, and sample in columns 3(6) includes families with at least 4 children, where the eldest 3 are females(males) and aged < 18 years. * Significant at 5%, ** Significant at 1%

Table 3: Sample Balance

Variable	fm=1	ff=1	Difference in means	p-value
Mother's age	36.869	36.845	-0.024	0.293
Mother's age at first birth	22.309	22.300	-0.010	0.647
Mother's years of schooling	3.405	3.407	0.002	0.868
Mother has no schooling	0.248	0.251	0.003	0.116
Mother has at least completed primary school	0.297	0.298	0.001	0.634
Mother is married	0.912	0.909	-0.003	0.039*
Mother is divorced	0.039	0.041	0.002	0.035*
Father is present in household	0.884	0.879	-0.004	0.001**
Birth spacing b/w 1st and 2nd births	2.926	2.939	0.014	0.066
Family size	4.082	4.187	0.105	0.000**
N	131,360	128,896		
	ffm=1	**fff=1**		
Mother's age	36.033	36.006	-0.027	0.477
Mother's age at first birth	21.471	21.457	-0.014	0.690
Mother's years of schooling	3.402	3.395	-0.007	0.683
Mother has no schooling	0.248	0.249	0.001	0.604
Mother has at least completed primary school	0.296	0.297	0.000	0.923
Mother is married	0.927	0.924	-0.003	0.036*
Mother is divorced	0.033	0.033	0.001	0.626
Father is present in household	0.899	0.896	-0.003	0.032*
Birth spacing b/w 1st and 2nd births	2.630	2.657	0.027	0.004**
Birth spacing b/w 2nd and 3rd births	3.043	3.042	0.000	0.976
Family size	4.547	4.654	0.107	0.000**
N	54,557	53,841		
	fffm=1	**ffff=1**		
Mother's age	35.549	35.538	-0.010	0.841
Mother's age at first birth	20.960	20.905	-0.055	0.253
Mother's years of schooling	3.226	3.209	-0.017	0.552
Mother has no schooling	0.262	0.264	0.002	0.597
Mother has at least completed primary school	0.274	0.272	-0.002	0.668
Birth spacing b/w 1st and 2nd births	2.372	2.385	0.012	0.390
Birth spacing b/w 2nd and 3rd births	2.631	2.636	0.005	0.740
Birth spacing b/w 3rd and 4th births	2.900	2.887	-0.013	0.443
Family size	5.191	5.319	0.128	0.000*
N	21,158	21,304		

N refers to the number of first-born female children. fm=1 indicates female at 1st birth, male at 2nd birth; ff=1 indicates female at 1st 2 births, and so on. Birth orders coded based on age of children born to the same mother.
** Significant at 1%, * Significant at 5%

Table 4: Birth Interval Preceding a Female Birth

	[1]	[2]	[3]
	Birth interval before 2^{nd} born	Birth interval before 3^{rd} born	Birth interval before 4^{th} born
	[1]	[2]	[3]
f_2=1	0.012		
	[0.007]		
f_3=1		0.002	
		[0.011]	
f_4=1			0.020
			[0.017]
Observations	259,961	108,150	39,857
R-squared	0.02	0.02	0.02
Sample	2+, f=1	3+, ff=1	4+, fff=1

Estimates from equation (4) shown. All regressions include controls for mother's age and education, family and village characteristics and state dummies. Standard errors clustered at the village level are in parentheses. Birth interval is measured as the difference in ages (in years) between the n-1^{th} and n^{th} births. Sample in column 1 includes all families with at least 2 children, and where the eldest child is a female aged < 18 years; sample in column 2 includes all families with at least 3 children, where the eldest 2 are females aged < 18 years; sample in column 3 includes all families with at least 4 children, where the eldest 3 are females aged < 18 years. * Significant at 5%, ** Significant at 1%

Table 5: Effects of Family Size on Education, Female First-Borns

Instrument	n/a	ff	n/a	fff	n/a	ffff
	[1]	[2]	[3]	[4]	[5]	[6]
First Stage						
Family Size		0.110**		0.105**		0.118**
		[0.005]		[0.008]		[0.012]
Observations		259131		107928		39838
F Test		429.8		154.6		101.8
Second Stage, Outcome ↓	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
School enrolment						
Family Size	-0.021**	-0.017	-0.021**	0.014	-0.020**	-0.020
	[0.001]	[0.016]	[0.001]	[0.027]	[0.002]	[0.038]
Observations	259131		107928		39838	
p-value of test of exogeneity		0.80		0.19		1.00
Years of schooling						
Family Size	-0.111**	-0.024	-0.109**	-0.015	-0.104**	-0.243
	[0.004]	[0.074]	[0.006]	[0.125]	[0.011]	[0.191]
Observations	257566		107286		39611	
p-value of test of exogeneity		0.25		0.45		0.46
Primary school						
Family Size	-0.015**	0.01	-0.014**	-0.008	-0.013**	-0.063+
	[0.001]	[0.014]	[0.001]	[0.026]	[0.002]	[0.035]
Observations	257672		107327		39621	
p-value of test of exogeneity		0.07		0.84		0.14
Lower secondary						
Family Size	-0.020**	0.007	-0.021**	-0.0003	-0.022**	0.013
	[0.001]	[0.017]	[0.001]	[0.023]	[0.002]	[0.035]
Observations	177539		74081		27923	
p-value of test of exogeneity		0.12		0.36		0.31
Sample	2+		3+, ff=1		4+, fff=1	

All regressions include controls for mother's age and education, family and village characteristics and state dummies. Note also that we condition implicitly on the sex composition of the first n-1 births by restricting the sample to the first n-1 births being all-female. For the lower secondary schooling outcome, sample is restricted to 14–17 year olds. Standard errors clustered at the village level in parentheses. Sample in columns 1 and 2 includes the first born in families with at least 2 children where the first born is a female aged < 18 years, sample in columns 3 and 4 includes the first born in families with at least 3 children, where the eldest 2 are female aged < 18 years, and sample in columns 5 and 6 includes the first born in families with at least 4 children, where the eldest 3 are female aged < 18 years. * Significant at 5%, ** Significant at 1%

Table 6: Characteristics of ff Compliers relative to Population

	Ratio of 1st stage for sub-sample listed in column (A) to overall 1st stage
Column A	**ff**
Mother's education:	
No qualification	0.690
Some primary	1.029
Min completed primary school	1.295
Mother is married	1.066
Father is present	1.074
Mother age 35+	0.842
Head works in agriculture	1.016
Head indigenous	0.874
Utilities:	
Availability of water in house	1.248
Availability of light/electricity	1.051
Has own toilet	1.096
Has water in toilet	1.350
Asset ownership:	
Blender	1.221
Fridge	1.229
Gas stove	1.262
Radio	1.049
Gas heater for water	1.083
Record player	1.262
TV	1.152
Video	1.273
Washing machine	1.073
Fan	1.046
Car	1.645
Truck	1.165
Land for agric/forestry	1.029
Animals	0.918

Sample comprises families with at least 2 children where the first-born is a female aged <18 years. Italicised items are those for which ≤10% of the population own one.

Table 7: Effects of Family Size on Education, Female Second- and Third-Borns

	Second-borns				Third-borns	
Instrument →	**n/a**	**fff**	**n/a**	**ffff**	**n/a**	**ffff**
	[1]	[2]	[3]	[4]	[5]	[6]
First Stage						
Family Size		0.125**		0.133**		0.159**
		[0.011]		[0.015]		[0.029]
Observations		64024		25331		9483
F Test		121.30		84.50		31.00
Second Stage, Outcome ↓	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
School Enrolment						
Family Size	-0.023**	-0.002	-0.023**	-0.018	-0.020**	-0.004
	[0.001]	[0.028]	[0.002]	[0.044]	[0.003]	[0.049]
Observations	64024		25331		9483	
p-value of test of exogeneity		0.46		0.91		0.75
Years of schooling						
Family Size	-0.094**	0.122	-0.088**	-0.052	-0.083**	0.012
	[0.006]	[0.116]	[0.010]	[0.185]	[0.012]	[0.188]
Observations	63657		25213		9433	
p-value of test of exogeneity		0.06		0.84		0.61
Primary school						
Family Size	-0.017**	0.037	-0.016**	-0.012	-0.018**	0.020
	[0.001]	[0.030]	[0.002]	[0.043]	[0.003]	[0.058]
Observations	63659		25214		9434	
p-value of test of exogeneity		0.06		0.93		0.51
Lower secondary						
Family Size	-0.012**	0.028	-0.015**	-0.034	n/a	
	[0.001]	[0.033]	[0.002]	[0.040]		
Observations	28592		11651			
p-value of test of exogeneity		0.22		0.63		
Sample	3+, ff=1		4+, fff=1		4+, fff=1	

All regressions include controls for mother's age and education, family and village characteristics and state dummies. Note also that we condition implicitly on the sex composition of the first n-1 births by restricting the sample to the first n-1 births being all-female. For the lower secondary schooling outcome, sample is restricted to 14–17 year olds. Standard errors clustered at the village level in parentheses. Sample in columns 1 and 2 includes the second born in families with at least 3 children where the first two are females aged < 18 years, sample in columns 3 and 4 includes the second born in families with at least 4 children, where the eldest 3 are females aged < 18 years, and sample in columns 5 and 6 includes the third born in families with at least 4 children, where the eldest 3 are females aged < 18 years. * Significant at 5%, ** Significant at 1%

Table 8: Effects of Family Size on Education of Females, Pooled Birth Parities

First Stage.		**Family Size**
Instrument: Subsequent birth a female		0.114**
		[0.005]
F - Stat		562.19
Observations		332638
Second Stage.		
Outcome ↓	**OLS**	**IV**
School enrolment		
Family Size	-0.021**	-0.013
	[0.001]	[0.014]
Observations	332638	332638
p-value of test of exogeneity		0.56
Years of schooling		
Family Size	-0.106**	0.010
	[0.004]	[0.061]
Observations	330656	330656
p-value of test of exogeneity		0.06
Primary school completion		
Family Size	-0.015**	0.015
	[0.001]	[0.012]
Observations	330765	330765
p-value of test of exogeneity		0.01
Lower secondary school completion		
Family Size	-0.019**	0.009
	[0.001]	[0.016]
Observations	207677	207677
p-value of test of exogeneity		0.07
Sample	2+, 3+ & ff=1, 4+ & fff=1	

All regressions include controls for mother's age and education, family and village characteristics and state dummies. For the lower secondary schooling outcome, sample is restricted to 14–17 year olds. Standard errors clustered at the village level in parentheses. Sample includes the first born in families with at least 2 children, where the eldest child is a female aged < 18 years, the second born in families with at least 3 children where the eldest 2 are females aged <18 years, and the third born in families with at least 4 children, where the eldest 3 are females aged < 18 years. * Significant at 5%, ** Significant at 1%

Table 9: Testing Economies of Scale

Dependent Variable→	Purchased children's clothing	Expenditure on children's clothing
	[1]	[2]
	0.004	0.790
All-male	[0.016]	[2.16]
Observations	4,205	4,203

PROGRESA data from October 1998 and May 1999, control villages only. Marginal effects from equations (7) and (8) shown in columns [1] and [2] respectively. Regressions include controls for household demographics, family size, maternal age and education, village size and distance to the nearest town. Sample pools families with at least 2 children where the first is a male, families with at least 3 children where the first 2 are male, and families with at least 4 children where the first 3 are male. All-male=1 if the first n are male, for n=2...4; all-male=0 if first n-1 are male and n is female, for n=2...4. Standard errors clustered at the village level are in parentheses. * Significant at 5% ** Significant at 1%

Table 10: Estimated Bounds

	School enrolment	Years of schooling	Primary school completion	Lower secondary school completion
OLS	-0.021**	-0.106**	-0.015**	-0.019**
IV	-0.013	0.010	0.015	0.009
IV_{Z*}	-0.021	-0.100	-0.014	-0.018
Bounds	[-0.021,-0.013]	[-0.100,0.010]	[-0.014,0.015]	[-0.018,0.009]
Confidence Intervals	(-0.023,0.015)	(-0.108,0.141)	(-0.015,0.041)	(-0.019,0.041)
Observations	332,638	330,656	330,765	207,677

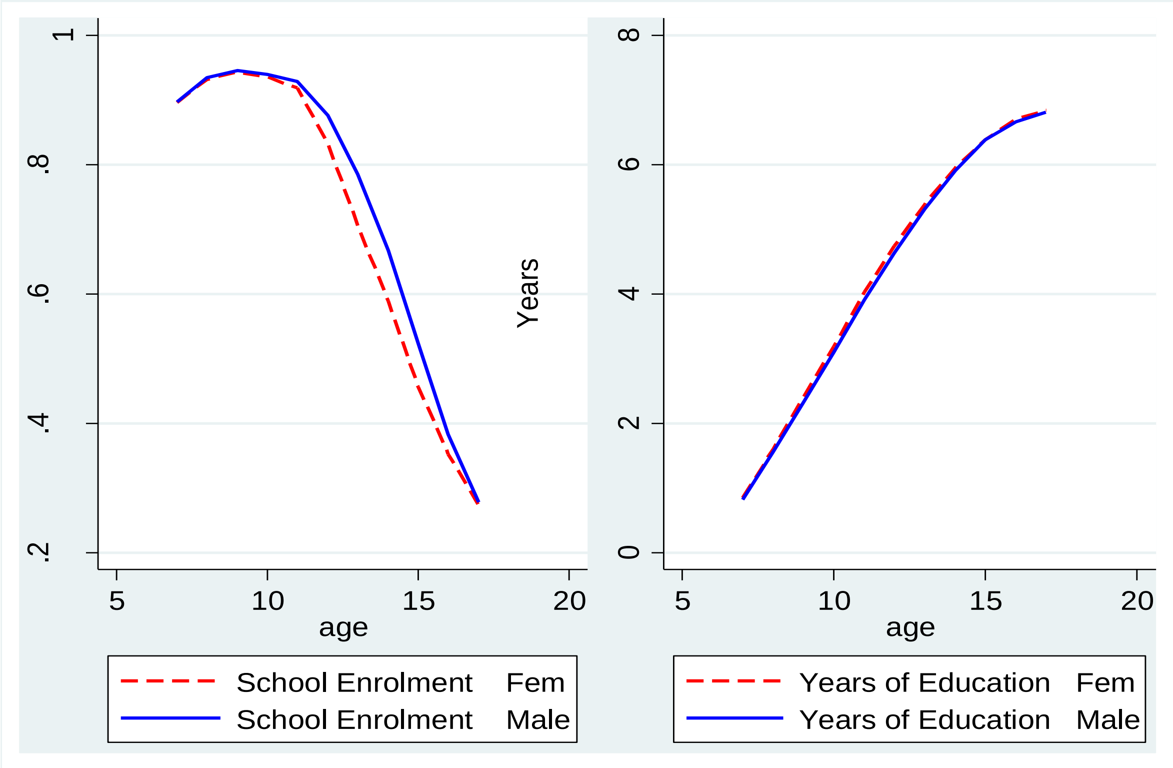
All regressions include controls for mother's age and education, family and village characteristics and state dummies. Sample as described in notes to Table 8. IV_{Z*} corresponds with the lower bound, IV corresponds with the upper bound. * Significant at 5%; ** Significant at 1%

Table 11: Effects of Family Size on Mother's Labour Supply

	OLS	IV
Instrument →	**n/a**	**subsequent birth female**
Outcome ↓		
Mother's work		
Family Size	-0.016**	0.076**
	[0.001]	[0.011]
Observations	334026	
Sample	2+; 3+, ff=1; 4+, fff=1	

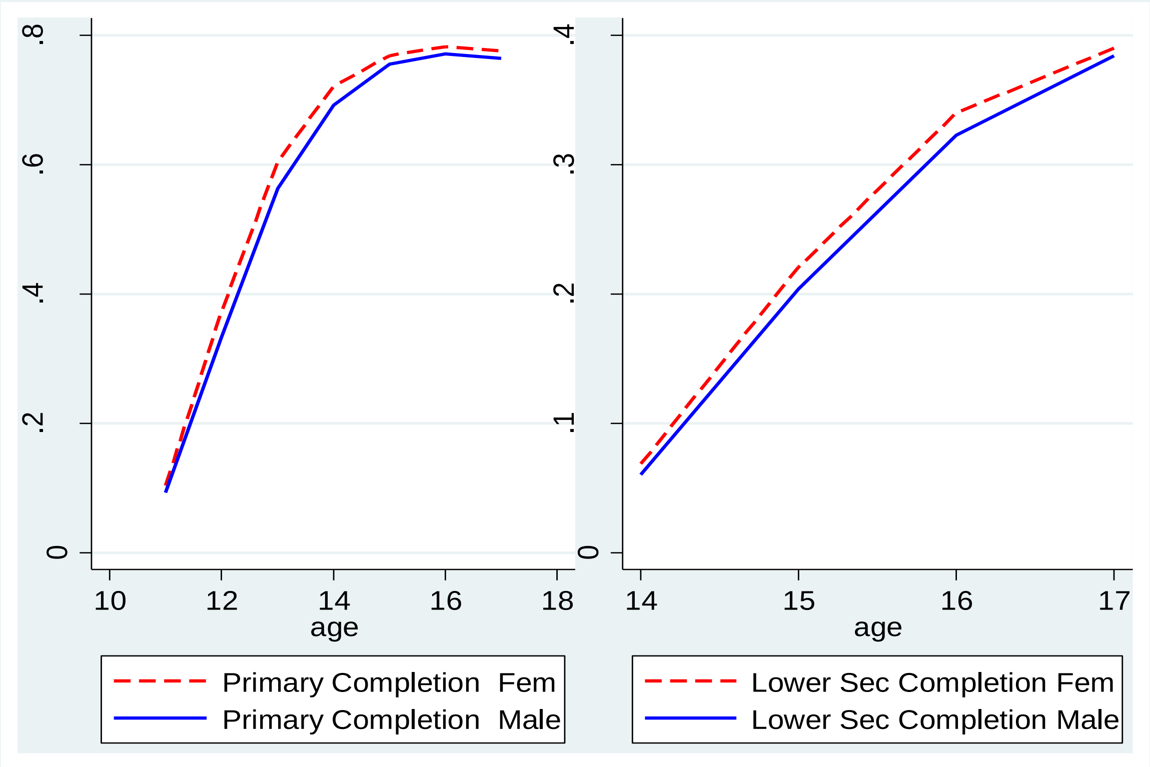
All regressions include controls for mother's age and education, family and village characteristics and state dummies. For the lower secondary schooling outcome, sample is restricted to 14–17 year olds. Standard errors clustered at the village level in parentheses. Sample includes the first born in families with at least 2 children, where the eldest is a female aged < 18 years, the second born in families with at least 3 children where the eldest 2 are females aged <18 years, and the third born in families with at least 4 children, where the eldest 3 are females aged < 18 years. * Significant at 5%, ** Significant at 1%

Figure 1 School enrolment and years of schooling, by age and gender



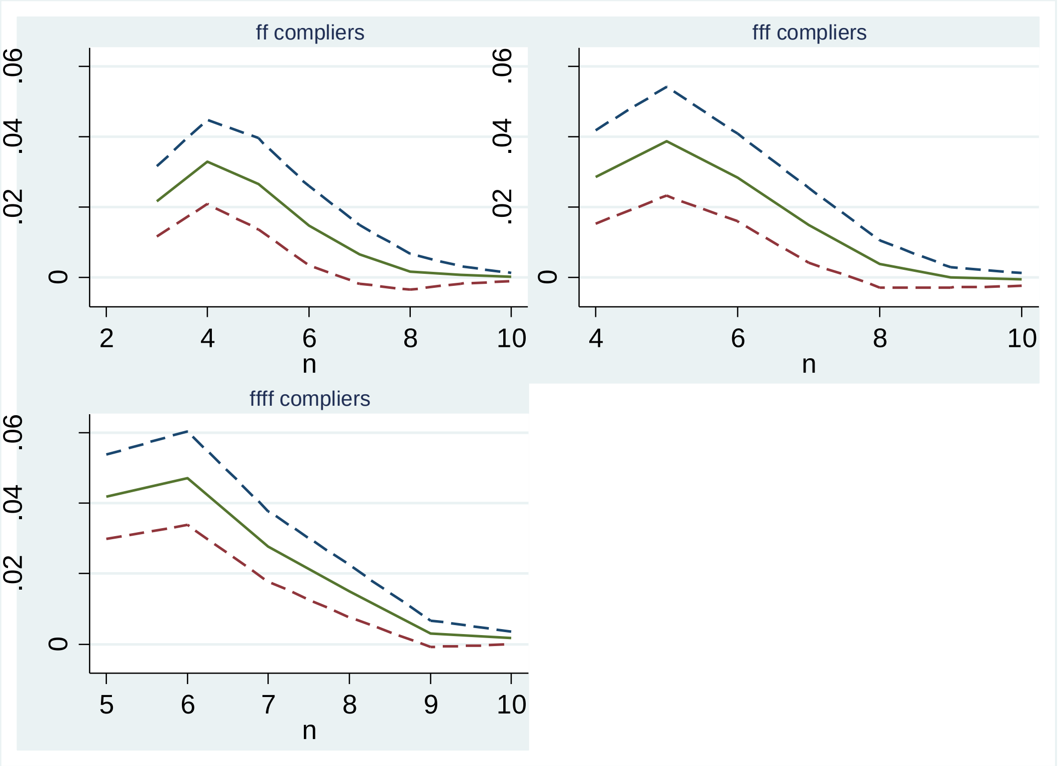
Notes: Graphs plot enrolment in school at time of survey and years of education for all children aged 6-17 years living with their mother

Figure 2 Primary and lower secondary school completion, by age and gender



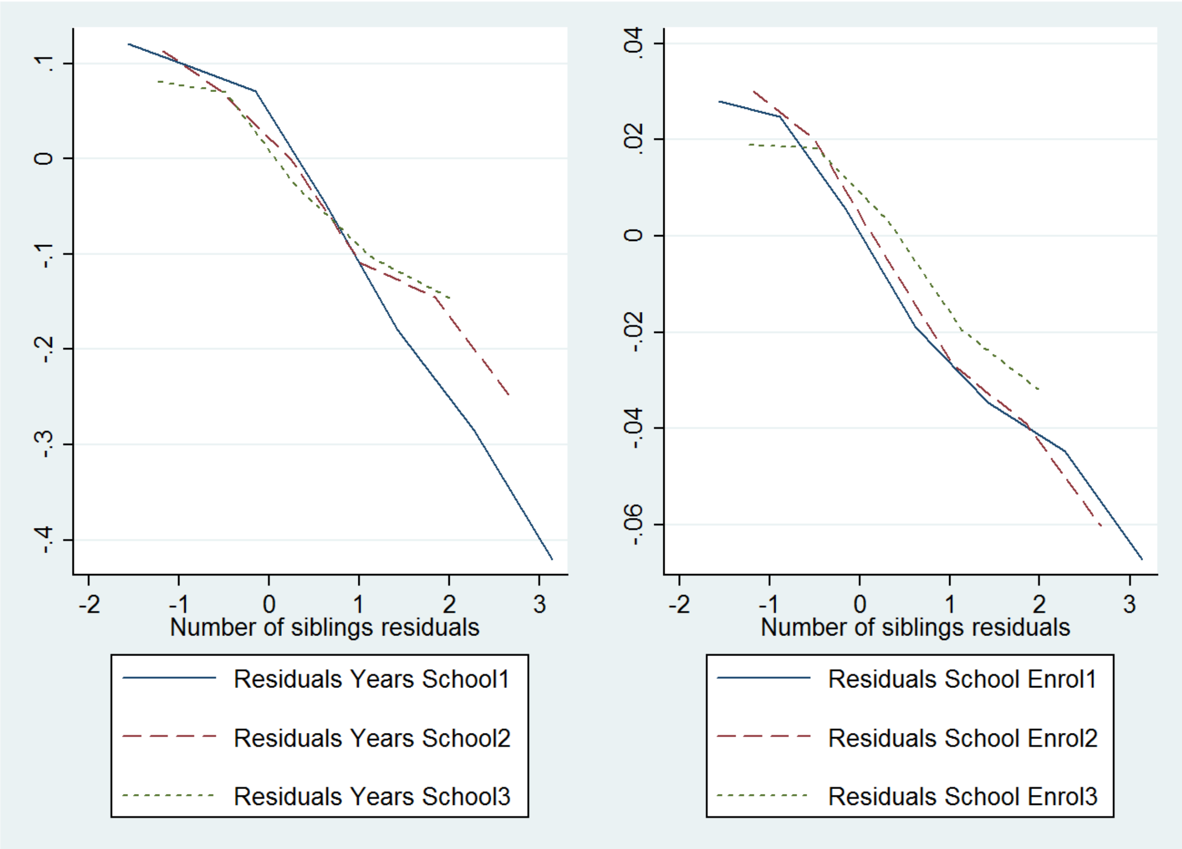
Notes: Graphs plot primary school completion for children aged 11-17 years living with their mother, and lower secondary completion for children aged 14-17 years living with their mother

Figure 3 Compliers, all-female instruments



Dashed lines are 95% bootstrapped confidence intervals. Figures shown are for first-born females; figures for other parities are very similar.

Figure 4 Relationship between family size and education, by birth parity



The figure displays graphs plotting on the vertical (horizontal) axis residuals from a regression of the education outcome (family size) on the control variables. Years of schooling shown on left hand panel; school enrolment on right hand panel. Note that each line denotes a different birth parity, for parities 1 through 3. Figures for the other outcomes reveal similar patterns and are available on request.

References

Agüero J, Marks M (2008) Motherhood and female labor force participation: evidence from infertility shocks. Am Econ Rev: Pap & Proc 98:500–504

Airola J, Juhn C (2005) Wage inequality in post-reform Mexico. Discussion Paper 1525, IZA

Angelucci M, De Giorgi G, Rangel M, Rasul I (2010) Family networks and schooling outcomes: evidence from a randomised social experiment. J Public Econ 94:197–221

Angelucci M, De Giorgi G, Rangel M, Rasul I (2012) Insurance and investment within family networks. Mimeo, University College London

Angrist J (2004) Treatment effect heterogeneity in theory and practice. Econ J 114:C52–C83

Angrist J, Evans, W (1998) Children and their parents' labor supply: evidence from exogenous variation in family size. Am Econ Rev 88:450–477

Angrist J, Imbens, G (1995) Two stage least squares estimates of average causal response models with variable treatment intensity. J Am Stat Assoc 90:431–442

Angrist J, Imbens G, Rubin D (1996) Identification of causal effects using instrumental variables. J Am Stat Assoc 91:444–455

Angrist J, Lavy V, Schlosser A (2010) Multiple experiments on the causal link between the quantity and quality of children. J Labor Econ 28:773–824

Arnold F (1992) Sex preference and its demographic and health implications. Int Fam Plan Perspect 18:93–101

Attanasio O, Binelli C (2010) Mexico in the 1990s: the main cross-sectional facts. Rev Econ Dyn 13:238–264

Becker G (1960) An economic analysis of fertility. In: NBER (ed) Demographic and economic change in developed countries. Columbia University Press, New York, pp 209–231

Becker G, Lewis HG (1973) On the interaction between the quantity and quality of children. J Polit Econ 81:S279–S288

Becker G, Tomes N (1976) Child endowments and the quantity and quality of children. J Polit Econ 84:S143–S162

Behrman JR, Rosenzweig MR, Taubman P (1994) Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment. J Polit Econ 102:1131–1174

Behrman JR, Sengupta P, Todd P (2005) Progressing through PROGRESA: an impact assessment of a school subsidy experiment in rural Mexico. Econ Dev & Cult Chang 54:237–275

Bhaskar V (2011) Sex selection and gender balance. Am Econ J: Microeconomics 3:214–244

Black S, Devereux P, Salvanes K (2005) The more the merrier? The effect of family size and birth order on children's education. Q J Econ 120:669–700

Black S, Devereux P, Salvanes K (2010) Small family, smart family? Family size and the IQ scores of young men. J Hum Resour 45:33–58

Black S, Devereux P, Salvanes K (2011) Older and wiser? Birth order and IQ of young men. CESifo Econ Stud 57:103–120

Brown M, Flinn C (2011) Family law effects on divorce, fertility and child investment. Mimeo, Federal Reserve Bank of NY, New York University

Butcher K, Case A (1994) The effect of sibling sex composition on women's education and earnings. Q J Econ 109:531–563

Cáceres-Delpiano J (2006) The impacts of family size on investment in child quality. J Hum Resour 41:738–754

Chakravarty A (2010) Supply shocks and gender bias in child health investments: evidence from the ICDS programme in India. B.E. J Econ Anal & Policy: Topics. doi:10.2202/1935-1682.2402

Conley D, Glauber R (2006) Parental educational investment and children's academic risk: estimates of the impact of sibship size and birth order from exogenous variation in fertility. J Hum Resour 41:722–737

Dahl GB, Moretti E (2008) The demand for sons. Rev Econ Stud 75:1085–1120

Deaton A, Subramaniam S (1996) The demand for food and calories. J Polit Econ 104:133–162

Duryea S, Galiani S, Nopo H, Piras C (2007) The educational gender gap in Latin America and the Caribbean. Working Paper 600, Inter-American Development Bank

Ebenstein A (2008) The 'missing girls' of China and the unintended consequences of the One Child Policy. J Hum Resour 45:87–115

Imbens G, Angrist J (1994) Identification and estimation of local average treatment effects. Econom 62:467–475

Koch H (1955) Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children. Genet Psychol Monogr 52:3–50

Lee J (2008) Sibling size and investment in children's education: an Asian instrument. J Popul Econ 21:855–875

Li H, Zhang J, Zhu Y (2008) The quantity-quality tradeoff of children in a developing country: identification using Chinese twins. Demogr 45:223–243

Millimet D, Wang L (2011) Is the quantity-quality trade-off a trade-off for all, none or some? Econ Dev & Cult Chang 60:155–195

Mogstad M, Wiswall M (2010) Testing the quantity-quality model of fertility: linearity, marginal effects, and total effects. Working Paper 5216, Department of Economics, NYU

Nevo A, Rosen A (2012) Identification with imperfect instruments. Rev Econ & Stat 94:659–671

Oster E (2009) Does increased access increase equality? Gender and child health investments in India. J Dev Econ 89:62–76

Parazzini F, La Vecchia C, Levi F, Franceschi S (1998) Trends in male:female ratio among newborn infants in 29 countries among 5 continents. Hum Reprod 13:1394–1396

Parker SW, Pederzini C (2000) *Gender Differences in Education in Mexico*. World Bank, Washington DC

Ponczek V, Souza AP (2012) New evidence of the causal effect of family size on child quality in a developing country. J Hum Resour 47:64–106

Qian N (2009) Quantity-quality and the One Child Policy: the only-child disadvantage in school enrollment in rural China. Working Paper 14973, NBER

Rosenblum D (2013) The effect of fertility decisions on excess female mortality in India. J Popul Econ 26:147–180

Rosenzweig MR, Wolpin K (1980) Testing the quantity-quality fertility model: using twins as a natural experiment. Econom 48:227–240

Rosenzweig MR, Wolpin K (2000) Natural 'natural' experiments in economics. J Econ Lit 38:827–874

Rosenzweig MR, Zhang J (2009) Do population control policies induce more human capital investment? Twins, birthweight, and China's 'One Child' Policy. Rev Econ Stud 76:1149–1174

Schultz TP (2004) School subsidies for the poor: evaluating the Mexican Progresa poverty program. J Dev Econ 74:199–250

Schultz TP (2008) Population policies, fertility, women's human capital and child quality. In: Schultz TP, Strauss J (eds) Handbook of development economics, volume 4. North Holland, Amsterdam, pp 3249–3303

Stoye J (2009) More on confidence intervals for partially identified parameters. Econom 77:1299–1315

UNESCO (2011) UIS Statistics in Brief, Latin America and the Caribbean. http://stats.uis.unesco.org/unesco

Wu X, Li L (2012) Family size and maternal health: evidence from the One-Child policy in China. J Popul Econ 25:1341-1364

Zajonc RB (1976) Family configuration and intelligence. Sci 192:227–236

APPENDIX 1

Using Sex Composition as an Instrument for Family Size

In this section, we present results using sex composition as an instrument for family size. The analysis follows closely upon Angrist et al (2010). The sex composition instrument can be defined in two ways, both of which are used (separately) below: (i) a single indicator for whether the first n births are of the same sex, regardless of whether male or female, and (ii) two indicators - one for an all-female composition and one for an all-male composition.

We conduct this analysis on the first n-1 children from families with at least n births, separately for n=2, 3, 4. These samples are labelled 2+, 3+ and 4+ respectively below. For illustrative purposes, we show the first-stage specifications for n=2+ and for n=3+.[40]

The first-stage specifications for the sample of first-borns in families with at least two children (2+ families) are:

$$F_i = \quad \alpha_0 + \alpha_1 X + \beta_1 m_{1i} + \beta_2 m_{2i} + \pi Z + \xi_i \qquad \text{(A1)}$$

$$F_i = \quad \alpha_0 + \alpha_1 X + \beta_1 m_{1i} + \eta_f ff_i + \eta_m mm_i + \xi_i \qquad \text{(A2)}$$

where equation (A1) uses as an instrument an indicator for whether the first 2 births in i's family are of the same sex, Z, and equation (A2) uses as an instrument indicators for the first 2 births being female (ff_i) and the first 2 births being male (mm_i). Note also that $m_{1i}(m_{2i})$ is an indicator for a male first (second) birth in i's family.[41]

The first-stage specifications for the sample of families with at least 3 children (3+ families) are:

$$F_i = \quad \alpha_0 + \alpha_1 X_i + \beta_1 m_{1i} + \beta_2 m_{2i} + \ \beta_3 m_{3i} + \alpha_2 s_{12i} + \pi Z + \beta k_i + \xi_i \qquad \text{(A3)}$$

$$F_i = \quad \alpha_0 + \alpha_1 X_i + \alpha_2 m_{1i} + \eta_1 ff_i + \eta_2 mm_i + \eta_3 (1 - s_{12i}) * m_{3i} + \eta_4 fff_i + \eta_5 mmm_i + \beta k_i + \xi_i \qquad \text{(A4)}$$

[40] The first-stage equation for the sample of families with at least 4 children follows a similar pattern. Note also that the second-stage specifications corresponding to (A1)-(A4) are the same as equation (1) in the main text.
[41] In equation A2, we must drop m_{2i} since $\{m_{1i}, m_{2i}, ff_i, mm_i\}$ are linearly dependent.

where Z in equation (A3) is an indicator for whether the first 3 births in i's family are of the same gender, and fff_i and mmm_i (in equation (A4)) are respectively indicators for the first 3 births being male and female in i's family. In addition, we control in both specifications for the sex composition of earlier births (in equation (A3), through the terms m_{1i}, m_{2i}, m_{3i}as defined already, and s_{12i}, which is an indicator that the first 2 births in i's family are of the same gender; in equation (A4) through the terms $m_{1i}, ff_i,\ mm_i$ and $(1 - s_{12i}) * m_{3i}$). This means that the parameter π in equation (A3) estimates the difference in family size between families with a succession of 3 same-sex children (fff_i or mmm_i) and those with a specific mixed sex composition (2 same-sex children followed by a different sex child, i.e. - ffm_i or mmf_i, consistent the conditioning in section 2). The parameter η_4 (η_5) in equation (A4) captures the difference in family size between those with an fff_i (or mmm_i) composition relative to a specific mixed composition - ffm_i (mmf_i).

The top panel of Table A1 shows the first-stage estimates for the 2+, 3+ and 4+ samples. It shows that families where the first n children are of the same sex have an additional 0.06-0.089 children compared to families with a mixed-sex composition. However, when we allow for differential effects by gender (as per equation (A4); columns 3, 6, 9), it becomes clear that this correlation is driven primarily by families with an all-female composition. In all cases, the correlation for the all-female composition is at least twice that of the all-male instrument (and almost 7 times for the ff and mm instruments), highlighting the strong son preferences in this population.

Turning to the second-stage estimates shown in the lower panels of Table A1, we see that the manner in which the instrument is specified matters importantly for the identified effect, particularly in the 2+ case (columns 2, 3). The reason for this is that the second-stage TSLS estimate is equivalent to a weighted average of instrument-specific causal effects (i.e. TSLS estimates computed using a particular instrument on its own), where the weights depend on the relative magnitudes of the first stages (Imbens and Angrist 1994; Angrist and Imbens 1995). So, the TSLS estimates reported in column 3, where the first stage coefficient for the ff instrument is almost 7 times as large as that for the mm instrument, are driven primarily by causal effects for the ff compliers. By contrast, the TSLS estimates displayed in column 2 are computed in a manner that weights equally the causal effects for ff and mm compliers, thereby generating different estimates to those in column 3.

[Insert Table A1 here]

Table A1: Effect of Family Size on Education, Mixed Sex and Same-Sex Instruments

	First-Borns, 2+			First- and Second-Borns, 3+			First-, Second- and Third-Borns, 4+		
			n=2			n=3			n=4
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Dep var -->		**Family Size**	**Family Size**		**Family Size**	**Family Size**		**Family Size**	**Family Size**
First Stage, Instrument ↓									
samesex		0.063**			0.084**			0.089**	
		[0.004]			[0.006]			[0.009]	
n^{th} birth=female[a]			0.110**			0.113**			0.129**
			[0.006]			[0.010]			[0.013]
n^{th} birth=male[a]			0.016**			0.0549**			0.051**
			[0.006]			[0.008]			[0.013]
Observations		548,693	548,693		728,616	728,616		619,042	619,042
F Test		290.80	213.30		185.10	93.46		99.14	57.20
Second Stage, Outcome ↓	**OLS**	**IV**	**IV**	**OLS**	**IV**	**IV**	**OLS**	**IV**	**IV**
School Enrolment	-0.019**	0.052**	-0.005	-0.022**	0.012	0.010	-0.021**	0.010	-0.004
Family Size	[0.001]	[0.018]	[0.015]	[0.001]	[0.020]	[0.020]	[0.001]	[0.032]	[0.028]
Observations	548,693	548,693	548,693	728,616	728,616	728,616	619,042	619,042	619,042
p-value of test of exogeneity		0.00	0.34		0.08	0.10		0.32	0.53
Years of schooling									
Family Size	-0.108**	-0.001	-0.023	-0.114**	0.001	0.010	-0.106**	-0.168	-0.162
	[0.003]	[0.087]	[0.073]	[0.003]	[0.090]	[0.091]	[0.004]	[0.147]	[0.140]
Observations	545,565	545,565	545,565	724,696	724,696	724,696	615,820	615,820	615,820
p-value of test of exogeneity		0.27	0.24		0.20	0.17		0.66	0.68
Primary school									
Family Size	-0.015**	0.021	0.013	-0.017**	0.010	0.010	-0.017**	-0.030	-0.034
	[0.0004]	[0.018]	[0.014]	[0.0005]	[0.020]	[0.020]	[0.001]	[0.030]	[0.029]
Observations	545,765	545,765	545,765	724,864	724,864	724,864	615,924	615,924	615,924
p-value of test of exogeneity		0.04	0.05		0.18	0.19		0.65	0.54
Lower secondary									
Family Size	-0.019**	-0.010	0.004	-0.018**	0.002	0.003	-0.017**	-0.023	-0.018
	[0.001]	[0.020]	[0.017]	[0.001]	[0.018]	[0.018]	[0.001]	[0.028]	[0.026]
Observations	381,337	381,337	381,337	443,672	443,672	443,672	347,863	347,863	347,863
p-value of test of exogeneity		0.64	0.19		0.26	0.24		0.84	0.97
Sample		2+			3+			4+	

All regressions include controls for mother's age and education, family and village characteristics and state dummies. For the lower secondary schooling outcome, sample is restricted to 14–17 year olds. Standard errors clustered at the village level in parentheses. Sample in columns 1 and 2 contains first-borns in families with at least 2 children where the eldest < 18 years, columns 3 and 4 includes the first- and second-borns in families with at least 3 children and where the eldest is < 18 years, sample in columns 5 and 6 includes the first-, second- and third-borns in households with at least 4 children, where the eldest is < 18 years. * Significant at 5%, ** Significant at 1%. [a] All specifications control for the sex compositions of previous (n-1) births. Estimates in col. 3 relative to an fm/mf composition, those in col. 6 relative to an ffm/mmf composition, those in col. 9 relative to an fffm/mmmf composition.