

Fitzsimons, Emla; Malde, Bansi

**Working Paper**

## Empirically probing the quantity-quality model

IFS Working Papers, No. 10,20

**Provided in Cooperation with:**

The Institute for Fiscal Studies (IFS), London

*Suggested Citation:* Fitzsimons, Emla; Malde, Bansi (2010) : Empirically probing the quantity-quality model, IFS Working Papers, No. 10,20, Institute for Fiscal Studies (IFS), London

This Version is available at:

<https://hdl.handle.net/10419/47493>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Empirically Probing the Quantity-Quality Model

IFS Working Paper 10/20

Emla Fitzsimons  
Bansi Malde

# Empirically Probing the Quantity-Quality Model

Emla Fitzsimons\* and Bansi Malde<sup>†</sup>

September 2010

**Abstract:** This paper tests whether family size has a causal effect on girls' education in Mexico. It exploits son preference as the main source of random variation in the propensity to have more children, and estimates causal effects using instrumental variables. Overall, it finds no evidence of family size having an adverse effect on education, once the endogeneity of family size is accounted for. Results are robust to another commonly used instrument in this literature, the occurrence of twin births. A divisive concern throughout this literature is that the instruments are invalid, so that inferences including policy recommendations may be misleading. An important contribution of this paper is to allow for the possibility that the instruments are invalid and to provide an answer to the question of just how much the assumption of instrument exogeneity drives findings. It concludes that the assumption of exogeneity does not affect the results that much, and the effects of family size on girls' schooling remain extremely modest at most.

**Keywords:** Fertility; Education; Instrumental Variables; Mexico.

**JEL classification:** I20, J13

**Acknowledgements:** The authors are grateful to the Economic and Social Research Council for funding through research grant RES-000-22-2877. All errors are the responsibility of the authors. We are grateful to conference participants at the European Economic Association Annual Conference, Milan 2008; also to Orazio Attanasio, Valerie Lechene, David Phillips, Nancy Qian, Marta Rubio-Codina, and Marcos Vera-Hernandez for useful discussions.

Correspondence: emla\_f@ifs.org.uk, bansi\_m@ifs.org.uk

---

\* Institute for Fiscal Studies, London, <sup>†</sup> Institute for Fiscal Studies, London and University College London.

# 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 beneficial 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, hence, largely an empirical one, and indeed one on which causal evidence in developing countries remains scarce.<sup>1</sup>

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, 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 (Rosenzweig and Zhang, 2009; Li, Zhang and Zhu,

---

<sup>1</sup> More generally, 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. See Schultz (2005) for a review.

2007; Qian, 2009).<sup>2</sup> Another paper using a new instrument - infertility shocks - finds no evidence of a negative effect of family size on children's education in a sample of Latin American countries (Agüero and Marks, 2008).<sup>3</sup> Our paper contributes to this gap in the literature 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 main source of exogenous variation in family size exploited is parental preferences for having at least one son. We find no evidence to support the quantity-quality trade-off in our setting: the negative correlation between family size and schooling observed in the data 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 school outcomes. It is also robust to another source of exogenous variation in family size, the birth of twins.

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 (2011) on the other hand find no evidence invalidating the identifying restrictions in an Israeli context.<sup>4</sup> Very few other studies directly examine the extent to which concerns about instrument validity underlies 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

---

<sup>2</sup> Li, Zhang and Zu (2007) 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 (Qian, 2009).

<sup>3</sup> Other than these studies, work that estimates the effects of family size on children's education generally relates to developed countries, and generally shows no or only very weak evidence of a quantity-quality trade-off (Black, Devereux and Salvanes (2005), Cáceres-Delpiano (2006), Conley and Glauber (2006) – all for the U.S.; Angrist, Lavy and Schlosser (2011) for Israel).

<sup>4</sup> 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.

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 (2008). 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 modest at best. Another contribution of the paper is to show that although what is identified are local average treatment effects, our findings are likely to generalise to our entire sample, which comprises the population of rural indigent households in Mexico. We do this by characterising the complier sub-populations across different instruments. These contributions of the paper set it apart from other related studies and fill an important gap in a literature.

Furthermore, the data used in this paper, spanning over half a million relatively poor households in marginalised communities in rural Mexico, allow 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 Mexican sample used here is just over 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

$$Y_i = \beta_0 + \beta_1 X + \beta_2 F_i + u_i \quad (1)$$

where the outcome variables,  $Y_i$ , pertain to child  $i$ 's education and include a 0-1 indicator of participation in school, accumulated 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 individual, parental, household and village characteristics;  $F_i$  is family size of child  $i$ ; and the error term  $u_i$  denotes unobserved factors that affect  $Y_i$  and that may be correlated with  $F_i$ .<sup>5</sup> This model is estimated using pooled cross sections of data from 1996 through 1999, covering the entire population of rural indigent communities in Mexico (detailed in section 3).

Estimating equation (1) by ordinary least squares (OLS) would render the coefficient of interest,  $\beta_2$ , biased and inconsistent if omitted variables, such as parental preferences, influence both children's outcomes and family size. To obtain a consistent estimate of  $\beta_2$ , an instrumental variable method is used, that requires the existence of a variable,  $Z$ , that is correlated with  $F_i$  but uncorrelated with  $u_i$ . In a first-stage regression, we estimate

$$F_i = \alpha_0 + \alpha_1 Z + \alpha_3 X + \xi_i \quad (2)$$

---

<sup>5</sup> We use a linear specification in this paper, given that the instrumental variables are binary.

The main source of exogenous variation in family size used in this paper is all-female births.<sup>6</sup> Our population exhibits a strong son preference: that the first  $n$  births are female is highly correlated with further childbearing; that the first  $n$  births are male is not. In a later section we also assess robustness of findings to another instrument commonly used in this literature, the birth of twins.<sup>7</sup>

The instrument is effectively the sex of the  $n^{\text{th}}$  child in households in which the first  $n-1$  births are female<sup>8</sup>: we expect (and later show) family size to be higher in households where the  $n^{\text{th}}$  birth is also female. We do this for  $n=2\ldots 4$ . We consider the outcomes of the first  $n-1$  children, all female by definition.<sup>9,10</sup> 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. The importance of birth order for education choices has been highlighted in the recent literature (Black et al (2007, 2010), Rosenzweig and Zhang (2009)), yet the majority of estimates of family size effects thus far pertain to first-borns only.<sup>11</sup> Another advantage of our strategy is that we can distinguish whether the effects of family size vary depending on the margin of increase considered, an important point raised by Black et al (2007, 2010). These different analyses are possible due to the very large

---

<sup>6</sup> The use of sex composition as an instrument for family size was pioneered by Angrist and Evans (1998), and has since been applied by researchers such as Angrist, Lavy and Schlosser (2009) and Conley and Glauber (2006). These studies use same-sex births as the instrument, whether all boys or all-females; Lee (2008) on the other hand uses all-female births.

<sup>7</sup> Rosenzweig and Wolpin (1980) were the first to use twin births as an instrument for family size; it has since been used by Caceres (2005), Black et al (2005), and Angrist et al (2010), among others.

<sup>8</sup> We condition on the first  $n-1$  births being female as the instrument is preference for at least one son.

<sup>9</sup> The natural reason for this is that children of the  $n^{\text{th}}$  birth may be of different sexes; another important reason is that it avoids any selection bias arising from families who go on to have children after a male birth being different from those who do not.

<sup>10</sup> Whilst the all-female instrument does not allow us to obtain effects for boys, we believe that the advantage, in terms of robustness, outweighs this drawback. We note also that we have estimated effects for boys using the twin births instruments and have found no effects of family size on any education outcome considered.

<sup>11</sup> There is some research to suggest that the effects on the first-born may provide a lower bound of the effect of family size, as the eldest child is exposed to a lower family size for some time compared to the rest of the children (Kessler, 1991).



sample we have as well as high fertility rates, at an average of just under 4 children per woman.

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 correlation between it and the error term (Nevo and Rosen, 2008). This allows us, for the first time in this literature, to provide bounds on the magnitude 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  $n$  girls that go on to have an additional child solely because they wish to have a boy. This sub-population is called the compliers (Angrist, Imbens and Rubin (1996)). A potential limitation is that the effect only pertains to this specific group and is not applicable to the wider population. To investigate this, we compare effects obtained from another instrument widely used in the literature, twin births, which affects different complier sub-populations. We also compare characteristics of the compliers for both instruments. This allows us 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 marginalised rural areas throughout 31 states in Mexico between 1996 and 1999.<sup>12</sup> Our sample therefore comprises particularly poor households, as the descriptive statistics will confirm later on. The survey - the Survey of Household Socio-Economic Characteristics (*Encuesta de Características Socioeconómicas de los Hogares, ENCASEH*) - was conducted in order to aid in the targeting of the PROGRESA (now *Oportunidades*) welfare programme, introduced in selected marginalised 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 locality 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 instrumental variables, different margins of increase in family size, and different birth orders.

The analysis is restricted to 12-17 year olds, as school enrolment before age 12 is practically universal, at just over 97%. Further, we drop households in which the eldest child is 18 or above (39%).<sup>13</sup> A potential concern with the remaining sample is that we may miscode family size (and birth orders) if older children have left the household permanently. We believe this to be a relatively minor concern: only 2.1% of households report having a household member who migrated permanently in the past 5 years. Note also that we retain households in which both parents are married, thus dropping any divorced parents from the analysis. This is

---

<sup>12</sup> Most localities were chosen on the basis of having been graded with a high degree of marginalisation from the 1995 Census data.

<sup>13</sup> 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 reporting of implausible ages.

relatively innocuous, as divorce is extremely uncommon in the sample, at below 1%. This leaves us with a sample of just over half a million households across just under 1,500 villages. Family size is defined as the number of biological siblings in the household, i.e. the number of children born to the same parents. Other children present in the household (such as nephews and nieces), are not considered to be part of the sibship but their presence is controlled for in all specifications.

## 3.2 Descriptive Statistics

### 3.2.1 Our Sample

We first show some characteristics of the sample of households in Table 1. The average family size is 4. Around 50% of households have children of the same sex in the first two births: just under half of these have two girls. Mothers are 38 years old on average and have just over 3 years of education; fathers are 42 years old and have just below 4 years of education. Less than 30% of mothers have at least completed primary schooling, while the corresponding figure for fathers is just over 30%. Agricultural work is widespread, with almost 80% of households involved 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.

**Table 1 Mean characteristics of households**

<b>Variable</b>	<b>Mean</b>	<b>Std. Dev.</b>
Number of siblings	3.951	1.838
Proportion of households with 1 <sup>st</sup> 2 births of the same sex	0.504	0.500
Proportion of households with 1 <sup>st</sup> 2 births female (ff)	0.234	0.423
Proportion of households with 1 <sup>st</sup> 3 births female (fff)	0.119	0.323
Proportion of households with 1 <sup>st</sup> 4 births female (ffff)	0.062	0.241
<b>Socio-economic variables</b>		
Father's age	42.073	8.535

Mother's age	37.887	7.279
Father's years of schooling	3.607	3.021
Mother's years of schooling	3.252	2.859
Father has no schooling	0.201	0.401
Mother has no schooling	0.267	0.442
Father has at least completed primary schooling	0.308	0.462
Mother has at least completed primary schooling	0.278	0.448
Birth spacing b/w 1 <sup>st</sup> and 2 <sup>nd</sup> borns	2.837	1.927
Indigenous language speakers	0.340	0.474
Household owns dwelling	0.925	0.264
Water supply in dwelling	0.244	0.429
Electricity in dwelling	0.797	0.402
Number of rooms in dwelling	1.921	1.207
Household has own toilet	0.631	0.482
Household has water in toilet	0.199	0.399
Household owns land	0.518	0.500
Household head works in agriculture	0.773	0.419
Wall materials of dwelling (0 = poor quality)	0.875	0.331
Roof materials of dwelling (0 = poor quality)	0.409	0.492
N	529,857	

---

*Notes:* Sample of households with at least one 12-17 year old, in which the eldest child is <age 18.

### 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, as well as three different measures of the stock of education: years of schooling, completion of primary schooling and completion of lower secondary schooling.<sup>14</sup>

---

<sup>14</sup> 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.

The latter two variables are potentially the more relevant outcomes for policymakers: whilst around 85% of children complete primary school, just over half 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.<sup>15</sup> Completion of levels is also of interest in the presence of “sheepskin effects” in the returns to schooling. Another point in their favour is that they are likely to be less affected by measurement error than years of education.

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. The fact that measures of education are similar across males and females suggests that son preferences do not affect intra-household allocation choices once a child is born.<sup>16</sup> As we will see in section 4, this is reassuring from the point of view of the validity of the instrument.

The figures also show a sharp drop in school enrolment at age 12, which corresponds to the first year of lower secondary school (see Figure 1). Before that age, school enrolment is practically universal (corresponding to primary schooling). For this reason we consider school

---

<sup>15</sup> 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.

<sup>16</sup> Schultz (2004) and Behrman et al (2003) document higher 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 to availability of schools/distance to schools/marriage markets rather than preferences for boys’ schooling *per se*.

choices from age 12 onwards only. Figure 1 also shows that years of schooling are increasing with age, though not one-to-one.

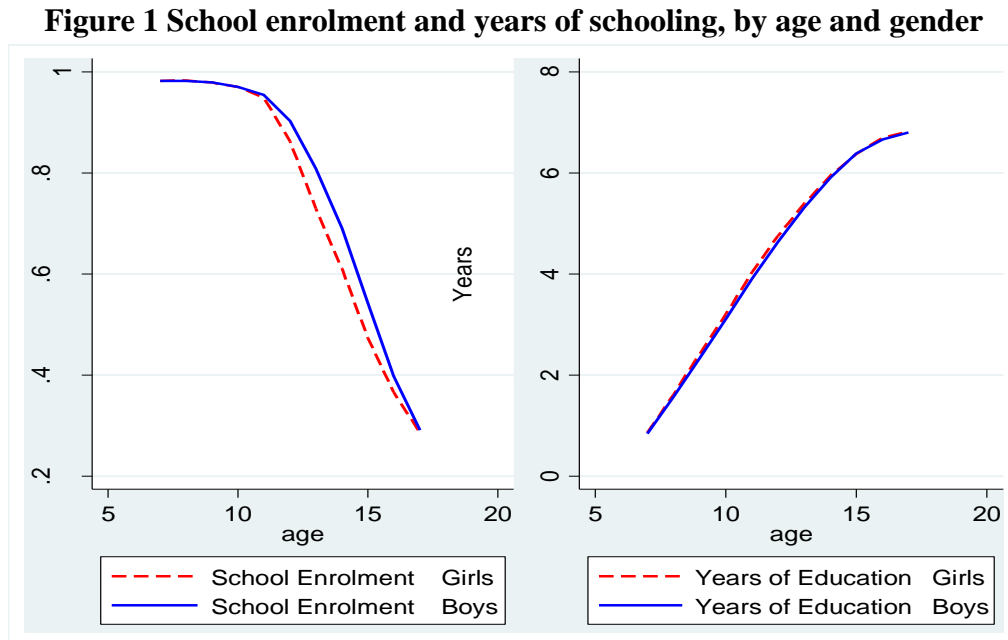
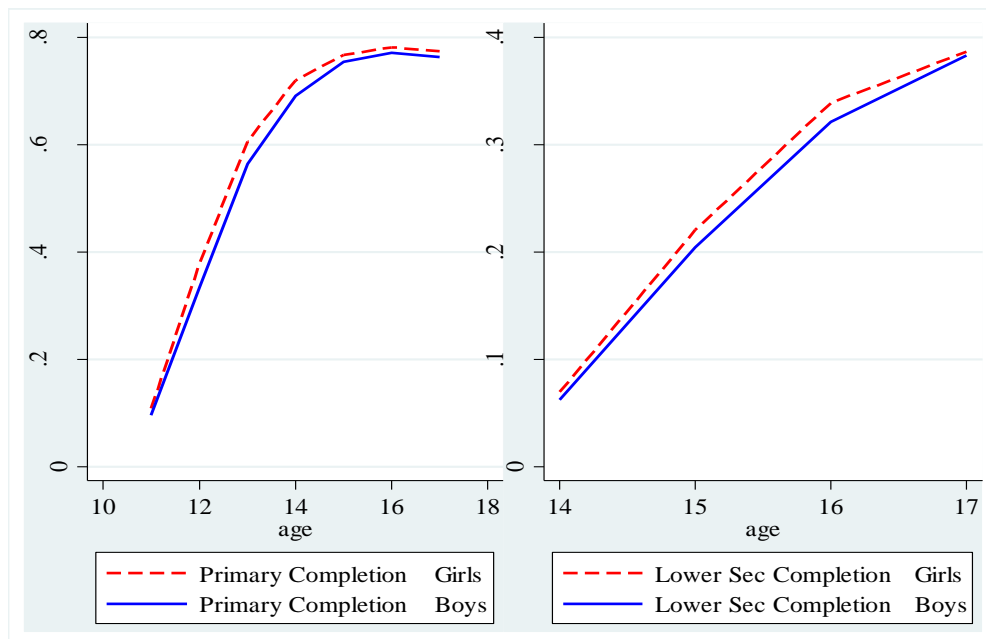


Figure 2 shows primary school completion and lower secondary school completion for 12-17 year olds (both of which are free and publicly provided).<sup>17</sup> The proportions completing primary school and lower secondary school are low. By age 12, the age at which a child should have completed primary schooling, less than 40% of children has done so, and less than 80% of boys and girls 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.

<sup>17</sup> 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.

**Figure 2 Primary and Lower Secondary School Completion by age and gender**



### 3.2.3 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). We believe that this issue is unlikely to arise in our sample: Mexico is a predominantly Catholic country where abortion is highly legally restricted. Indeed Table 2, which compares characteristics of parents (age and education) whose first  $n-1$  births are girls, and who have either a girl or a boy at the  $n^{\text{th}}$  birth, confirms that the samples are well-balanced, giving us no reason to believe that the instruments are not random.

**Table 2 Mean characteristics by sex composition of earlier births**

<b>Variable</b>	<b>fm=1</b>	<b>ff=1</b>	<b>Difference in means</b>	<b>p-value</b>
Father's age	40.912	40.885	-0.027	0.394
Mother's age	36.806	36.781	-0.024	0.354
Mother's age at first birth	22.234	22.229	-0.005	0.830
Father's years of schooling	3.744	3.744	0.000	0.973
Mother's years of schooling	3.389	3.386	-0.002	0.856
Father has no schooling	0.186	0.188	0.002	0.214
Mother has no schooling	0.248	0.251	0.004	0.030*
Father has at least completed primary school	0.325	0.327	0.002	0.327
Mother has at least completed primary school	0.294	0.295	0.001	0.672
Birth spacing b/w 1 <sup>st</sup> and 2 <sup>nd</sup> births	2.836	2.845	0.009	0.244
Family size	4.168	4.282	0.114	0.000*
N	111,588	108,911		
	<b>ffm=1</b>	<b>fff=1</b>		
Father's age	40.102	40.045	-0.057	0.302
Mother's age	36.027	35.994	-0.032	0.425
Mother's age at first birth	21.462	21.440	-0.023	0.540
Father's years of schooling	3.761	3.762	0.001	0.978
Mother's years of schooling	3.386	3.375	-0.011	0.551
Father has no schooling	0.184	0.185	0.001	0.852
Mother has no schooling	0.248	0.250	0.002	0.458
Father has at least completed primary school	0.330	0.331	0.001	0.690
Mother has at least completed primary school	0.294	0.294	0.000	0.886
Birth spacing b/w 1 <sup>st</sup> and 2 <sup>nd</sup> births	2.573	2.604	0.031	0.002*
Birth spacing b/w 2 <sup>nd</sup> and 3 <sup>rd</sup> births	3.016	3.017	0.001	0.919
Family size	4.597	4.717	0.119	0.000*
N	47,207	46,348		
	<b>fffm=1</b>	<b>ffff=1</b>		



Father's age	39.555	39.563	0.008	0.923
Mother's age	35.479	35.501	0.021	0.704
Mother's age at first birth	20.889	20.859	-0.030	0.567
Father's years of schooling	3.666	3.606	-0.060	0.020*
Mother's years of schooling	3.222	3.166	-0.056	0.062
Father has no schooling	0.188	0.193	0.004	0.232
Mother has no schooling	0.261	0.269	0.008	0.107
Father has at least completed primary school	0.318	0.313	-0.005	0.255
Mother has at least completed primary school	0.273	0.266	-0.007	0.138
Birth spacing b/w 1 <sup>st</sup> and 2 <sup>nd</sup> births	2.378	2.398	0.020	0.109
Birth spacing b/w 2 <sup>nd</sup> and 3 <sup>rd</sup> births	2.637	2.637	0.000	0.991
Birth spacing b/w 3 <sup>rd</sup> and 4 <sup>th</sup> births	2.896	2.907	0.011	0.563
Family size	5.194	5.332	0.138	0.000*
N	17,571	17,588		

*Notes:* N refers to the number of first-born female children. fm=1 indicates female at 1<sup>st</sup> birth, male at 2<sup>nd</sup> birth; ff=1 indicates female at 1<sup>st</sup> 2 births, and so on. A \* indicates that the variable is statistically different from 0 at the 5% level or less.

## 4 Results

In this section we first display estimates from the first-stage relationships between family size and the instruments. We then show the two-stage least squares (TSLS) estimates, alongside the linear probability model (LPM) estimates for comparison.<sup>18</sup>

### 4.1 First-stage relationships

Table 3 shows the first-stage correlations between family size and the instruments. The top panel of each table shows the first-stage coefficients for first-borns, the middle and lower panels show those for second- and third-borns respectively.

<sup>18</sup> One of our outcome variables, years of schooling, is not binary: thus we use OLS estimation in its case. For convenience we use the term LPM throughout the text.

**Table 3 Effect of sex-composition on family size**

	[1]	[2]	[3]
	ff	fff	ffff
<b>First-Borns</b>			
Family size	0.117** [0.006]	0.115** [0.009]	0.125** [0.012]
Observations	219563	93153	34998
F test	349.72	148.63	114.14
<b>Second-Borns</b>			
Family size	n/a	0.136** [0.012]	0.141** [0.015]
Observations		55922	22407
F test		123.66	86.93
<b>Third-Borns</b>			
Family size	n/a	n/a	0.173** [0.03]
Observations			8447
F test			33.53
Sample	2+, 1 <sup>st</sup> born =female	3+, ff=1	4+, fff=1

*Notes:* Dependent variable is family size. All regressions control for the socio-economic variables listed in Table 1. \* Denotes statistical significance at the 1% - 5% level. \*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level are in parentheses. 2+ indicates households with 2 or more children, etc; ff indicates females at 1<sup>st</sup> 2 births, etc.

The 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, around 1 in 10 first-born

girls gain an additional sibling due to the instrument. Put differently, the proportion of households in the complier subpopulation ranges between 10% and 15%.

We further 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 below.<sup>19</sup> The horizontal axis gives completed family size<sup>20</sup>; the vertical axis gives the proportion of households that has that family size because the instrument is switched on, and that would not otherwise have continued their fertility. So for instance, Figure 3 shows 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 not 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 to 8 children.<sup>21</sup> 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 household is 4).

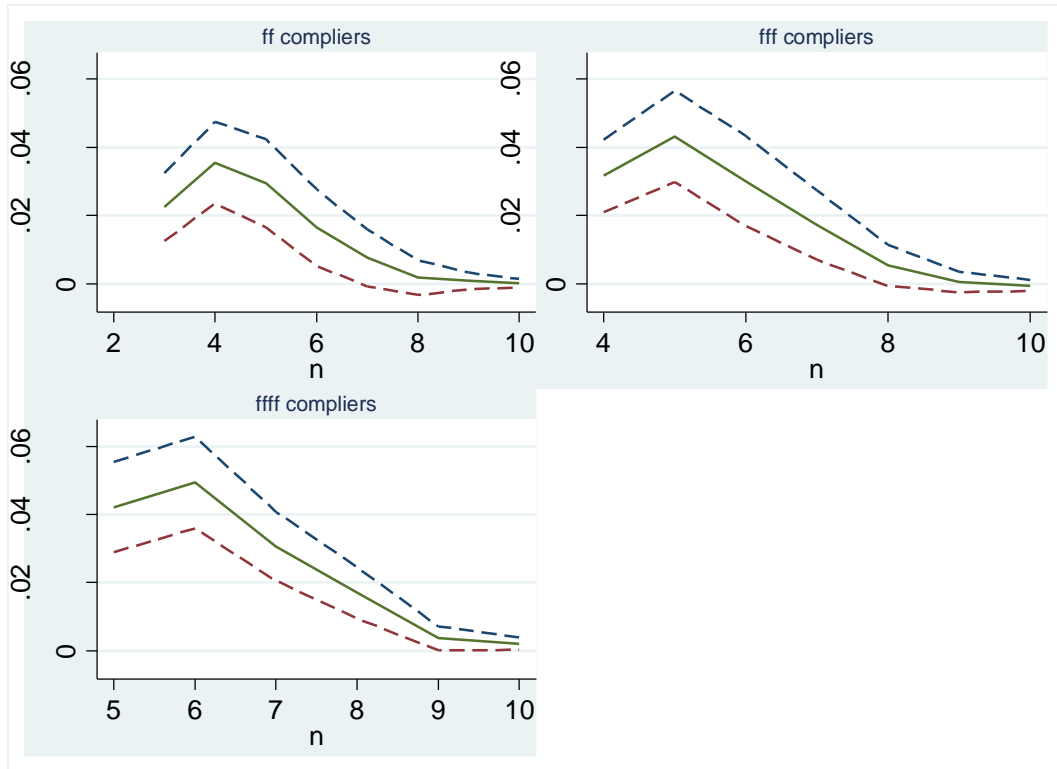
---

<sup>19</sup> The first stages are decomposed following Angrist and Imbens (1995).

<sup>20</sup> Though we use the term completed family size, it refers to completed as at the time of the survey.

<sup>21</sup> Whilst we do not explicitly consider non-linear effects of family size in this paper (see Mogstad and Wiswall (2010) for an analysis), 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.

**Figure 3 Compliers, all-female instruments**



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

## 4.2 Two-stage least squares estimates

In this section we display the TSLS estimates of the effect of family size on children's education. The LPM estimates, which do not account for the endogeneity of family size, are also shown. Results are shown separately by birth order, for four outcomes: school enrolment, years of schooling, primary school completion, and lower secondary school completion.

We see from the estimates in Table 4 that regardless of birth parity or outcome considered, the LPM estimates are negative and significantly different from zero. In terms of magnitude, an extra child is associated with a reduction of 2 percentage points in school enrolment, primary school completion and lower secondary school completion for all birth orders. The magnitude for years of education is around 0.1 years. These magnitudes are larger than those

found by Agüero & Marks (2008) for school enrolment in a sample of Latin American countries. They are, however, in line with those found by Angrist et al (2011) for Israel.

When we instrument for family size, the magnitude of the effect of family size on schooling outcomes changes, regardless of birth parity: the coefficient remains generally negative, but in no case statistically significant from zero, suggesting that family size has no effect on education. This finding is consistent across the outcomes, birth orders and family sizes considered.

**Table 4 Effects of family size on education, first-borns**

	LPM	IV	LPM	IV	LPM	IV
<b>Instrument →</b>	<b>n/a</b>	<b>ff</b>	<b>n/a</b>	<b>fff</b>	<b>n/a</b>	<b>ffff</b>
<b>Outcome ↓</b>						
<b>School enrolment</b>						
Family Size	-0.020** [0.001]	-0.019 [0.015]	-0.020** [0.001]	0.007 [0.025]	-0.018** [0.002]	-0.032 [0.036]
Observations	219597		93158		35007	
<b>Years of schooling</b>						
Family Size	-0.110** [0.004]	0.035 [0.085]	-0.110** [0.006]	-0.017 [0.123]	-0.105** [0.011]	-0.149 [0.176]
Observations	218380		92652		34813	
<b>Primary school</b>						
Family Size	-0.014** [0.001]	0.022 [0.017]	-0.014** [0.001]	0.003 [0.026]	-0.013** [0.002]	-0.051 [0.035]
Observations	218469		92688		34823	
<b>Lower secondary</b>						
Family Size	-0.020** [0.001]	0.009 [0.019]	-0.021** [0.001]	-0.017 [0.022]	-0.021** [0.002]	0.035 [0.034]
Observations	150722		63961		24508	

Sample	2+	2+	3+, ff=1	3+, ff=1	4+, fff=1	4+, fff=1
--------	----	----	----------	----------	-----------	-----------

*Notes:* Control for socio-economic variables listed in Table 1. Note also that using the  $n^{\text{th}}$  birth as an instrument, we condition implicitly on the sex of the first  $n-1$  births. For the lower secondary schooling outcome, sample is restricted to 14–17 year olds.

\* Denotes statistical significance at the 1% - 5% level.

\*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level in parentheses.

**Table 5 Effects of family size on education, second- and third-borns**

	Second borns				Third-borns	
	LPM	IV	LPM	IV	LPM	IV
Instrument →	n/a	fff		ffff		fff
<b>Outcome ↓</b>						
<b>School Enrolment</b>						
Family Size	-0.022**	-0.008	-0.021**	-0.023	-0.018**	-0.008
	[0.001]	[0.028]	[0.002]	[0.045]	[0.003]	[0.047]
Observations	55921		22411		8443	
<b>Years of schooling</b>						
Family Size	-0.094**	0.07	-0.086**	-0.118	-0.074**	-0.176
	[0.006]	[0.111]	[0.009]	[0.197]	[0.012]	[0.181]
Observations	55624		22310		8398	
<b>Primary school</b>						
Family Size	-0.017**	0.035	-0.015**	-0.026	-0.016**	-0.024
	[0.001]	[0.028]	[0.002]	[0.044]	[0.003]	[0.058]
Observations	55626		22311		8399	
<b>Lower secondary</b>						
Family Size	-0.012**	-0.007	-0.015**	-0.009	n/a	
	[0.001]	[0.030]	[0.002]	[0.040]		
Observations	25114		10355			
Sample	3+, ff=1	3+, ff=1	4+, fff=1	4+, fff=1	3+, ff=1	3+, ff=1

*Notes:* see notes to Table 5. Additionally, note that lower secondary school completion is not considered as an outcome variable for third-borns as they are too young to have completed it.

While these findings are similar to those found by Agüero & Marks (2008), they differ markedly from those of Qian (2009) and Rosenzweig and Zhang (2009) for China. However, given the strict fertility restrictions in China, these latter papers concern effects in families with at most 2 or 3 children, whereas we consider a much wider range of fertility change.<sup>22</sup>

Much of the literature concludes on the basis of this evidence alone that when the endogeneity of family size is taken into account, there is no evidence of a quantity-quality trade-off for children's education. In this paper, we investigate 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 look into how representative the findings are for the population at large: by first estimating effects using another instrument – twin births – which picks up a different set of compliers, and then by characterising the compliers for whom the LATE effects are identified. Third, we investigate whether families are adjusting on margins other than children's education, in particular mother's labour supply.

## 5 Robustness

A key concern throughout this literature relates to 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, more often than not, instrument validity is not directly addressed.<sup>23</sup> 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

---

<sup>22</sup> By contrast, the findings are very much in line with previous findings for developed countries (Cáceres-Delpiano (2006) for the US; Black et al (2005) for Norway; Angrist et al (2009) for Israel). Moreover, Cáceres-Delpiano (2006) and Black et al (2005) use a different source of identification to our main one – twin births.

<sup>23</sup> Exceptions include Rosenzweig and Zhang (2009) and Rosenzweig and Wolpin (2000), who provide direct evidence on the likely validity of twin and same-sex 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.

a new approach in this paper by testing directly how robust the findings are 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 (2008). 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. As a final robustness exercise, we use twin births as an instrument for family size. This instrument affects a different set of compliers, which is useful as a way of assessing whether our findings are specific to the all-female compliers, and whether another instrument may show up evidence of variability in the effects of family size across other groups in the population. Indeed we go on to characterise the two sets of compliers, to see just how comparable they are.

### **5.1.1 Evidence on instrument validity**

As has been discussed, the exclusion restriction is that the sex of the  $n^{\text{th}}$  born has no direct effect on education. There are at least two concerns with this. The first is that son preferences may directly affect 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.<sup>24</sup> In both cases, the direction of the resulting bias of the IV estimate is positive: if postnatal son preferences exist (and if they affect education decisions), then a sister is more beneficial for girls' schooling than a brother; if scale economies are important, accumulated savings may be higher in all-female households, yet difficult to control for.

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

---

<sup>24</sup> This is a widely recognised concern in this literature; see for instance Rosenzweig and Wolpin (2001), Rosenzweig and Zhang (2009).



daughters, and not that parents treat sons more favourably than daughters. However if postnatal son preferences exist, the sibship gender composition may affect intra-household schooling choices. We are not unduly concerned about this: we have seen in section 3 that education outcomes for boys and girls are very similar (see Figures 1 and 2), and this conforms with recent trends in Mexico showing convergence in education between the sexes.<sup>25</sup> In further investigation of this, we estimated a school participation model for girls, including as regressors the number of sisters and brothers (above and below age 5, separately). Neither of the coefficients is statistically different from zero, suggesting that males do not have a detrimental effect on their sisters' education.

A potentially more serious concern - and one that has received much attention in the literature - is economies of scale resulting in savings from all-female births 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 industrialised countries that the scope for economies of scale is much more limited. Traditional hand-me-downs that can result in economies of scale include children's clothing and shoes, though in the environment we consider these goods tend to be unisex, especially at young ages. Other common gender-specific items such as school books are likely to be common to both given the predominance of mixed-sex schools in our setting. Moreover, the sharing of gender-specific goods is unlikely to be restricted to *within* the household, but to take part across the extended family and social network.<sup>26</sup>

---

<sup>25</sup> Recent UNESCO statistics for Mexico show that 98% of girls and 98% of boys are in primary school; 72% of girls and 70% of boys are in secondary school (UNESCO, 2007); 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 girls and boys below the age of 20 no longer display significant differences in educational attainment, as measured by years of schooling. Duryea et al (2007) analyse the educational gender gap in Latin America and the Caribbean and find that the most striking differences are across income groups and not gender.

<sup>26</sup> Angelucci et al (2009, 2010) document the importance of extended family networks for this population in providing mutual support to households, and making schooling decisions.

To provide more factual evidence however, we take a look at data on expenditures on children's clothing and shoes. As this information is not available in the ENCASEH survey, we instead use data from the Progresa evaluation survey, which provides information on expenditures on children's clothing and shoes in the previous 6 months for around 26,000 households from 7 states in rural Mexico in 1998/99. These data are informative about our population, as the Progresa sample was drawn from the ENCASEH survey.<sup>27</sup>

The evidence suggests strongly that economies of scale are not an important concern. First, the purchase of children's clothing and shoes is very infrequent: a large proportion of households (61% and 45% respectively) have purchased neither over a 6 month period; amongst those that have purchased these items, expenditures account for just 1% of their monthly non-durable consumption. This is consistent with Attanasio et al (2009), who find that households in this population spend around 70% of their budget on food, leaving little scope for scale economies (compared to less than 20% in western industrialised economies).

When we test more directly whether the sex composition of children affects the household's decision to purchase children's clothes and shoes, we find no evidence that it does. We estimate the following equation:

$$D_h = \lambda_0 + \lambda_1 f_{nh} + \lambda_2 X + \lambda_3 share_h + \xi_h \quad (3)$$

where  $D_h$  is a dummy variable equal to 1 if a household reports positive expenditure on children's clothing/shoes (separately) and 0 otherwise,  $f_{nh}$  is a dummy variable equal to 1 if the first  $n$  children in a household are female and 0 otherwise,  $X$  is a vector of control variables including household demographics, parental age and education, ages of the first  $n$ -

---

<sup>27</sup> We pool post-programme data from surveys in October 1998 and May 1999, from control villages only, to keep the analysis uncontaminated by any potential programme effects. We retain households 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, households here have fewer children on average; parents are also on average younger, but more educated, as we retain younger households here.

borns (to control for age differences between children), family size (to disentangle sex composition effects from family size effects), locality variables such as locality size and distance to the nearest large town (to proxy for costs of purchasing these goods), and  $share_h$  is the share of non-durable consumption a household spends on food (to control for available household resources). Equation (3) is estimated at the household level using a probit model. Estimates are shown in Table 6. We find no evidence of sex composition affecting these purchase decisions, across any sub-sample considered. Moreover, for the 2+ and 3+ subsamples, the coefficients are close to zero across all 4 outcomes considered.

**Table 6 Effects of sex composition on purchase of children's clothing/shoes**

	Dep Vble: Purchased children's clothes			Dep Vble: Purchased children's shoes		
	[1]	[2]	[3]	[4]	[5]	[6]
ff	-0.014 [0.023]			0.032 [0.023]		
fff		-0.001 [0.033]			0.005 [0.036]	
ffff			-0.077 [0.050]			0.064 [0.046]
Observations	2570	981	318	2570	981	318
Sample	2+	3+, ff=1	4+, fff=1	2+	3+, ff=1	4+, fff=1

*Notes:* Progres data from October 1998 and May 1999, control villages only. Sample includes households where the eldest child is <18 and is female. Marginal effects from probit estimates of equation (3) shown.

\* Denotes statistical significance at the 1% - 5% level. \*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level are in parentheses.

Nor do we find any evidence of sex composition affecting the amount spent on children's clothing/shoes. Table 7 shows tobit coefficient estimates from the following model

$$M_h = \theta_0 + \theta_1 f_{nh} + \theta_2 X + \theta_3 share_h + u_h \quad (4)$$

where  $M_h$  is household expenditure on children's clothing/shoes (separately) in pesos, and all other variables are as previously defined.

**Table 7 Effects of sex composition on expenditure on children's clothing/shoes**

	Dep Vble: Expenditure on children's clothing			Dep Vble: Expenditure on children's shoes		
	[1]	[2]	[3]	[4]	[5]	[6]
ff	-0.997 [2.171]			0.653 [1.146]		
fff		0.414 [3.265]			-1.457 [1.912]	
ffff			-9.27 [6.045]			0.858 [2.889]
Observations	2528	965	311	2528	621	311
Sample	2+	3+, ff=1	4+, fff=1	2+	3+, ff=1	4+, fff=1

Notes: See notes to Table 6. Coefficients from tobit estimates of equation (4) shown.

Again, this shows no consistent evidence of differences related to sex composition in expenditures on children's clothing or shoes for the subsamples considered.

Another way to assess whether economies of scale associated with same-sex births have an independent effect on education is to test whether an all-male sex composition (which has no effect on family size) affects education.<sup>28</sup> If it does, the concern may be that it is due to economies of scale. To look at this, we estimate the relation between all-male composition and education as follows

$$Y_i = \mu_0 + \mu_1 X + \mu_2 m_{nh} + v_i \quad (5)$$

where  $Y_i$  denotes the 4 schooling outcomes considered throughout, and  $m_{nh}$  is a dummy variable that takes the value 1 if the first  $n$  children are male and 0 otherwise. If an all-male composition has a direct effect on outcomes through economies of scale, we expect  $\mu_2$  to be  $>0$ . Table 8 displays  $\mu_2$  as estimated from equation (5) using the 2+, 3+ and 4+ samples for

<sup>28</sup> This is similar to the strategy used by Angrist et al (2009) who estimate reduced-form twins effects on outcomes in samples where twins have little effect on family size, which they term the "no-first-stage samples".

first-born males. It provides no cause for concern: apart from the lack of statistical significance of all but one of the estimates, the magnitudes are all extremely low, exceeding half of a percentage point in just two instances.

**Table 8 Effect of all-male composition on education outcomes, first-born males**

	mm	mmm	mmmm
<b>Dep vble ↓</b>			
School Participation	0.008**	0.002	0.003
	[0.002]	[0.003]	[0.005]
Observations	245811	106097	39446
Years of education	-0.003	-0.003	0.013
	[0.009]	[0.012]	[0.018]
Observations	244565	105596	39277
Primary Completion	0.001	-0.001	0.003
	[0.002]	[0.002]	[0.004]
Observations	244653	105630	39288
Lower Secondary Completion	-0.002	-0.001	-0.002
	[0.002]	[0.003]	[0.004]
Observations	173247	75404	28764
Sample	2+	3+	4+

*Notes:* OLS estimates for first-borns displayed. Each column refers to a separate regression of the dependent variable on the variable listed in the top of each column. mm denotes first two births males, and so on. Also control for socio-economic variables listed in Table 1. \* Denotes statistical significance at the 1% - 5% level. \*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level are in parentheses. Estimates for second- and third-borns are similar and not shown.

Taken together, all of this evidence suggests that the threats to the validity of the all-female instrument are not very serious in this context.<sup>29</sup> Still, this evidence alone does not (and could not) establish validity of the instrument. A new contribution of the paper is to allow for the

<sup>29</sup> As further reassuring evidence, we re-emphasise that there is no relation between all-female births and any of the covariates in our model— see Table 2. 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 and father's education, household assets, home and land ownership).

instrument to be imperfect and under weaker identification assumptions, derive bounds on the effects of family size on education, which is what we do next.

### 5.1.2 Bounds

In this section we consider explicitly just how much the assumption of instrument exogeneity drives the results. We do this using the method of Nevo and Rosen (2008), imposing weaker assumptions on the degree of correlation between the instrument and the error term, and estimating bounds on the effects of family size on education. So whilst we no longer point identify the model parameters, the benefit is that inferences made are robust to a lack of instrument exogeneity. And more importantly, it 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.

As in Nevo and Rosen (2008), we consider cases where

1. *The instruments have the same direction of correlation with the error term as the endogenous regressor [A3]:* the potential correlation between the instrument and the error term in the outcome equation is positive (see section 5.1.1). The correlation between the endogenous regressor ( $F$ ) and the error term is negative however. To satisfy [A3], we simply specify the treatment variable as  $-F$ .
2. *The instruments are less correlated with the error term than the endogenous regressor [A4]:* this assumption tightens the bounds further in many cases. We believe it is reasonable to expect the all-female instrument to be less correlated with the error term in the outcome equation than is family size.

Nevo and Rosen show that the correlation between the instrumental variable (all-females) and the endogenous regressor ( $-F$ ) plays a key role in estimating the bounds: the key condition is

that this correlation is negative, which we know to be the case. They also show that the larger its magnitude, the tighter the bounds.<sup>30</sup>

When we implement this method, we derive bounds on the effects of family size on education, as shown in Table 9 and Table 10. The bounds are informative in all cases. They suggest that even if we allow for the instrument to be invalid, this does not affect findings by much. The magnitudes of the effects remain very modest. 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.

<b>Table 9 First-born females, Bounds</b>			
	<b>ff</b>	<b>fff</b>	<b>ffff</b>
<b>School enrolment</b>			
OLS	-0.020**	-0.020**	-0.018**
IV	-0.019	0.007	-0.032
Bounds	[-.019,-.018]	[-.018,.007]	[-.032,-.019]
Observations	219597	93158	35007
<b>Years of schooling</b>			
OLS	-0.110**	-0.110**	-0.105**
IV	0.035	-0.017	-0.149
Bounds	[-.103,.035]	[-.106,-.016]	[-.149,-.107]
Observations	218380	92652	34813
<b>Primary school completion</b>			
OLS	-0.014**	-0.014**	-0.013**
IV	0.022	0.003	-0.051
Bounds	[-.012,.022]	[-.013,.003]	[-.051,.015]
Observations	218469	92688	34823

<sup>30</sup> For the twin birth instrument on the other hand, the direction of the correlation between the instrument and the error term is unclear (see section 5.1.3) so this method is not useful without imposing assumptions to sign the bias.

<b>Secondary school completion</b>			
OLS	-0.020**	-0.021**	-0.021**
IV	0.009	-0.017	0.035
Bounds	[-.018,.009]	[-.020,-.017]	[-.017,.035]
Observations	150722	63961	24508

*Notes:* Control for variables listed in Table 1. \* Denotes statistical significance at the 1% - 5% level, \*\* Denotes statistical significance at the 1% level or less. Additionally, work on methods for deriving confidence sets for the bounds is currently ongoing in the theoretical literature and so are not displayed here.

Useful benchmarks for contextualising the bounds computed above come from empirical estimates for developing countries. Estimates obtained by Agüero & Marks (2008) for school enrolment and years of schooling for first-borns in a similar context to ours, a range of Latin American countries, lie within or are very close to the bounds estimated in this paper. Qian (2009) and Rosenzweig and Zhang (2009) on the other hand obtain estimates that lie outside these bounds, though their setting of China and its one child policy is very different, which is a likely explanation. More reassuringly, the bounds indicate that even when the identification assumption is relaxed, the effects are zero or very modest at worst, a conclusion similar to that reached by Rosenzweig and Zhang (2009).

**Table 10 Second- and third-born females, Bounds**

	<b>Second-borns</b>		<b>Third-borns</b>
	fff	ffff	ffff
<b>School enrolment</b>			
OLS	-0.022**	-0.021**	-0.018**
IV	-0.008	-0.023	-0.008
Bounds	[-.021,-.008]	[-.023, -.021]	[-.018,-.006]
Observations	55921	22411	8443
<b>Years of schooling</b>			
OLS	-0.094**	-0.086**	-0.074**



IV	0.07	-0.118	-0.176
Bounds	[-.085,.070]	[-.118, -.088,]	[-.174, -.081]
Observations	55624	22310	8398
<b>Primary school completion</b>			
OLS	-0.017**	-0.015**	-0.016**
IV	0.035	-0.026	-0.024
Bounds	[-.014,.035]	[-.026,-.015]	[-.025, -.017]
Observations	55626	22311	8399
<b>Secondary school completion</b>			
OLS	-0.012**	-0.015**	
IV	-0.007	-0.008	n/a <sup>a</sup>
Bounds	[-.012,-.007]	[-.014,-.009]	
Observations	25114	10355	

Notes: See notes to Table 9. <sup>a</sup> lower secondary school completion not considered as an outcome variable for third-borns as they are too young to have completed it.

### 5.1.3 Twin births as an instrument

To assess further the robustness of our results, we consider another instrument commonly used in this literature: the birth of twins. There are at least two instrument validity concerns, highlighted in a recent paper by Rosenzweig and Zhang (2009) - differential endowments of twins and singletons, and differential birth spacing between twins than between two singleton births.<sup>31</sup> Unlike the all-female instrument, the direction of the ensuing bias on the IV estimate is either positive or negative, depending on the interaction of the two effects.<sup>32</sup> This rules out the method of Nevo and Rosen (Assumption [A3] is violated), unless one is willing to make assumptions on the direction of the likely correlation. Moreover, we have insufficient data to

<sup>31</sup> Differential birth spacing may place differing demands on household resources (both monetary and time), with older siblings of twins likely to have fewer resources available to them than older siblings of singletons. Differential birth endowments directly affect outcomes if parents choose to reinforce or compensate for these differences in their investment decisions. Economies of scale may also be higher in twin households, of which there is evidence in China (Rosenzweig and Zhang, 2009).

<sup>32</sup> It is positive if parents reinforce birth endowments and the birth spacing effect is small, and negative otherwise.

allow us to assess empirically the severity of these concerns in our context. It is for these reasons that twin births are not our preferred source of exogenous variation. Nonetheless, they are a useful specification check since this instrument is affected by different sources of bias, which should affect estimates in a different way.

Just under 1% of our sample reports a twin birth, though our samples are so large that we can consider separately as instruments twins at the  $n^{\text{th}}$  birth, for  $n=2\ldots 4$ . The sample is restricted to households with at least  $n$  births, and the outcomes of children born before the  $n^{\text{th}}$  birth are considered. This is to ensure that on average, preferences over family size are the same in households with and without twins at  $n^{\text{th}}$  birth. It also avoids the problem that families with more births are more likely to have a twin birth. As for the all-female instrument, for  $n=2$  we consider the outcomes of first-borns; for  $n=3$  first- and second-borns; for  $n=4$ , first-, second- and third-borns.

The first stage estimates are shown in Table 11. The coefficients are larger than those for the all-female instruments, as by definition, twin births result in an additional family member. The magnitudes all exceed 0.4 and are increasing with the birth at which the twin birth occurs.

<b>Table 11 The effect of twins on family size</b>			
	<b>[1]</b>	<b>[2]</b>	<b>[3]</b>
	<b>twin2</b>	<b>twin3</b>	<b>twin4</b>
<b>First-Borns</b>			
Family size	0.560**	0.626**	0.707**
	[0.037]	[0.035]	[0.030]
Observations	221196	187616	136998
F test	232.83	324.98	565.42

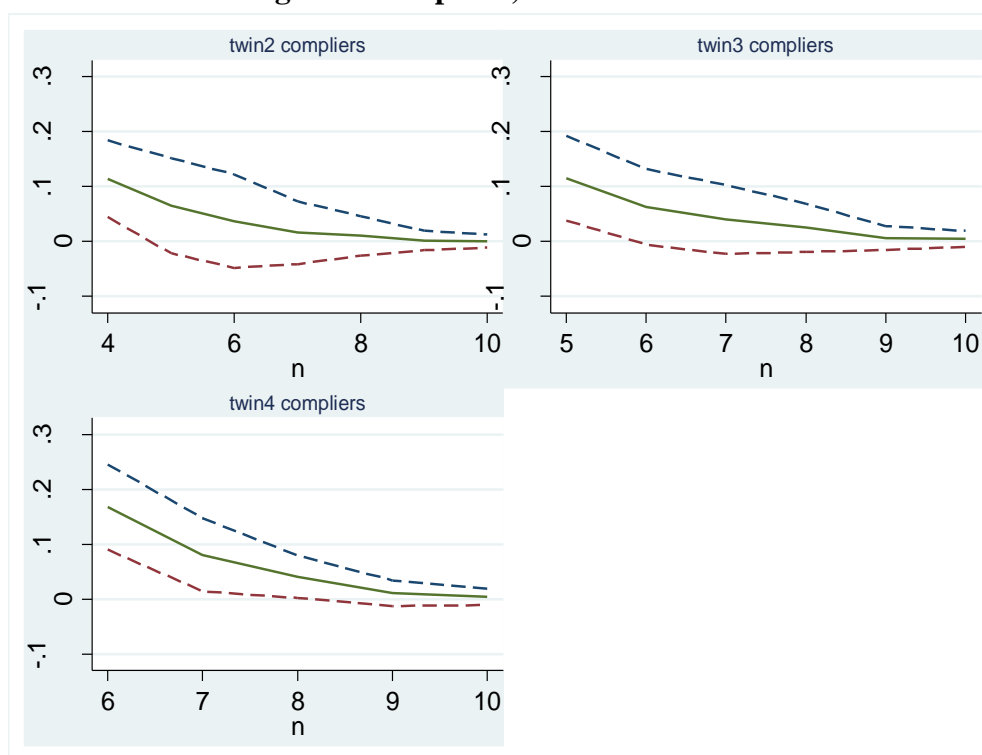
<b>Second-Borns</b>			
Family size	n/a	0.603**	0.657**
		[0.044]	[0.039]
Observations		119402	94261
F test		185.66	280.69
<b>Third-Borns</b>			
Family size	n/a	n/a	0.442**
			[0.063]
Observations			36621
F test			48.73
Sample	2+	3+	4+

Notes: Dependent variable is family size. All regressions control for the socio-economic variables listed in Table 1. \* Denotes statistical significance at the 1% - 5% level. \*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level are in parentheses.  
2+ indicates households with 2 or more children, etc; twin2 indicates twins at 2<sup>nd</sup> birth, etc.

We also see from this that the proportion of compliers is higher for the twins than for the all-female instrument. However, when we decompose the first stages further (see Figure 4) we see that the range of fertility variation induced by twins is narrower than for the all-female instrument.<sup>33</sup> We also see that upwards of 10% of households with twin births go on to have one additional birth even after the twin birth. The fact that we are picking up different sets of compliers is useful, to show us how robust findings are if estimated on a different subpopulation, and to give a sense as to how representative our main findings are likely to be.

<sup>33</sup> Thus, in line with Angrist et al (2010), we see that while both instruments have a strong effect on the propensity to have an additional child, the sex composition instrument induces an effect on a wider range of family size than the twin births instrument.

**Figure 4 Compliers, twin instruments**



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

Turning now to the TSLS estimates, we see from Table 12 that - as for our main specification - the majority of estimates are not statistically different from zero. There are a few exceptions - in two instances we find negative effects on schooling (enrolment and primary school completion) for first-borns, but the vast majority of estimates for first-borns show no significant effects. Similarly, in just one instance we see a positive effect of family size on schooling (years completed) for second-borns. However overall, we conclude from this robustness exercise that there is no or extremely weak evidence of a quantity-quality trade-off for poor families living in rural Mexico.

**Table 12 Effects of family size on education, twins instrument, females**

	First-borns			Second-borns		Third-borns
Instrument →	Twin2	Twin3	Twin4	Twin3	Twin4	Twin4
Outcome ↓						

<b>School enrolment</b>						
Family Size	-0.021	-0.047**	0.006	-0.016	-0.034	-0.032
	[0.019]	[0.018]	[0.020]	[0.023]	[0.025]	[0.055]
Observations	221196	187616	137002	119402	94274	36605
<b>Years of schooling</b>						
Family Size	0.023	0.057	0.01	0.173+	-0.028	0.057
	[0.078]	[0.080]	[0.087]	[0.094]	[0.105]	[0.200]
Observations	219970	186633	136334	118762	93775	36396
<b>Primary school</b>						
Family Size	0.013	0.027	-0.038*	0.032	-0.021	0.016
	[0.017]	[0.017]	[0.019]	[0.020]	[0.024]	[0.054]
Observations	220060	186707	136374	118769	93782	36397
<b>Lower secondary</b>						
Family Size	-0.033	0.02	-0.004	0.051	0.015	n/a
	[0.023]	[0.020]	[0.017]	[0.035]	[0.022]	
Observations	151834	129351	97012	55010	44959	
Sample	2+	3+	4+	3+	4+	

*Notes:* Control for socio-economic variables listed in Table 1. LPM estimates not shown, but all are of magnitude similar to those in tables 4 and 5.

\* Denotes statistical significance at the 1% - 5% level. \*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level are in parentheses.

## 5.1.4 Discussion

These findings raise at least three additional questions. The first is, given that they identify a local average treatment effect, just how representative are they for the population (in our survey) at large? The second is whether the lack of variability in treatment effects across instruments is due to similarities in compliant sub-populations, or whether it is evidence of lack of heterogeneity in the effects of family size in the population. Finally, given the lack of adverse effects of family size on children's education, a natural question is to what extent families are adjusting on other margins, in particular using mother's labour supply.

### 5.1.4.1 Characterising Compliers

We address the first two of these questions using Table 13: whilst compliers are not an identifiable subpopulation, the table describes them in relation to the general population. This is in the spirit of Angrist and Imbens, 1995. It shows, separately for each instrument, how the two sets of compliers compare to the general population in terms of the observed characteristics listed in the left-hand column of the tables. 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 parental education, mother's age, household head occupation, and measures of household wealth including dummy variables for asset ownership. A number of interesting features emerge from the table. First, compliers of the all-female instrument are relatively better off than the population in our survey at large: they include parents from considerably more educated backgrounds compared to the general population.<sup>34</sup> They are also relatively more likely to own most of the listed assets. Twin compliers, on the other hand, tend to be more similar to the population at large, though twin2 compliers are relatively less educated. Whilst the comparison of asset ownership suggests that they are also better off than the population, they are generally more similar to them than are the all-female compliers. Another interesting feature is that there is considerable variability in complier sub-populations within instrument: compare for instance parental education across the 3 sets of all-female compliers. This is useful in showing that the effects pertain to a range of different groups, yet all point in the same direction.

---

<sup>34</sup> 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  $\frac{3}{4}$  as likely to have a non-qualified father compared to the overall population.

**Table 13 Characteristics of compliers vs. entire population**

Column A	Ratio of 1 <sup>st</sup> stage for sub-sample listed					
	in column (A) to overall 1 <sup>st</sup> stage					
	ff	fff	ffff	twin2	twin3	twin4
Father's education:						
No qualification	0.741	0.587	0.465	1.259	0.910	1.003
Some primary	0.867	0.923	0.830	0.976	1.07	1.061
Min completed primary school	1.386	1.286	1.492	0.879	0.943	0.899
Mother's education:						
No qualification	0.719	0.869	0.610	1.193	0.861	1.071
Some primary	1.019	0.931	1.124	0.925	1.018	0.934
Min completed primary school	1.288	1.251	1.155	0.961	1.024	1.054
Mother age 35+	0.837	0.841	0.982	1.084	1.018	1.008
Head works in agriculture	0.991	0.989	1.041	1.006	0.978	0.988
Head indigenous	0.852	0.876	1.005	0.819	0.942	0.979
<b>Utilities:</b>						
Availability of water in house	1.223	1.337	1.517	1.111	0.946	0.956
Availability of light/electricity	1.049	1.065	1.042	0.983	1.017	0.991
Has own toilet	1.100	1.109	0.990	0.973	1.039	1.000
<i>Has water in toilet</i>	<i>1.378</i>	<i>1.356</i>	<i>1.815</i>	<i>1.070</i>	<i>1.093</i>	<i>1.164</i>
<b>Asset ownership:</b>						
Blender	1.245	1.113	1.085	1.104	0.994	1.077
Fridge	1.219	1.217	1.107	1.137	1.229	1.079
Gas stove	1.264	1.129	1.218	0.999	1.061	0.990
Radio	1.066	1.058	1.040	1.022	0.993	1.007
<i>Gas heater for water</i>	<i>1.021</i>	<i>1.026</i>	<i>0.972</i>	<i>1.018</i>	<i>1.033</i>	<i>0.996</i>
<i>Record player</i>	<i>1.306</i>	<i>1.034</i>	<i>1.314</i>	<i>1.186</i>	<i>1.202</i>	<i>1.090</i>
TV	1.140	1.142	1.012	1.047	1.009	1.108
<i>Video</i>	<i>1.321</i>	<i>1.182</i>	<i>1.390</i>	<i>1.142</i>	<i>0.954</i>	<i>1.098</i>
<i>Washing machine</i>	<i>1.042</i>	<i>1.155</i>	<i>0.760</i>	<i>1.219</i>	<i>1.099</i>	<i>0.799</i>

Fan	1.006	1.071	1.555	1.373	1.284	1.073
<i>Car</i>	<i>1.607</i>	<i>1.271</i>	<i>2.053</i>	<i>1.444</i>	<i>0.847</i>	<i>1.078</i>
<i>Truck</i>	<i>1.154</i>	<i>1.323</i>	<i>2.067</i>	<i>1.087</i>	<i>1.023</i>	<i>1.006</i>
Land for agric/forestry	1.006	0.962	0.900	0.940	1.032	1.017
# animals	0.920	0.801	0.810	1.002	0.946	1.119

Notes: sample is first-born females; results for other females similar and not shown. Italicised items are those for which  $\leq 10\%$  of the population own one.

We conclude the following from this analysis. First, comparing the complier sub-populations to the population at large we see that the all-female compliers tend to be relatively better off than the population, whereas the twin compliers are much more similar. So the LATE effects identified in this paper using twin births as instruments are indeed informative about effects for the population at large. Second, as just noted, we do observe differences in complier sub-populations for both instruments (and even within instruments), suggesting that the lack of any effect of family size on children's education for either instrument considered may indeed reflect a lack of heterogeneity of effects in the population of rural indigent households in Mexico.

#### 5.1.4.2 Investigating Channels

We have found no evidence in this paper that family size affects the schooling of girls: we conclude this discussion with an investigation of 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 Wolping, 1980; Angrist and Evans, 1998). We here investigate the extent to which mothers are working more in large versus small families. 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 LPM estimates in Table 14 that, in line with previous work, mothers with large families work less than those with small families. However, the IV estimates show the opposite: in 2 out of 3 cases, 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 health investments, found to be important by Agüero and Marks (2008), though beyond the scope of this present study.

**Table 14 Effects of family size on mother's labour supply**

	LPM	IV	LPM	IV	LPM	IV
<b>Instrument →</b>	n/a	ff	n/a	fff	n/a	ffff
<b>Outcome ↓</b>						
<b>Mother's work</b>						
Family Size	-0.004** [0.001]	0.058** [0.010]	-0.006** [0.001]	0.053** [0.015]	-0.008** [0.001]	0.0232 [0.0251]
Observations	220534		93560		35168	
Sample	2+		3+, ff=1		4+, ffff=1	

Notes: Control for socio-economic variables listed in Table 1. Note also that using the  $n^{\text{th}}$  birth as an instrument, we condition implicitly on the sex of the first  $n-1$  births.

\* Denotes statistical significance at the 1% - 5% level. \*\* Denotes statistical significance at the 1% level or less. Standard errors clustered at the village level are in parentheses.

## 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 using a succession of female-only births as its main source of exogenous variation in family size. The paper exploits extremely large samples and high fertility rates to consider a range of instruments for family size across different birth orders and schooling outcomes. All findings point in the same direction - the negative correlation between family size and schooling observed in the data disappears when we allow for the endogeneity of family size. These

findings are robust to another instrument commonly used in the literature, twin births. This suggests that though they are local average treatment effects, they may be generalisable to other sub-populations. Indeed, when we look at the characteristics of compliers we find that both instruments identify different complier sub-populations, suggesting that the lack of any effect may apply more generally to the population of rural indigent households in Mexico.

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 (2008). 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 for our findings of no effect is 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, found to be important by Agüero and Marks (2008). Investigation of these margins is unfortunately outside the scope of this study, though it is an important agenda for future work.

## Bibliography

Agüero, J., and M. Marks. 2008. "A New Approach to Estimate the Effect of Family Size on Child Development." *mimeo*, University of California, Riverside.

Angelucci, M., De Giorgi, G., Rangel, M., and I. Rasul. 2010. "Family Networks and Schooling Outcomes: Evidence from a Randomised Social Experiment." *Journal of Public Economics*, 94(3-4): 197-221

Angelucci, M., De Giorgi, G., Rangel, M., and I. Rasul. 2009. "Insurance and Investment within Family Networks." *mimeo*, University College London.

Angrist, J., 2004. "Treatment Effect Heterogeneity in Theory and Practice", *The Economic Journal*, 114(March): C52-C83.

Angrist, J., and G. Imbens. 1995. "Two Stage Least Squares Estimates of Average Causal Response Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90(430): 431-442.

Angrist, J., Imbens, G., and D. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434): 444-455.

Angrist, J. and W. Evans. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, 88(3): 450-77.

Angrist, J., V. Lavy and A. Schlosser. 2011. "Multiple Experiments on the Causal Link between the Quantity and Quality of Children." Forthcoming *Journal of Labor Economics*.

Attanasio, O., DiMaro, V., Lechene, V., and D. Phillips. 2009. "The Welfare Consequence of Increases in Food Prices in Mexico and Colombia", *mimeo*, UCL.

Becker, G. 1960. "An Economic Analysis of Fertility." In *Demographic and Economic Change in Developed Countries*. Princeton, NJ: Princeton University Press and NBER.

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

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

Behrman, J., Sengupta, P., and P. Todd. 2005. "Progressing Through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico. *Economic Development and Cultural Change*, 54(1): 237-275.

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

Black, S., P. Devereux and K. Salvanes. 2007. "Older and Wiser? Birth Order and IQ of Young Men." IZA Discussion Paper No. 3007.

Black, S., P. Devereux and K. Salvanes. 2010. "Small Family, Smart Family? Family Size and the IQ Scores of Young Men." *Journal of Human Resources*, 45(1): 33-58.

Cáceres-Delpiano, J. 2006. "The Impacts of Family Size on Investment in Child Quality." *Journal of Human Resources*, 41(4): 738-754.

Conley, D. and R. Glauber. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *Journal of Human Resources*, 41(4): 722-737.

Kessler, D. 1991. "Birth Order, Family Size, and Achievement: Family Structure and Wage Determination." *Journal of Labor Economics*, 9(4): 413-426.

Lee, J. 2008. "Sibling Size and Investment in Children's Education: An Asian Instrument." *Journal of Population Economics*, 21(4): 855-875.

Li, H., J. Zhang and Y. Zhu. 2007. "The Quantity-Quality Tradeoff of Children in a Developing Country: Identification using Chinese Twins." IZA Discussion Paper No. 3012.

Mogstad, M., and M. Wiswall (2010). "Testing the Quantity-Quality Model of Fertility: Linearity, Marginal Effects, and Total Effects." *Mimeo*, Department of Economics, New York University.

Nevo, A., and A. Rosen. 2008. "Identification with Imperfect Instruments." *CEMMAP Working Paper CWP16/08*.

Qian, N. 2009. "Quantity-Quality and the One Child Policy: The Only Child Disadvantage in School Enrolment in Rural China", NBER Working Paper w14973.

Rosenzweig, M. and K. Wolpin. 1980. "Testing the Quantity-Quality Fertility Model: Using Twins as a Natural Experiment." *Econometrica*, 48(1): 227-240.

Rosenzweig, M. and K. Wolpin. 2000. "Natural "Natural" Experiments in Economics." *Journal of Economic Literature*, 38(4): 827-874.

Rosenzweig, M. and J. Zhang. 2009. "Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy." *Review of Economic Studies*, 76(3): 1149-1174.

Schultz, T.P. 2004. "School Subsidies For the Poor: Evaluating the Mexican Progresa Poverty Program", *Journal of Development Economics*, 74(1): 199-250.

Schultz, T.P. 2005. "Population Policies, Fertility, Women's Human Capital and Child Quality." In *Handbook of Development Economics*, vol 4., ed. T.P. Schultz, J. Strauss. Amsterdam: North Holland.

Zajonc, Robert B. 1976. "Family Configuration and Intelligence." *Science*, April 1976, 192:227-36.