A New Approach to Estimate the Effect of Family Size on Child Development*

Jorge M. Agüero[†] **Mindy S. Marks**[‡]

November 2008

Abstract

We introduce a novel instrument to test for the quantity-quality trade-off in child development. Specifically, we exploit exogenous changes in family size generated by infertility shocks after the first child is born. We present evidence that an indicator variable for the infertility status of women is a plausible instrument for family size and show that infertility is not correlated with the child's or the mother's background characteristics and has a robust impact on family size. Using data from Demographic and Health Surveys in Latin America we evaluate the possible trade-off for different aspects of child development including education and health. Our results indicate that, after instrumenting for the number of children using the mother's infertility status, there is a negative and significant relationship between the quantity and quality of children for health indicators (weight-for-age and breastfeeding practice) but not for education.

Keywords: Quantity-quality trade-off, child development, fertility, Latin America.
JEL codes: I31, J13, O15, N36.

*We would like to thank seminar participants at the University of Southern California, the University of KwaZulu Natal, the NBER's Program on Children and the NEUDC 2008 meetings for their comments and suggestions. All remaining errors are ours.

[†]Department of Economics, University of California, Riverside. Email: jorge.aguero@ucr.edu.

[‡]Department of Economics, University of California, Riverside. Email: mindy.marks@ucr.edu.

"To put it simply: the greatest single obstacle to economic and social advantage of the majority of the peoples in the underdeveloped world is the rampant population growth."
-Robert S. MacNamara[1].

1 Introduction

It has been suggested that high levels of fertility are behind the low levels of human capital investment observed in children from developing countries. For example, a recent report by the UN Population Fund (UNFPA 2002) assigns an important role to family planning policies as a tool to achieve the Millennium Development Goals (see Table 1 below). The idea of a possible trade-off between the quantity and the quality of children is shared by other social scientists as well[2].

The influential paper by Becker and Lewis (1973) provides theoretical support for these claims. Although prevalent, the negative correlation between family size and child development observed in cross sectional micro (or macro) datasets can not be taken as conclusive evidence of a quantity-quality trade-off. Even after controlling for important characteristics of the children and their families the association could be spurious due to reverse causality (i.e., families might decide to have more children based on the "quality" of the first child.) Additionally, omitted factors could influence both family and child outcomes. Investigations of the causal evidence for the possible trade-off between quantity (family size) and quality (child development) are crucial to validate or not the claims made by policy makers and the prediction of theoretical models.

[Table 1 about here.]

Several papers have attempted to find exogenous sources of variation in family size as a way to test for this trade-off. Starting with Rosenzweig and Wolpin (1980), twinning has

[1]MacNamara, R. S. (1974) "The World Bank Perspective on Population Growth" in McCoy, T. L., editor, *The Dynamics of Population Policies in Latin America*, Ballinger Publishing Co.

[2]See Schultz (1997, 2008).

been used as an exogenous source of variation in family size by comparing multiple births versus singleton. More recently Black, Devereux, and Salvanes (2005) and Cáceres-Delpiano (2006) as well as Black, Devereux, and Salvanes (2007) and Li, Zhang, and Zhu (2007) have followed the same strategy. Other papers have exploited the sex-composition of the first two children and argue that a family with two boys (or two girls) has a higher probability of trying for a girl (boy) and so they are more likely to have a third child (Angrist, Lavy, and Schlosser 2006). All these studies, with the exception of Rosenzweig and Wolpin (1980) and Rosenzweig and Zhang (2006), tend to find no support for the quantity-quality trade-off hypothesis, at least at the "relevant" margins[3].

Rosenzweig and Zhang (2006) criticize the existing approaches. First, they show that twins have a lower birth weight (endowment). Thus if one does not control for birth weight, twinning is an inappropriate instrument because poorly endowed twin siblings can directly impact the outcomes of the first-born child if parents reallocate resources in response to initial endowments. Additionally, the authors provide evidence of household economies of scale for twins and children of the same gender. This suggests that households with same-sex children may experience a lower loss in resources compared to those with mixed-sex children which may translate into higher child quality. As noted by Angrist, Lavy, and Schlosser (2006) and others, parents with preferences for a mixed sex-composition of their offspring could also have unobserved preferences about the way to raise their children. This could constitute a violation of the exclusion restriction or at least, imply a very local average treatment effect.

This paper introduces a new source of exogenous variation in family size that does not suffer from the above problems. Specifically, we exploit exogenous changes in family size generated by infertility shocks and tests for the quantity-quality trade-off using demographic

[3]Cáceres-Delpiano (2006) finds no impact of an additional child on education attainment. He does find that additional siblings lower the probability of going to private school and reduces the mother's labor force participation.

and health data from Latin America. The medical literature defines infertility as the failure to conceive after a year of regular intercourse without contraception. In this paper we focus on secondary infertility, that is, failure to conceive after having at least one child. We show that secondary *infertile* women (those who became infertile after $K > 0$ children) tend to have fewer children than their *fertile* counterparts (those still fertile after having K children). For example, for households with first-born between 6 and 14 at the time of the survey, secondary infertile women have 1.9 children while the full sample average is 2.7.

Also, one important feature of the surveys used in our paper is that they contains information about mother's preferences for the number and gender composition of their children. Using this information we show that secondary infertile women do not differ in these margins when compared to their fertile counterparts.

We provide evidence from the medical literature showing that the main predictor for infertility is age. We complement this by showing that infertile women do not differ in their fertility preferences and place of birth compared from their fertile counterparts. We also show that their first-born children have the same predetermined characteristics as those born to fertile women, when we consider variables such as birth weight, survival rates, access to vaccines among others. Thus, we argue that (secondary) infertility mimics an "experiment" in which nature assigns, to each woman, a positive upper bound for the number of children, independent of background and preferences.

Our approach differs from the existing literature in several important ways. After having the first child, infertility can occur after any parity $K \geq 1$. Thus, we can test for the quantity-quality trade-off at any (positive) parity. Papers that relay on twinning or sex-composition must limit their analysis to families with at least two children.

Second, instruments such as twins at parity K and sex-composition only affect families whose desired fertility was equal or less than to K but not those with preferences for a family size greater than K. On the other hand, becoming infertile after the K-th child

affects all families with desired fertility larger than K so the set of "compliers" generated by this new instrument is different than the one generated by twins or sex-composition. Thus, by looking at a different set of compliers we can learn more about the possible existence of quantity-quality trade-off.

Third, we look at poorer countries compared to most of the literature and are the first to look at the Latin American case. We expect the Latin American countries included in our analysis to be constrained by the scarcity of resources and thus more likely to face a quantity-quality trade-off. In this regard, Qian (2005), Rosenzweig and Zhang (2006) and Li, Zhang, and Zhu (2007) are closer to our paper as they focus on China. However, China's one-child policies (and their relaxation allowing for two children) imposed an enormous restriction on fertility that might not be representative of developing countries at large. Finally, the richness of the data used in this paper allow us to evaluate the trade-off by looking at health-related outcomes in addition to the education variables that are common in the literature[4].

For education-related outcomes we find no causal relationship between additional siblings and decreased school attendance or grade advancement of the first born. However, we do find that the exogenous change in family size bought about by the onset of secondary infertility matters for children's health. In particular, additional children reduce the first-born child's age-adjusted weight percentile and also reduce the likelihood of breastfeeding.

The next section introduces the topic of infertility by presenting the definition and the evidence from the medical literature on this topic. The Demographic and Health Surveys used in this paper are described in section three followed by the econometric methods applied in this paper. In section five, we present evidence that women's infertility status is a valid exogenous source of variation for family size. The main results and robustness tests are introduced in section six. Section seven closes the paper summarizing the findings and outlining the agenda for future research and policy considerations.

[4]Most papers rely on population census data and therefore have a limited number of outcomes to explore.

2 Background information about infertility

The medical literature defines infertility as the failure to conceive after a year of regular intercourse without contraception (Chandra 1994). Infertility can be further broken down into primary infertility, which describes women who have never been able to conceive, and secondary infertility, describing those who have had at least one successful pregnancy, but are unable to achieve future pregnancies. In this paper we will focus on secondary infertility as women need to have at least one child in order to investigate the quantity-quality trade-off. However, for simplicity, throughout the paper we will use the term "infertility" even though we just refer to the secondary case[5].

Women's ability to have children is highly heterogeneous. It is well established that infertility increases with women's age (e.g., Dunson, Baird, and Colombo 2004, Buck, Sever, Batt, and Mendola 1997). Determining the sources of such heterogeneity among women remains a challenge for fertility research (Weinberg and Dunson 2000). Beyond age, the medical literature is not in agreement about what other factors, if any, influence infertility. For example, in an article summarizing the epidemiological literature on the role of life-style factors (cigarette smoking, alcohol and caffeine consumption, exercise, BMI, and drug use) on female infertility Buck, Sever, Batt, and Mendola (1997) conclude that "[f]ew risk factors have been assessed or identified for secondary infertility." Additionally, Negro-Vilar (1993) reviews the role of stress and other environmental factors affecting infertility in men and women. He concludes that consistent and systematic methods are needed to properly asses the environmental influences on human reproductive health. Until then, he argues, the cases of female (and male) infertility will remain unexplained.

There is some evidence suggesting that previous use of birth control as well as indicators of poor health such as sexually transmitted diseases, smoking, drinking, extreme body mass index (BMI), and miscarriages are associated with infertility (e.g., Grodstein, Goldman, and

[5]See Agüero and Marks (2008) for a study where both definitions of infertility are used.

Cramer 1994, Hassan and Killick 2005). However, the majority of this evidence is suspect as it comes from couples recruited for prospective studies. That is, women who identify themselves as having difficulties conceiving are recruited for a study and their observable characteristics are then correlated with their time to pregnancy. This type of study design may produce a spurious association.

For example, assume that drinking is not a factor associated with infertility. In other words, drinkers and non-drinkers are equally distributed between low and high levels of infertility. However, consider the case where drinking is associated with riskier behaviors including having unprotected sex. Drinkers with low levels of infertility may have had all of their desired pregnancies through unintended conceptions compared to their non-drinker counterparts (with the same low infertility levels). This implies that in a prospective study on infertility that only recruits women having difficulties conceiving, the sample will have more low-infertility non-drinkers than low-infertility drinkers (high-infertility women will be equally distributed by drinking pattern). After some time, those who have low levels of infertility will be able to have babies and the pregnancy rates will appear to be higher for non-drinkers than for drinkers even though both drinking groups were equally distributed among infertility levels in the first place. Thus, the sample selection could create a spurious correlation relating infertility to variables that are not truly associated with it (see Weinberg and Dunson 2000, Tielemans, Burdorf, te Velde, Weber, van Kooij, and Heederik 2002).

Additionally, there is some evidence that infertility appears to be independent of the background characteristics of infertile women. For example, characteristics such as the women's mother's BMI status, height and smoking habits cannot predict whether these women will become infertile. Similarly, variables such as father's social class and parity are unrelated to observed heterogeneity in fertility (Joffe and Barnes 2000). In section 5, using our data, we present new evidence where secondary infertility is not correlated with "pre-determined" or background characteristics of a woman and her first-born child. The next section describes

the surveys we use in our analysis.

3 Demographic and health surveys

In this paper we use cross-sectional data from the Demographic and Health Surveys (DHS) in Bolivia (conducted in 1994 and 1998), Brazil (1996), Colombia (1995 and 2000), the Dominican Republic (1996), Guatemala (1998), Nicaragua (1998), and Peru (1996.) We select these Latin American surveys because the questions related to infertility are exactly the same[6] The DHS are standardized nationally representative household surveys in developing countries[7]. Women between the ages of 15 and 49 answer questions about their birth history, current and future contraceptive use, fertility preferences, their socio-economic, marital and health status, among other topics.

These surveys allow us to identify self-reported infertility in two ways. The first is when women mentioned sub-fertility or infertility as their reason for not using contraceptives at the time of the survey (*Infertility 1*). Additionally, when non-sterilized women are asked about their desire for future children, we define *Infertility 2* as a binary variable for capturing those women who responded that they were unable to have more children. We define a woman as infertile in either of these cases. Thus, in keeping with the medical definition of infertility, we can only identify infertility for non-sterilized women who are not currently taking contraceptives. These women constitute more than 40% of original the sample. We exclude from the final sample children whose mothers were using contraceptives as well as those who were sterilized at the time of the survey. Restricting our sample to a non-random subgroup of the population affects the extrapolation of the results. However, the subpop-

[6]Later and earlier surveys are available for a subset of countries and for others not included in our sample (e.g., Mexico, Paraguay, Haiti). However, at least one of the questions related to infertility, if not all, are not asked in the same way in those additional surveys.

[7]Further explanations about the data and sampling strategies can be found in the DHS website (http://www.measuredhs.com/). The data can be downloaded from their website after registering and submitting a short proposal for the project.

ulation under study, women not actively controlling the fertility, are the group most likely to be targeted by family planning programs. Hence, the gains from the sample restriction (testing the trade-off for this subgroup) are likely to be higher than the cost in terms of lack of national representability of the results.

Relevant to this paper is the information about child development contained in the DHS. We will measure child development through education and health outcomes for children under the age of 15. For all children older than six (or five, depending on the country and year) the household roster collects information about their education. However, it is only possible to match, with certainty, the roster information with the women's questionnaire for those children under the age of 15.[8] Thus, when using education as our "quality" indicator we limit our study to families where the first-born child is between the ages of 6 and 14 at the time of the survey. Two separate education questions are available from the roster. First, we know if the child is currently attending school (recorded as a binary variable) and second, the highest grade ever passed is recorded. With the latter, DHS constructs a measure of years of schooling taking into account the differences in school systems by year and country. However, raw comparisons of years schooling across countries and cohorts could be misleading as children could be ahead (or falling behind) of the rest of their cohort-nation group. Hence, we standardize years of schooling as follows.

Let $s_{i,n,c}$ be the years of schooling of child i from nation n born to cohort c, where cohorts are defined by the year and quarter of birth. The standardized years of schooling are computed as following

$$\tilde{s}_{i,n,c} = \frac{s_{i,n,c} - median(s_{i,n,c})}{\sigma(s_{i,n,c})} \tag{1}$$

where $\sigma()$ represents the standard deviation of the schooling for all children born in cohort c

[8]However, for Bolivia in 1994, the questionnaire does not allow us to match roster and women's questionnaire at all. Hence we do not include this survey for the education outcomes.

and nation n and $median()$ is the corresponding 50th percentile for the pair cohort-nation. A value of zero for $\tilde{s}$ implies that the child's years of schooling is the same as the median of her cohort-nation group. A positive value reflects a child with more years of schooling than the median while a negative value characterized those lagging behind the median, relative to the standard deviation of the cohort-nation group[9].

The DHS also asks mothers a series of question about their children. Women respond to questions about the health status of their children born in the five years previous the survey. In particular, questions about prenatal care, the delivery of the baby and breastfeeding practices are available. Also, anthropometric measures are taken for children present in the household at the time of the survey. Specifically, we have information about the age-adjusted height and weight distributions for these children as well as information about birth weight[10]. Hence, when focusing on health outcomes, we are limited to households where the first-born child is at most five years of age at the time of the survey.

To summarize, the sample for our analysis is composed by (non-twin) first-born children from households where the eldest resident child was at most five years of age when considering health outcomes and 6 to 14 when considering education-related outcomes. In addition, we restrict our samples to children whose mother's age at first birth was between 15 and 40 year of age and were not using contraceptive, nor sterilized at the time of the survey.

The sample size for children between 6 and 14 years of age is 6,035. For children under the age of six the sample size varies depending on the outcome variable as some children where not present and lack anthropometric measures. The biggest sample size is 7,154 and the smallest is 6,153 when the mother's BMI is included in the regression. The summary statistics of these samples are described in Table 2 below.

[9]Cáceres-Delpiano (2006) uses a similar methodology to transform years of schooling but it does not control for quarter of birth, just year and location.

[10]We are missing birth weight for a subset of our sample. Many children were born at home and did not have their birth weight recorded.

[Table 2 about here.]

The average child between 6 and 14 has a 0.88 probability of attending school and tends to have the same years of schooling as the median child in her country of birth and cohort. The average number of children at home is 2.7, including the first-born. Around 7% of the women in this sample are (secondary) infertile and the average women was over 21 years old when she had her first child. For the sample of children under 6 years of age, the average child is 1.7 years old, her age-adjusted height and weight are around the 28 and 38 percentile, respectively. Mother's for this sample are younger and as, expected, the proportion that is infertile after the first birth is lower than in the older sample. In both samples, the population is split evenly between boys and girls, most children live in urban areas and around a quarter of the sample comes from the Peruvian survey.

4 Methodology

The main specification to test for a trade-off in the quantity and quality of child development is given by the following equation

$$y_{i,j} = \alpha + \beta K_j + \gamma' CHILD_{i,j} + \delta' FAMILY_j + e_{i,j} \tag{2}$$

where $y_{i,j}$ represents a measure of child development for the first-born child i in family j. K_j is the variable of interest and captures child i's number of siblings, including i herself, as common in the literature. Thus, a negative value for β reflects a trade-off between quality ($y_{i,j}$) and quantity (K_j). We include two sets of control variables. First, the vector $CHILD$ contains child-specific characteristics such as age and sex. The vector $FAMILY$ contains information about the family including mother's age at first birth (in the form of grouped binary variables), mother's education, marital status, location (urban-rural) in addition to survey (country-year) fixed-effects.

Due to the data structure, when the first-born child i is between 6 and 14 the outcomes variables are related to education. We will consider school attendance (as a binary variable) and standardized years of schooling as described in the previous section. For first-born children under the age of six, $y_{i,j}$ will measure health-related variables such as age-adjusted percentiles for weight and height as well as breastfeeding status.

OLS estimates of β in equation (2) could be biased because family size is a choice variable for the household. To avoid this problem we instrument K_j with the infertility status of the mother. Let $Infertility_j$ be equal to one if the mother is infertile as defined in section 3 and zero otherwise. Thus, in the first stage we estimate the following equation

$$K_j = \alpha_1 + \pi Infertility_j + \lambda' CHILD_{i,j} + \phi' FAM_j + \nu_{i,j} \tag{3}$$

We will use the results from equation (3) to estimate the impact of family size on the quality of the children as generated by our proposed exogenous source of variation, infertility. As also discussed in the previous section, the infertility status of a women comes from two separate survey questions. When using the union of these questions will refer to the IV estimation and we will use the term GMM when each question is used as a separate source of identification[11].

The validity of the use of infertility status as an exogenous source of variation for family size depends on the assumptions made about the strength of the relationship between the variable *Infertility* and K as well as the exclusion restriction assumption regarding the lack of correlation between infertility and the unobserved variables captured by e_{ij} in equation (2). The next section explores these issues.

[11]GMM estimates use the continuously updated procedure with a constant weight matrix.

5 Exogenous source of variation in family size

In this section we discuss and present evidence to validate our approach using the infertility status of a women, after her first child, as an exogenous source of variation in family size. We first provide evidence in favor of the exclusion restriction and then show that our instrument alters significantly the number of children a women does have.

5.1 Infertility as a shock

The validity of our instrument relies on the, untestable, exclusion restriction. This restriction requires women's infertility status to be uncorrelated with all possible unobservable variables included in e_{ij} from equation (2). Here we argue in support of this assumption. We already mentioned in section 2 that the medical literature shows that infertility increases with mother's age but the evidence is less clear with regards to other important variables, including health. Also, the work by Joffe and Barnes (2000) suggests that the background characteristics of infertile women are not different from the corresponding background characteristics of their fertile counterparts.

The richness of our datasets allow us to test whether predetermined observable characteristics of first-born children from later-to-be infertile mothers are different from those born to still-fertile mothers. By predetermined characteristics we mean those that took place before a women could have realized here infertility status. To understand our strategy, consider the situation where a researcher argues about the validity of a randomized experiment. In such a case, the researcher will show how treatment and control groups do not vary when considering a set of predetermined characteristics obtained from a baseline survey as a way to show no systematic bias in the allocation of the treatment under study.

Our approach is an analog to this strategy. Infertility is not a random event –it increases with age– but we can test whether it correlates with some key predetermined observable

characteristics of first-born children[12]. This is close to the approach taken by Rosenzweig and Zhang (2006), who want to argue that mothers giving birth to twins are not different than their non-twin counterparts after controlling for age. This results are shown in Table 3. There we report the estimates for our infertility measure from a series of regressions (indexed by V_{ij}) where in addition to fertility status we control for the women's age at first birth as follows (without an intercept):

$$V_{ij} = \theta_1 \text{Infert}_j + \theta_2(1 - \text{Infert}_j) + \sum_s \rho_s \text{AGEFB}_{js} + \eta_{ij} \tag{4}$$

In essence, these regressions ask, controlling for age at first birth (AGEFB_{js}), whether variable V_{ij} is different for first born children from later-to-be infertile women compare to those born to still-fertile mothers ($\theta_1 - \theta_2 = 0$). Table 3, panel A, shows that many important predetermined variables of first-born children between the ages of 6 and 14 from infertile women mirror those born to fertile counterparts. Both groups have the same mother's place of birth, and mother's fertility preferences. The child's sex and month of birth do not vary by infertility status either. For first-born children between under 6 (Panel B) we have additional child-level characteristics. We observe that the length of labor when about to be delivered, the survival rates, gender, birth weight, size at birth and how parents treat them with respect to early vaccination and initial breastfeeding do not differ by mother's fertility status. Also importantly, when asked about their desired future fertility both groups of mothers show no difference by fertility status.

[Table 3 about here.]

Infertility thus mimics an experiment in which nature assigns, to each woman, an exoge-

[12]This approach is related to the work by Currie and Thomas (1995) where they analyze the effect of Head Start, a US early childhood intervention program for disadvantaged children, on child development. They show that, within a family, all children had the same probability to be taken to a clinic or hospital if sick, regardless of the children's enrollment in Head Start. In this way the authors suggest that within-sibling comparisons could be a valid approach when evaluating the impact of Head Start on children's outcomes.

nous upper bound for the number of children, independent of the woman's background, her preferences and of the characteristics of her first-born child. As expected, we also find that infertility is associated with a history of miscarriage and current obesity status. Importantly though, we do not find that a long-term indicator of women's health (height) is lower for infertile women as one would expect if poor health causes infertility. If anything, our results show that infertile women tend to be taller than their fertile counterparts.

5.2 First stage results

For our instrument to be valid, infertility should also be correlated with the number of children a woman has. Table 4 shows the first stage results for the two samples of children. Infertility is indeed highly correlated with family size. As shown in column (i), a first-born child between the ages of 6 and 14 will, on average, have .78 fewer siblings if his mother experiences secondary infertility. Note that the average child in the sample has 1.7 siblings so this is a sizable effect. Infertility has a larger impact on family size in rural settings and for families with less educated mothers - reducing family size by 1.15 and 1.05 children respectively. In Panel B we present the results of a modified equation (3) where both infertility indicators are included. We show that each indicator has predictive power on the number of siblings. In both panels, the Shea's partial R^2 and the F-tests show that our instruments have sufficient power in all specifications.

[Table 4 about here.]

For the sample of younger children, infertility again reduces family size by a significant amount (see Table 5). A first-born child younger than 5 will, on average, have .23 fewer siblings if his mother becomes infertile. Since the average number of siblings for this subsample is only .3 these results imply that the majority of young children whose mother experience

secondary infertility are only children. Again, the F-tests and Shea's partial R^2 confirm that our instruments have sufficient power in all specifications.

[Table 5 about here.]

The next section presents a test for the quantity-quality trade-off, in equation (2) using infertility as an exogenous source of variation for family size in the context of the Latin American countries described above.

6 The effects of family size

6.1 Impact on children's education

In Table 6 we report the estimated impact of the number of siblings on the probability that the first-born child is attending a school (β in equation (2)) for a sample of children between 6 and 14 years of age. In Panel A, column (ii) we show that an additional sibling is associated with a one percentage point *decrease* in the attendance status of the first-born child. This OLS estimate, while statistically significant, represents a small impact on attendance when compared to the mean value for the sample, as shown in column (i).

However, this result cannot be taken as evidence of a quantity-quality trade-off due to the possible presence of unobserved variables that could simultaneously affect the decision to have children and their school attendance. In column (iii) we show the results after we instrument for the number of siblings using infertility. The estimates show a negative but not statistically significant relationship between the number of siblings and school attendance. While the IV and OLS parameters are close in magnitude the former one is, as usual, a more imprecise parameter. However, as discussed in the previous section we can reject the possibility of weak instruments.

Continuing with panel A, column (iv) shows the estimates when the model is overidentified. Here we instrument using two distinct sources of infertility. As discussed earlier, women report their infertility status when answering questions as to why they are not currently using contraceptives (*Infertility 1*) and when asked about their desire to have an additional child (*Infertility 2*), for those not sterilized. The GMM results confirm the IV findings. There is no statistically significant relationship between the number of siblings and the school attendance of the first-born child. The Hansen J-statistic provides evidence in favor of the validity of the instruments.

As mentioned before, other studies have found similar results. For example, Angrist, Lavy, and Schlosser (2006), Black, Devereux, and Salvanes (2005) and Qian (2005), among others, show that there are not trade-offs between quantity and quality when considering school-related outcomes. However, their use of twinning or sex-compositions as exogenous sources of variations for family size are subject to the criticisms developed by Rosenzweig and Zhang (2006). For example, since twins tend to have lower birth weight than singletons, the authors argue that use of twinning may underestimate the trade-off when parents focus on efficiency and allocate resources to children born before the twins. In our case, by introducing a new instrument based on infertility shocks, we are not affected by these claims. Thus, the lack of trade-off in our findings is less likely to be the result of a underestimation of the trade-off.

[Table 6 about here.]

Our education findings are robust to the addition of health variables as shown in Table 6. Columns (vi) to (viii) from panel A show the result of adding mother's health indicator measure as BMI-based obesity status. The negative and significant association found in OLS estimates does not persist when we instrument for the number of siblings. However, there could be heterogenous effects. Since rural households tend to have less resources than

those in urban areas we could expect the trade-off to be more prevalent in those areas. In Table 6 we split the sample and compare the estimates from urban and rural areas (Panel B). Nonetheless, we find no evidence of a trade-off in rural areas (neither in urban) when considering exogenous changes in family size. Similarly, in Panel C, we divide the sample between households where the mother has high levels of education (i.e., those with more than just completed primary education) versus those with low levels of education (i.e., those who have at most completed primary education.) Again, we do not find causal evidence of a trade-off. For the high-education sample we also observed a change in sign implying a stronger rejection of the trade-off hypothesis.

[Table 7 about here.]

In Table 7 we repeat the exercise but now the outcome variable is the number of standardized years of schooling as defined in equation (1). The same patterns are found here. An additional sibling tends to reduce the years of schooling of the first-born but this finding disappears after instrumenting. Furthermore, our findings are robust to the inclusion of health variables. We do not find an impact of sibship either when splitting the sample into urban and rural or high and low education levels of the mother. Finally, in all our specifications for the overidentified models, the Hansen J-statistic show a strong validity of our instruments.

6.2 Impact on children's health

The section presents results for the sample of first-born children under the age of six. Column (i) of Panel A in Table 8 presents the OLS estimate, which suggests that each additional child in the household lowers the age-adjusted weight of the first-born by 3.2 percentiles. Column (ii) contains the corresponding IV estimate. It suggests that the effect of additional younger siblings on age-adjusted weight, using the variation in family size that comes through the infertility channel, persists. The IV point estimate is large, negative and statistically

significant suggesting that, if anything, the OLS estimate understates the quantity-quality trade-off. This is consistent with the case where unobserved variables, such as wealth or family support, are important factors explaining both child quality and number of children. In column (iii) of Panel A we take advantage of having two separate measures of infertility and use each as an instrument to run a Hansen J-test for overidentifying restrictions. The GMM model confirms our findings and the J-test statistic cannot reject the null hypothesis that our instruments are valid. In this estimation, the result suggests that each additional child reduces the age-adjusted weight of the first born by 20.5 percentiles (about $\frac{2}{3}$ of a standard deviation).

Column (iv) of Panel A, presents the results when current age-adjusted height is the health indicator (HAZ). The results for HAZ differ from the results for age-adjusted weight. The OLS estimates suggest a small quantity-quality trade-off in height. However, the IV and GMM estimates argue that there is no causal relationship between an increase in sibship and height as the coefficient on the kids-at-home variable is small and statistically insignificant. These results are consistent with Rosenzweig and Zhang (2006) who also find a trade-off for weight but not for height when looking at the impacts of twinning at first birth in China. Perhaps deficiencies in height reflect persistent resource deprivation, while current resource shortages impact current weight but not current height.

The final health outcome is an indicator for if the first-born child is currently breastfeeding. The OLS results suggest that having a sibling reduces the probability that the child will be breastfeeding by nine percentage points. The IV estimates are more negative than the OLS estimates but not precisely estimated owning to the less efficiency of IV. The GMM estimates suggest that additional children limit the first-born child's access to breast milk[13]. These results state that a first-born child has a twenty percentage point reduction in the

[13]Clearly the oldest children in this sample are not still breastfeeding. The results are unchanged if we limit the sample to children three years old or younger. Additionally, the breastfeeding results hold if child's age (measured in months) is included as a control instead of child age dummies measured in years.

likelihood of breastfeeding following the birth of an additional child. For both the weight and the breastfeeding outcomes a Hausman test says we cannot reject the null that probability limit of the OLS and the IV are the same.

In Panel B we add information about health (proxied by mother's BMI indicators) to the regressions. It is possible that our infertility measure is capturing poor health on the part of the mother and that mother's poor health could directly influence the health outcomes of her children, thus invalidating our identification strategy. However, the main finding persists when health indicators are added to the model[14]. Additional siblings have a significant and negative impact in the OLS, IV, and the GMM specifications when the health outcome is weight-for-age or breastfeeding. When a child is "removed" from the family the first-born kid is heavier and more likely to be breastfed. However, her height is not affected.

[Table 8 about here.]

7 Conclusions

This paper investigates the effects of family size on child development in a developing country setting. We employ a novel instrument (secondary infertility) that acts as removing a child from a family that desires more children instead of the traditional instruments which add a child to a family that would have stopped with few children. We focus on Latin America and argue that in developing countries parents may be more likely to face binding resource constraints which should result in a quantity-quality trade-off.

For the education outcomes our findings support the existing literature. Specifically, we find no causal relationship between additional siblings and decreased school attendance or performance. Our data also allows us to look at an important and understudied outcome:

[14]We have experimented with adding previous miscarriage as an additional health indicator and the results are unchanged. We have also included an indicator which captures previous birth control use (an indicator of infertility problems) and again the pattern of the results are unchanged.

children's health. We find that the exogenous change in family size brought about by the onset of secondary infertility matters for children health. In particular, additional children reduce the first-born child's age-adjusted weight percentile and also reduce the likelihood of breastfeeding. Rosenzweig and Zhang (2006) is the only other paper we are aware of that includes health outcomes. They too find a quantity-quality trade-off for health outcomes in China using twinning as an exogenous source of variation in family size. However, more evidence for health-related outcomes is needed.

This paper can be extended by comparing by including second-born children. Additionally, one can use twins and sex-composition of children to address the differences of our findings in relation to the currently used methods in the literature. Also, we can further explore for possible heterogenous effects by country and child's gender. On the latter, preliminary evidence suggests that an additional sibling to a first-born girl reduces her years of schooling by approximately 20 percentages point of a standard deviation.

An unanswered question in the quantity-quality literature is along which margins are parents really constrained. It may be that some outcomes (like children's health) are more resource intensive for parents to provide than others (like public education). It may also be the case that parents need time to adjust to fertility shocks. Thus we could observe declines in "quality" when the children are young, that will disappear as parents have time to adjust household resources. Note that the original models for the quantity-quality trade-off do not account for this effect and that this discussion is absent from the policy arena. These issues are important when designing family planning programs and broader interventions that wish to increase child welfare in developing countries.

References

Agüero, J. M., and M. S. Marks (2008): "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks," *American Economic Review*, 98(2), 500–504.

Angrist, J., V. Lavy, and A. Schlosser (2006): "New Evidence on the Causal Link between the Quantity and Quality of Children," IZA Discussion Papers 2075, Institute for the Study of Labor (IZA).

Becker, G. S., and H. G. Lewis (1973): "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, 81(2), S279–88.

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

——— (2007): "Small Family, Smart Family? Family Size and the IQ Scores of Young Men," NBER Working Papers, National Bureau of Economic Research, Inc.

Buck, G. M., L. E. Sever, R. E. Batt, and P. Mendola (1997): "Life-Style Factors and Female Infertility," *Epidemiology*, 8(4), 435–441.

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

Chandra, A. (1994): "Infertility," in *From Data to Action. CDC's Public Health Surveillance for Women, Infants, and Children*, ed. by L. S. Wilcoz, and J. Marks, pp. 63–72. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Currie, J., and D. Thomas (1995): "Does Head Start Make a Difference?," *American Economic Review*, 85(3), 341–64.

Dunson, D., D. Baird, and B. Colombo (2004): "Increased Infertility with Age in Men and Women," *Obstetrics and Gynecology*, 103(1), 51–56.

Grodstein, F., M. Goldman, and D. Cramer (1994): "Body Mass Index and Ovulatory Infertility," *Epidemiology*, 5(2), 247–250.

Hassan, M., and S. R. Killick (2005): "Is Previous Aberrant Reproductive Outcome Predictive of Subsequently Reduced Fecundity?," *Human Reproduction*, 20(3), 657–664.

Joffe, M., and I. Barnes (2000): "Do Parental Factors Affect Male and Female Fertility?," *Epidemiology*, 11(6), 700–705.

Li, H., J. Zhang, and Y. Zhu (2007): "The Quantity-Quality Tradeoff of Children in a Developing Country: Identification Using Chinese Twins," IZA Discussion Papers 3012, Institute for the Study of Labor (IZA).

Negro-Vilar, A. (1993): "Stress and Other Environment factors Affecting Fertility in Men and Women: Overview," *Environmental Health Perspectives*, 101(Supplement 2: Impact of the Environment on Reproductive Health), 59–64.

Qian, N. (2005): "Quantity-Quality: The Positive Effect of Family Size on School Enrollment in China," Mimeo, Department of Economics, Brown University.

Rosenzweig, M. R., and K. I. Wolpin (1980): "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment," *Econometrica*, 48(1), 227–40.

Rosenzweig, M. R., and J. Zhang (2006): "Do Population Control Policies Induce More Human Capital Investment? Twins, Birthweight, and China's 'One Child' Policy," IZA Discussion Papers 2082, Institute for the Study of Labor (IZA).

Schultz, T. P. (1997): "Demand for Children in Low Income Countries," in *Handbook of Population and Family Economics*, ed. by M. R. Rosenzweig, and O. Stark, vol. 1 of *Handbook of Population and Family Economics*, pp. 349–430. Elsevier.

——— (2008): "Population Policies, Fertility, Women's Human Capital, and Child Quality," vol. 4 of *Handbook of Development Economics*, pp. 3249–3303. Elsevier.

Tielemans, E., A. Burdorf, E. te Velde, R. Weber, R. van Kooij, and D. Heederik (2002): "Sources of Bias in Studies Among Infertility Clients," *American Journal of Epidemiology*, 156(1), 86–92.

UNFPA (2002): *State of the World Population 2002. People, Poverty and Possibilities.* United Nations Population Fund, New York.

Weinberg, C., and D. Dunson (2000): "Some Issues in Assessing Human Fertility. Invited Millennial Vignette," *Journal of the American Statistical Association*, 95.

Tables

Table 1: Role of family planning on the Millennium Development Goals

MDG	Effect
Goal 1: Poverty Reduction	Smaller families allow more investments in each child's health and education. Large families dilute the assets of poorer households
Goal 2: Universal Education	Large families numbers of children in poor families means that some children get no education. For others, education may be delayed, interrupted and shortened.
Goal 4: Child Mortality	Higher fertility reduces the provision of health care to children.

Source: UNFPA (2002) page 7.

Table 2: Summary Statistics

Variable	6 to 14 sample Mean	6 to 14 sample Std. Dev.	Under 6 sample Mean	Under 6 sample Std. Dev.
Children's characteristics				
School attendance (=1)	0.883	0.321		
Std. years of schooling	-0.001	1.058		
Breastfeeding (=1)			0.322	0.467
Age-adjusted height (pctile.)			37.87	30.11
Age-adjusted weight (pctile.)			27.65	28.49
Age (years)	9.386	2.719	1.729	1.420
Sex (girl=1)	0.497	0.500	0.492	0.500
No. of siblings	2.692	1.474	1.297	0.536
Mother's characteristics				
Infertility	0.071	0.257	0.021	0.144
Infertility 1	0.046	0.209	0.014	0.118
Infertility 2	0.030	0.171	0.008	0.087
Age at 1st birth	21.45	4.64	20.90	4.32
No education	0.123	0.328	0.083	0.276
Incomplete primary	0.310	0.462	0.272	0.445
Complete primary	0.142	0.349	0.156	0.362
Incomplete secondary	0.193	0.394	0.244	0.429
Complete secondary	0.139	0.346	0.153	0.360
Higher education	0.093	0.290	0.092	0.289
Underweight (=1)	0.023	0.151	0.058	0.234
Normal weight (=1)	0.682	0.466	0.629	0.483
Overweight (=1)	0.225	0.418	0.259	0.438
Obese (=1)	0.068	0.252	0.053	0.225
Location				
Urban (=1)	0.582	0.493	0.561	0.496
Bolivia (1994)			0.066	0.248
Bolivia (1998)	0.163	0.369	0.118	0.323
Brazil (1996)	0.087	0.283	0.078	0.268
Colombia (1995)	0.093	0.290	0.090	0.287
Colombia (2000)	0.088	0.284	0.080	0.272
Dom. Rep (1996)	0.070	0.255	0.089	0.285
Guatemala (1999)	0.099	0.299	0.098	0.297
Nicaragua (1998)	0.150	0.357	0.136	0.343
Peru (1996)	0.250	0.433	0.245	0.430

Note: Sample includes (non-twin) first-born children and those whose mother's age at first birth was between 15 and 40 year of age and were not using contraceptive, nor sterilized at the time of the survey.

Table 3: Women's and Children's Characteristics by Fertility Status

Characteristics (V_{ij})	Infertile (θ_1)	Fertile (θ_2)	$\theta_1 - \theta_2 = 0$
Panel A: 6-14 Sample ($N = 6,035$)			
Mother born in urban area (=1)	0.445 (0.038)	0.395 (0.033)	0.050 [1.85]
Child's sex (Female=1)	0.458 (0.037)	0.470 (0.032)	-0.012 [-0.46]
Child's month of birth	6.737 (0.246)	6.443 (0.213)	0.294 [1.70]
Mother's ideal number of children[d]	2.414 (0.094)	2.506 (0.089)	-0.092 [-1.39]
Mother's ideal number of boys[d]	1.144 (0.135)	1.102 (0.129)	0.042 [0.95]
Mother's ideal number of girls[d]	1.084 (0.116)	1.055 (0.114)	0.029 [0.61]
Mother is obese (=1)*	0.033 (0.016)	0.057 (0.014)	-0.024 [-2.40]
Mother's height (cm)*	152.86 (0.262)	152.29 (0.234)	0.566 [2.84]
Ever had a miscarriage (=1)[c]	0.298 (0.034)	0.204 (0.028)	0.094 [3.92]
Panel B: Under 6 Sample[1/]. ($N = 7,771$)			
Long labor(=1)[a]	-.005 (.020)	.005 (.020)	-.011 [-0.28]
Prenatal doctor(=1)	.863 (.171)	.758 (.117)	0.105 [2.83]
Child is alive(=1)	.785 (.128)	.802 (.128)	-.016 [-0.87]
Child's sex (Female=1)	1.39 (.160)	1.40 (.156)	-.009 [-0.22]
Child's birth weight (in grams)*	3356 (186)	3446 (184)	-89.6 [-1.46]
Child's Size at birth [a]	2.99 (.038)	3.01 (.038)	-.028 [-0.38]
Child received polio vaccine (=1)[b]	.732 (.158)	.752 (.155)	-0.020 [-0.46]
Breastfed 3 months or less (=1)	.499 (.162)	.500 (.158)	-.001 [-0.01]
Mother's ideal number of children[d]	1.93 (.231)	1.90 (.220)	0.037 [0.52]
Ever had a miscarriage (=1)[c]	.374 (.147)	.232 (.143)	0.141 [4.08]
Mother is obese (=1)*	.045 (.077)	-.006 (.075)	0.051 [2.79]
Mother's height (cm)*	153.5 (0.626)	151.9 (0.270)	1.58 [2.76]

Note: Standard errors in parentheses and t-statistics in brackets. Control variables include women's age at first birth. ([*]) Not gathered for all respondents; ([a]) not asked for Colombia in 2000 ; ([b]) not asked in Nicaragua; ([c]) not asked Bolivia (1998). ([d]) Respondents that gave non-numeric answers are excluded.
[1/] Sample includes deceased children who are not part of the main analysis.

Table 4: First Stage Results for Children Between 6 and 14.

Dependent Variable: Number of Children at Home (K_j)						
Sample:	All	Health	Urban	Rural	High educ.	Low educ.
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Panel A: Just-identified model						
Infertility	-0.784	-0.719	-0.612	-1.152	-0.478	-1.046
	(0.053)	(0.052)	(0.054)	(0.126)	(0.056)	(0.087)
Shea's R^2	$(0.172)^2$	$(0.159)^2$	$(0.166)^2$	$(0.182)^2$	$(0.147)^2$	$(0.188)^2$
F-stat	219.85	188.32	127.20	83.56	71.88	143.04
Panel B: Overidentified model						
Infertility 1	-0.409	-0.364	-0.325	-0.499	-0.161	-0.606
	(0.077)	(0.076)	(0.074)	(0.206)	(0.082)	(0.126)
Infertility 2	-0.919	-0.850	-0.719	-1.399	-0.638	-1.159
	(0.057)	(0.057)	(0.061)	(0.126)	(0.059)	(0.098)
Shea's R^2	$(0.181)^2$	$(0.168)^2$	$(0.175)^2$	$(0.196)^2$	$(0.167)^2$	$(0.194)^2$
F-stat	148.94	127.10	81.91	71.30	61.69	88.97
N	6,035	6,025	3,513	2,522	2,560	3,475

Note: Robust standard errors in parentheses. All regressions include controls for child's age, sex, mother's education, marital status, urban-rural location and survey fixed effects. Only in *Health*, BMI is included as a set of binary variables representing underweight, overweight and obese. The variable *Infertility* is the union of infertility 1 and 2. See text for definitions.

Table 5: First Stage results for children under 6.

Dependent variable: Number of children at home (K_j)				
	Anthropometric sample		Breastfeeding sample	
	No Health	With Health	No Health	With Health
Panel A: Just-identified model				
Infertility	-0.233	-0.231	-0.222	-0.23
	(0.034)	(0.034)	(0.032)	(0.033)
Shea's R^2	0.078	0.078	0.072	0.075
F-stat	48.01	46.64	46.87	48.98
Panel B: Overidentified model				
Infertility 1	-0.111	-0.11	-0.099	-0.1
	(0.063)	(0.064)	(0.059)	(0.061)
Infertility 2	-0.291	-0.289	-0.288	-0.295
	(0.036)	0.036	(0.035	(0.035)
Shea's R^2	0.085	0.085	0.079	0.083
F-stat	36.53	35.37	37.27	38.98
N	6,243	6,153	7,154	6,737

Note: Robust standard errors in parentheses. All regressions include controls for child's age, sex, mother's education, marital status, urban-rural location and survey fixed effects. Only in *Health*, BMI is included as a set of binary variables representing underweight, overweight and obese. The variable *Infertility* is the union of infertility 1 and 2. See text for definitions.

Table 6: Impact of Sibship on School Attendance (6-14 sample)

Dependent variable: Current School Attendance								
	Mean	OLS	IV	GMM	Mean	OLS	IV	GMM
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Panel A: Full Sample								
		Main Specification				Adding Mother's Health Status		
Number of siblings	0.883	-0.009	-0.010	-0.014	0.883	-0.011	-0.014	-0.018
		(0.004)	(0.015)	(0.014)		(0.004)	(0.017)	(0.015)
J-stat				[0.640]				[0.674]
Panel B: By Location								
		Urban				Rural		
Number of siblings	0.931	-0.004	-0.012	-0.021	0.816	-0.016	-0.020	-0.018
		(0.004)	(0.020)	(0.016)		(0.006)	(0.027)	(0.024)
J-stat				[0.424]				[0.912]
Panel C: By Mother's Education								
		High Education				Low Education		
Number of siblings	0.950	-0.007	0.006	0.008	0.834	-0.011	-0.022	-0.029
		(0.005)	(0.026)	(0.023)		(0.005)	(0.020)	(0.018)
J-stat				[0.574]				[0.317]

Notes: Robust standard errors in parentheses. P-value for Hansen J-statistic in brackets. All regressions include controls for child's age, sex, mother's education, marital status, urban-rural location and survey fixed effects. Only in *Health*, BMI is included as a set of binary variables representing underweight, overweight and obese.

Table 7: Impact of Sibship on Schooling (6-14 sample)

	Dependent Variable: Standardized Years of Schooling							
	Mean	OLS	IV	GMM	Mean	OLS	IV	GMM
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Panel A: Full Sample								
		Main Specification				Adding Mother's Health Status		
Number of siblings	-0.001	-0.092	-0.076	-0.084	-0.001	-0.100	-0.100	-0.107
		(0.011)	(0.062)	(0.062)		(0.011)	(0.067)	(0.067)
J-stat				[0.827]				[0.782]
Panel B: By Location								
		Urban				Rural		
Number of siblings	0.223	-0.093	-0.049	-0.073	-0.313	-0.069	-0.089	-0.085
		(0.015)	(0.092)	(0.094)		(0.015)	(0.083)	(0.076)
J-stat				[0.864]				[0.795]
Panel C: By Mother's Education								
		High Education				Low Education		
Number of siblings	0.374	-0.065	-0.059	-0.073	-0.277	-0.088	-0.083	-0.091
		(0.020)	(0.139)	(0.132)		(0.013)	(0.066)	(0.065)
J-stat				[0.915]				[0.975]

Notes: Robust standard errors in parentheses. P-value for Hansen J-statistic in brackets. All regressions include controls for child's age, sex, mother's education, marital status, urban-rural location and survey fixed effects. Only in *Health*, BMI is included as a set of binary variables representing underweight, overweight and obese.

Table 8: Impact of Sibship on Health Outcomes (Under 6 Sample)

Dep. variable:	Weight for Age			Height for Age			Currently Breastfeeding		
	OLS	IV	GMM	OLS	IV	GMM	OLS	IV	GMM
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
				Panel A: Main Specification					
Number of	-3.17	-21.19	-20.57	-4.23	-2.38	-2.06	-0.088	-0.169	-0.199
Siblings	(0.76)	(11.21)	(10.32)	(0.68)	(9.72)	(8.65)	(0.005)	(0.108)	(0.099)
J-stat			[0.799]			[0.866]			[0.137]
Mean of Dependent Variable			{37.85}			{27.66}			{.322}
				Panel B: Including Health Indicators					
Number of	-2.32	-18.87	-18.50	-3.94	-1.64	-1.43	-0.089	-0.165	-0.194
Siblings	(0.76)	(10.95)	(10.15)	(0.68)	(9.73)	(8.66)	(0.005)	(0.108)	(0.098)
J-stat			[0.805]			[0.874]			[0.152]
Mean of Dependent Variable			{37.89}			{27.67}			{.321}

Notes: Robust standard errors in parentheses. P-value for Hansen J-statistic in brackets. All regressions include controls for child's age, sex, mother's education, marital status, urban-rural location and survey fixed effects. Only in *Health*, BMI is included as a set of binary variables representing underweight, overweight and obese.