Does More Mean Better? A Quasi-experimental Analysis of the Link between Family Size and Children's Quality

Javier E. Baez *
Graduate Student
Department of Economics
Maxwell School of Citizenship and Public Affairs
Syracuse University
Syracuse, New York 13244-1020
jebaez@maxwell.syr.edu

May, 2006
(First draft)

Exogenous variation in fertility from parental preferences for sibling sex-mix composition is used to identify the causal effect of family size on several outcomes associated to either the allocation of resources within the household or its influence on children's wellbeing. Reduced form results using data from Colombia suggest that family size has negative effects on average child's quality. Children from larger families have accumulated almost 1.3 years less of education, are less likely (24-35 percentage points) to enroll in school and almost twice more likely to be held back in school. Regarding household resources, oldest children from large families have a higher probability of sharing a room, 22% more likely to live in a house with no connection to sewer or septic well and 28% less likely to have access to clean water. Mothers in these households have nearly 50% less labor participation and their oldest kids are about twice as likely to engage in labor activities or domestic chores. Children from large families are also more likely to be physically or psychologically affected by actions of domestic violence. Other less sophisticated but informative calculations using data on anthropometrics, morbidity and immunization records also fit well with the main results of the quasi-experimental research design. The evidence presented here supports the idea of an existing taste for the number instead of quality of children, regardless of the interdependence in prices between them.

* Special thanks go to Gary Engelhardt, Jeffrey Kubik and Dan Black for helpful comments and advice. I also want to thank Indhira Santos and Raul Abreu-Lastra for useful comments and discussions. This paper is still work in progress. Please do not cite without the permission of the author.

Introduction

In April of 2005 the Major of Cucuta, a Colombian town with a population of more than 600,000 inhabitants, announced a controversial program to subsidize sterilization among poor men and women and advertised it as an effort to reduce poverty. Almost at the same time, the Ministry of Social Policy launched a small program in a couple of Colombian towns to provide free sterilization to women with four or more children, an strategy that state officials said was intended to "fight against poverty". Multiple birth-control initiatives such as these have been implemented in numerous places in the developing world over the last decades. Most, if not all, have been motivated by a commonly observed association: families with lower standards of living usually have higher fertility rates.

Understanding whether family size is a central determinant of investment on children and their future economic success is an issue of great social importance. In that regard, the co-movement between high birth rates and economic status that has encouraged policies like the ones mentioned above is far from being refutable, particularly in low-income environments. The problem is that the presence of a correlation between birth rates and economic status, although necessary, is not a sufficient condition for determining causality. With that in mind, increasing research efforts from several disciplines have been devoted to identify whether any meaningful causal effect exists. A broad review of the available evidence suggests that the results are mixed. The debate about the tradeoff between the quantity and quality of children is still unresolved, a point I will return to in the next section.

A primary obstacle in determining a causal connection linking larger families with worse economic outcomes lies in both the existence of unobserved factors (e.g. tastes for number and quality of children) that are correlated with each other and their possible joint determination. Therefore, standard Ordinary Least Squares (OLS) techniques pick these confounders and produce biased estimates of the population parameters. A setting in which the quantities of children vary exogenously is needed to disentangle the interaction between the number of children and quality per child. Previous papers in the existing literature have addressed this issue using an exogenous source of variation in fertility coming from either multiple births or sibling-sex composition [Rosenzweig and Wolpin (1980), Duflo (1998), Caceres (2004), Black, Devereaux and Salvanes (2005), Conley and Glauber (2005), Angrist, Lavy and Schlosser (2005), Rosenzweig and Zhang (2006)].

The paper presented here builds upon a source of variation in family size initially developed by Angrist and Evans (1998). Using data from the Colombian Demographic and Health Survey (DHS) of 2005, I exploit the exogenous change in fertility induced by sibling-sex preferences in families with at least two (three) children to shed some light on the effects of family size on first (first- and second-born) children's outcomes. The contribution of this study is primarily empirical. First, I extend earlier papers by looking at measures rarely examined in this context that can be related to either the allocation of resources within the household or that potentially influence children's wellbeing. Apart from schooling performance, these variables include health care utilization, the probability of a child sharing a room with other kids, the access of children to clean water, the characteristics of the household they are living in, mothers' labor supply, children's use of time and domestic violence. Second, this work offers an additional opportunity for external validation in the sense that most of the conclusions so far –particularly those using same-sex instruments– have not been drawn from empirical evidence in developing countries, where the desired number of births and behavioral responses to family size may differ systematically.

Contrary to what has been identified in other works mostly examining developed countries, the findings of this paper indicate that OLS parameters underestimate the effect of family size. Several Two-stage Least Squares (2SLS) estimates appear to be considerably different from those obtained by OLS. In general, most of the empirical 2SLS models employed in this paper suggest that family size has negative effects on several variables associated with children's quality. More specifically, I find that children from larger families have accumulated almost 1.3 years less of education, are less likely (between 24 and 35 percentage points) to be enrolled in school and almost twice more likely to be held back, compared to their age reference cohorts. Regarding household resources, oldest children from larger families are more likely to be sharing a room with their siblings, 22%% more likely to be living in a house with no connection to sewer or with earth floor and 28% less likely to have access to clean water. Mothers in households with more children have almost 50% less labor participation and their oldest kids about twice as likely to engage in labor markets or domestic chores and spend around four more hours per week on these activities. On average, first- and second-born from large families are more likely to be physically or psychologically affected by actions of domestic violence within the household. Other less sophisticated but illustrative calculations using data on anthropometrics, morbidity and immunization records appear to match these findings as well. The evidence presented here seems to go along with the hypothesis that children are "cheap" and economically useful for poor parents in developing societies. The evidence also suggests the existence of a "taste" for the number rather than the quality of children, regardless of the interdependence in prices between each other.

The remainder of this paper is structured as follows. Section II provides a short background of the literature and briefly depicts the identification issues in the traditional methods of estimation. Section III illustrates the empirical strategy followed in this paper in order to circumvent these identification problems and includes a description of the data and some relevant descriptive statistics. Section IV presents the empirical findings of the 2SLS procedures undertaken, including a discussion of the first stage regressions and other complementary calculations Section VI concludes.

I. Background

Among the several interactions taking place within the household, not surprisingly, a great deal of attention has been devoted to the relationship between the quantity of children and their quality.[1] This interest is mostly driven by the negative association between income and fertility that is frequently observed within and across countries. The theoretical foundations of these patterns in fertility behavior are integrated in the "quantity-quality" model [Becker (1960), Becker and Lewis (1973), Becker and Tomes (1979)] which provides consistent predictions to this regularity. Modeling the optimal choices of the

[1] Quality is understood here as the level of direct and indirect investments on children that provide utility to the household and that affect positively their development, future earnings and living potential. This quality is also affected through other channels partially out of the control of the parents such as inherited ability, peer effects, public investments and other sort of random events.

quantity and quality of children in a similar fashion of other commodities in the household, this framework stresses a particular link between the two: the shadow price of the number children is positively related to the level of quality and vice versa. In other words, assuming no parental discrimination between their kids, an increase in the number of children increases the cost of raising a quality child (because it now applies to more of them) and, correspondingly, an increase in quality per child is more expensive the greater the number of children. This argument leads to an implicit trade-off between quality and quantity and reconciles the evidence of parents demanding children of higher quality as they get richer, without increasing their desired fertility.

Testing the “quantity-quality” trade-off is not only a matter of theoretical and empirical relevance but a subject with deep policy implications as well. Programs aimed to discourage large families in order to reduce poverty and promote economic development can be justified on the basis of analyses confirming the negative influence of family size on child quality. In fact, some existing evidence has systematically obtained results supporting this connection. For instance, Leibowitz (1974), Blake (1981), Hanushek (1992) and Hill and O’Neill (1994) find that children from larger families have lower schooling attainment in various dimensions.

From an empirical standpoint, there are concerns whether the causal effect of fertility of interest has been identified in works that adopt a standard OLS analysis without dealing with other confounders. This apprehension stems directly from the endogenous nature of the trade-off between number and quality of children.[2] In order to illustrate this formally,

[2] See Browning (1992) for a longer and more detailed discussion of the traditional limitations in modeling the effects of children on household behavior.

we can assume that the measure of investment or the outcome of investment of children i, C_i, at point in time, t, can be described by:

$$C_i = \beta X_i + \delta N_i + \varepsilon_i \qquad (1)$$

where N measures the number of siblings at home, X represents demographic characteristics and other controls and the error term can be decomposed in two terms, $\varepsilon_i = \nu_i + \eta_i$, with ν_i being a standard stochastic disturbance and η_i being a household fixed unobserved effect. Therefore, using the data in terms of deviations of the mean and calling S the sample covariance between any two variables, the relationship between C and N from the model in (1) can be written as:

$$S_{C,N} = \beta * S_{X,N} + \delta * S_{N,N} + S_{\varepsilon,N} \qquad (2)$$

The parameter δ that captures the relationship of interest can be recovered only if some assumptions are satisfied. Initially, we can assume that the number of children is orthogonal to other observable characteristics (i.e. $S_{X,N} = 0$) and focus on more challenging issues. For example, several findings linking children from larger families with worse outcomes have implicitly assumed that the third term in the right hand side of (2) is negligible. That is, the decisions about the number of children are completely independent of any fixed characteristic of the household and other unobservables do not condition fertility whatsoever, thus the parameter δ is fully identified.

The notion of interdependence between quantity and quality of children sketched in the theory suggests the presence of an omitted variable bias in that type of approach. We can think of a group of families with relatively higher (and unobservable) preferences for more children facing a higher marginal cost of quality per child ($S_{\eta,N} > 0$, $S_{\eta,C} < 0$), whereas families with preferences for fewer children would deal with a lower shadow price for quality ($S_{\eta,N} < 0, S_{\eta,C} > 0$). Indeed, there is an apparent preference of some households –particularly those with low income and low education in the developing world– for a larger number of children with rather low investment per child. For these families, raising children is cheap in the sense that they become wage earners at relatively young ages and, due to low female wages and employment opportunities, the cost of maternal time devoted to childcare is not significant. Not accounting for this unobserved characteristic would yield upward biased estimates of the population parameter of interest.

In contrast, the potential sources of individual heterogeneity can also go in the opposite direction to the one just pointed out above, namely towards zero. An example of this being the common belief that children raised in only-child households are growing up without the benefit of child company and, thus are disadvantaged (e.g. selfish, lonely, spoiled, maladjusted, etc.). That is, families can have preferences for larger families just because they might perceive the numerous interactions taking place among siblings as children's quality enhancing and highly beneficial for their future development. Likewise, non-random errors in the measurement of children's outcomes may produce attenuation bias in estimates of the impact of family size on children's quality (e.g. one-child families over-reporting the quality of their kids). Even though these two problems are of different nature, they both would result in underestimating the magnitudes of the effects of fertility on children's well-being. In short, a final determination on which one of these or other sources of bias dominates is, therefore, a subject of empirical analysis.

Disentangling the effects of the number of children is not a trivial task because a researcher needs an exogenous and measurable change in fertility that can be used as a source of identification to produce credible inference of its effects. In principle there are only two options available to get around these obstacles in the absence of longitudinal data. First, one could assign new births randomly among different families and examine whether investment on children vary with the number of them. Although ideal, this research design is unattainable for obvious reasons. Secondly, and more feasibly, one can try to find an exogenous change in fertility that simulates that experiment by inducing a pseudo-randomization of families between treatment and comparison groups.

Studies that have reproduced that simulation have exploited mostly two plausible exogenous factors of variation in family size: multiple births and mixed-sex sibling preferences. For instance, Rosenzweig and Wolpin (1980a) use the occurrence of twins to test the trade-off between quantity and quality of children in Indian households. They found that an increase in fertility had a negative impact on children's educational attainment and expenditures on consumer durables. Similarly, Rosenzweig and Wolpin (1980b) exploited U.S. data on multiple births to find a negative effect on women's labor participation. Conley and Glauber (2005) use the sibling-sex composition type of variation to find a negative impact of family size on children's educational outcomes using 1990 U.S. census data.

While this evidence seems to confirm the inverse link between the number and quality of children, other empirical studies do not support this view. For example, Caceres (2004) uses U.S. Census data of multiple births to find that an increase in the number of children reduces their chance to attend private school and their mother's labor participation, while increasing their probability of the children sharing a room and parental divorced. Despite of the influence of family size on these proxies of allocation of resources within the household, he finds no effect on other measures closer to children's wellbeing such as school grade or the likelihood of dropping out. Qian (2004) uses the relaxation in China's "one child policy" and the event of multiple births to find that for single child families, an exogenous increase in family size has a positive effect on first child's school attendance, an effect that is reversed for two-child families. Using Norwegian data and twin births as instruments as well, Black, Deveraux and Salvanes (2005) find no effects of sibship size on measures of educational attainment. Finally, Angrist, Lavy and Schlosser (2005) exploit variation in fertility from both multiple births and sibling-sex preferences in Israeli Census data and find no evidence of a quantity-quality trade-off.

III. Empirical Methodology

A. Identifying Strategy

Following Angrist and Evans (1998), I exploit plausible exogenous changes in fertility due to parental preferences for mixed-sibling sex composition in at-least-two-child and at-least-three-child families. Since the decision of having more kids is in part driven by the desire of having at least one of each sex and gender is determined exogenously, a dummy variable to identify women having same-sex births produces an instrumental variable (IV) for changes in fertility among these families. In other words, this instrument induces exogenous differences in family size at the time that appears to be orthogonal to resource allocation decisions and its determinants.

This empirical strategy is suitable to simulate the desired experiment of changing family size randomly and construct proper counterfactuals. For example, assume the researcher is interested here in two potential outcomes C_{oi} and C_{1i}, which represent the investment or product of investment of children *i* with, let's say, *n* and *n+1* children, respectively. Using the constant-effect models, these two potential outcomes can be written as:

$$Y_i = \alpha + \delta D_i + X_i \beta + \upsilon_i \tag{3}$$

where D_i is used to denote treatment status. The difference in children's outcomes between children that were raised in *n+1*-children families ($D_i = 1$) and those that were raised in *n*-children families ($D_i = 0$) is given by:

$$\begin{aligned} E[C_{i1} \mid D_i = 1] - E[C_{i0} \mid D_i = 0] &= E[C_{i1} - C_{i0} \mid D_i = 1] \\ &+ \{E[C_{i0} \mid D_i = 1] - E[C_{i0} \mid D_i = 0]\} \end{aligned} \tag{4}$$

$$= \delta + \{E[\upsilon_i \mid D_i = 1] - E[\upsilon_i \mid D_i = 0]\} \tag{5}$$

The first term in the right hand side of (4) is the average causal effect of increasing the number of children from *n* to *n+1* for those who had *n* children. Nevertheless, it is not possible to observe the same family with *n* and *n+1* children at the same time. The counterfactual average $E[C_{i0} \mid D_i = 1]$, therefore, cannot be observed. The second term in the right hand side of (4) represents the omitted variable bias if the outcome of families with *n* children is not a good counterfactual of those with *n+1* children had they had *n* children, which would be the case of the sort of unobservable preferences discussed in the previous section.

In principle, an instrument Z_i constructed out of the sibling-sex composition (SSC) of the first two and first three births removes that potential bias. This sex composition is used to construct a dummy IV taking the value of one for those families with the first two or three births having the same sex, and zero otherwise. Since the SSC is randomly assigned and appears to have significant effects on family size, a quasi-experimental setting involving this IV can be used to approximate a randomized trial. The effect of fertility on children's quality can be generalized by the following model:

$$C_i = X_i\beta + \delta D_i + \upsilon_i \tag{6}$$

where D_i is a dummy variable for women having more than two children (in families with at least two births) or more than three children (in families with at least three births) and the other variables as defined before. If only the variation of the instrument that is associated with D_i is used to identify the parameter of interest, it follows that:

$$\delta^{IV} = \frac{S_{C,Z}}{S_{D,Z}} \tag{7}$$

And given that Z_i is a binary variable, the IV sample estimate of δ can be expressed as follows:

$$\delta^{IV} = \frac{\{E[C_i \mid Z_i = 1] - E[C_i \mid Z_i = 0]\}}{\{E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]\}} \tag{8}$$

The expression in (8), a Wald estimate, depicts the reduced-form relationships between Y_i and Z_i (numerator) and Y_i and D_i (denominator). This parameter is the ratio of the difference in the measure of child quality and the probability of having more than two (three) children for those women with the first two (three) births of the same sex and those with sibling-sex mixed. Exactly the same way of thinking applies for the case in which the number of children, instead of D_i, is instrumented by Z_i.

In line with the identifying assumption, it follows that $E\left[C_i \mid Z_i\right]$ varies with the instrument only through its effect on D_i and thus a causal relationship between family size and children's quality can be recovered. Nonetheless, it is worth mentioning that this source of variation is only identifying the effect of having more births on those women in at-least-two-child and at-least three-child families whose treatment status was modified by Z_i, a parameter usually called local average treatment effect (LATE, Imbens and Angrist, 1994). Accordingly, the IV strategy used here is estimating the effect of fertility on families who had more children because of the SSC of their births but would not otherwise have increased their family size.

B. Data, Samples and Descriptive Statistics

The data employed here come from the Colombian Demographic and Health Survey (DHS) of 2005, a nationally representative cross-sectional survey that collected information from nearly 37,000 households on several indicators of fertility, education, health, nutrition, standards of living and domestic violence, among others. One of advantage of this survey is that it provides very precise pregnancy rosters and birth registration records that allowed me to match children with their biological mothers whose first birth was between the age of 15 and 49.

The samples used for the empirical exercises were restricted in several ways. I looked only at households with women being either heads or spouses of male householders, who have children below the age of 18 that were recorded as their births and that are living in the household. Those families with siblings living somewhere else were not included to avoid complications (e.g. inter-household transfers) and multiple families living in the same dwelling were treated independently. The SSC was constructed out of births that remained alive after one year old because of the non-trivial number of newborns dying under that age.

The unit of analysis are first births in the sub-sample of families with at least two children and first and second births in the sub-sample of families with at least three children. Each of these two sub-samples consists of roughly 7,600 observations. The reason of retaining only these sub-samples is because they contain those children that are exposed to the quasi-experiment derived from the S-S-C, unconditional on the treatment status.

For each family I constructed several variables that, although difficult to relate strictly to either investment or the product of it, may be associated somehow with children's quality. The first set of outcomes describe educational achievement and comprises the number of years of school completed, school attendance and children's progress in school (relative position of the child with respect to his/her reference cohort of the same age). The second group of variables describes the resources of the household and includes measures such as the probability of sharing a room with other siblings, the access to clean water and sewer, and some other physical characteristics of the dwelling (e.g. likelihood of living in a household with walls of mud and earth floor). The third set of variables is intended to examine the impact of family size on the attachment to labor market and includes mother's labor participation, children's use of time (labor activities both in the market and the household) and children's time spent working. The fourth group assesses health care utilization by estimating the likelihood that a sick child is taken for consultation. I also look at the probability that a child is involved or affected by actions of domestic violence at the household level as a proxy of allocation or dilution of resources. Finally, a set of anthropometrics, morbidity and vaccination variables for children under four are used for informational purposes to shed some light on the quantity-quality relationship.

Table 1 reports descriptive statistics of some relevant characteristics and variables used in the construction of the IV's for treatment and comparison groups in the two sub-samples of analysis. Overall, these summary statistics suggest that both experimental and non-experimental groups are comparable in terms of the variables reported and not any of the differences between them is statistically significant. Nearly 25% of the children with one or more siblings live in rural areas and have similar levels of wealth.[3] Most of their mothers are married (87%), are on average 32 years old, have spouses that are almost 4 years older and have approximately 7.8 years of schooling. Approximately 17% of the households are single headed. Characteristics associated with birth spacing are also very similar. Mothers had the first and second births at the ages of 21.5 and 25, respectively.

The analogous means for families with at least three children (also reported in Table 1) are slightly different to the ones just described but very comparable between treatment

[3] Unfortunately the DHS dataset does not collect detailed information on incomes. However, a wealth index constructed out of some assets of the family (e.g. radio, television, refrigerator, other electronic devices, characteristics of the dwelling, motorcycle, car, etc.) and other proxies such as the years of schooling of the parents and socioeconomic strata were used in this analysis. The results are not very sensitive to any of these alternative measures.

and controls as well. The wealth index is somewhat lower for this sub-sample; approximately 30% of these children live in rural areas and have parents over one year older. In addition, these mothers have almost one year less of education and had the births when they were roughly one year younger. On the whole, around 51% of the births are male regardless of the sub-sample, which implies that half of the first children in two-or-more-child families and 25.9% of the first two children in three-or-more-child families were born into a same-sex sibling pair and threesome, respectively.

IV. Results

A. First Stage Estimates

The exogenous variation derived from the SSC was exploited as an instrument for the total number of children in the household, and as well as an instrument for a dummy variable taking the value one in households having more than two (three) children in families with at least two (three) children, and zero otherwise. The first stage estimates for the first-born sample are based in the following model, which can be straightforwardly extended to the case of the first- and second-born sample:

$$N_i = X_i \beta + \lambda S_i + \eta_i \qquad (9)$$

where N can be either the total number of children in the household or the binary variable just described above, S is the dummy variable associated with the S-S-C (=1 if the sex of the first two children is boy-boy or girl-girl) and X stands for a set of other covariates that include the age and gender of the child, mother's age at first birth, birth spacing, mother's years of schooling, a wealth index, and a set of dummies to identify rural families, control for state effects and other characteristics of the households.

Table 2 summarizes unconditional and conditional effects from sex-compositions instruments using the model in equation (9). On average, families with either two boys or two girls (treatments) have 2.75 children, while their counterparts (controls) have 2.68, namely a difference of 0.073. Correspondingly, 48.1% of the households with same-sex siblings have a third child, whereas only 43.3% of those with sibling-sex mix have another birth. These unconditional estimates are not only statistically significant but robust to specifications that include several controls.

These results are roughly replicated by the models run in the sub-sample of families with at least three children and presented in the second panel of Table 2. Although somewhat weaker for the unconditional specifications, the F-value of the conditional estimates –the ones used in the second stage– suggests that the joint power of these first stages is comparable to that presented in the first panel. In brief, these robust and statistically significant differences in fertility resulting from parental preferences for mixed sex represent the input of variation in family size used in the second stages, whose results are presented in the next section.

B. Family Size and Children's Quality

OLS and 2SLS estimates of the model in equation (3) that describe the effects of family size on children's quality are reported in Table 3 and Table 4. For most of the outcomes, 2SLS estimators were obtained by running both IV linear probability models and IV probit Newey's efficient two-step minimum chi-squared procedures.[4] Each of the tables summarizes the results when the total number of children and dummies for the

[4] Traditional IV methods were employed for non-binary dependent variable models.

probability of having more than two (three) children in families with two (three) or more children are used as measures of fertility. The first panel presents the results of the oldest children in families with at least two children while the second reports the estimates of the first and second-born in families with at least three children. The outcome variables of interest approximate the influence of family size on characteristics of the children and their environment that can be seen as proxies of investments on them or the product of these investments.

Before discussing the results in detail, two main conclusions derived from the empirical exercises are worth mentioning. On one hand, OLS estimates suggest that family size has adverse effects on several variables of children's quality. On the other hand, 2SLS estimates that account for potential endogeneity bias appear to be considerably higher than OLS estimates, difference that in some cases is statistically significant. Contrary to what has been identified in other works, mostly for developed countries, I consistently find that the OLS parameters underestimate the effect of family size on measures related to the allocation of resources within the household. This conclusion could be undermined by some issues about the precision of some 2SLS estimates reported here to be informative enough. Although this is a legitimate concern, it can be argued that all the empirical models run for the different outcomes found the 2SLS estimates to be systematically higher (in absolute value) than their OLS counterparts, regardless of the estimator employed. In a few cases their 95% confidence intervals suggest they are statistically distinct as well. In my opinion, this can be an indication that OLS methods provide undervalued inference of the true quantity-quality trade-off and that still remarkable precise 2SLS estimates in the more conservative scenario would be at least slightly higher than OLS estimates.

Overall, the findings are indicative of children growing in large families being disadvantaged. More specifically and regarding school attainment, I find that children with more siblings have accumulated almost 1.3 years less of education and those in school-age are between 24 and 35 percentage points less likely to be attending school and almost twice more likely to be held back when compared to their age reference cohorts (e.g. school repetition, dropouts). In terms of household resources, oldest children from large families are more likely to be sharing a room with their siblings, 22% more likely to be living in a house with no connection to sewer or septic well or with earth floor and 28% less likely to have access to clean water, the latter being apparently a strong predictor of life expectancy. In addition, these children seem to be over twice as likely to be living in a dwelling with earth floor and walls of mud, although some of these results are not particularly strong in statistical sense.

The results for labor market outcomes indicate that mothers in households with more children have almost 50% less labor participation and their oldest kids are about twice as likely to engage in labor markets or domestic chores and spend around four more hours per week on these activities. First-borns from large families are also more likely to be physically or psychologically affected by actions of domestic violence within the household. Conditional on being sick, they also have a lower probability of being taken to the doctor, although this effect is highly imprecise. These effects appear to be uniformly larger for latter-born children, namely those studied in the at-least-three-child families, although with a lower level of precision in some cases.

The last set of results is for anthropometric, morbidity and immunization outcomes, which are close proxies to evaluate intra-household investments on children. However, since these measures were collected only for children under four, only a very limited number of experimental first-born and first- and second-born observations lie in this age range.[5] Consequently, the size of these sub-samples does not provide adequate statistical power to embark on IV methods. In order to get around this restriction, I run OLS and Probit models which –under the assumption of a prevailing negative endogenous bias in this selected sample– would yield lower bound estimates of the influence of family size on children's quality and may be informational useful.

The results of these empirical exercises are reported in Table 5. Initially, three of the most common anthropometric measures are used to assess children's nutrition: Body Mass Index (BMI), weight-for-age and height-for-age. Regardless of the index, family size seems to be associated with a higher risk of malnourished children, in particular for families with six or more children (see Figure 1). In view of the fact that anthropometric indicators of nutrition seem to reflect children's health status more accurately for the left tail of the distribution, I restrict the calculations to the 50^{th} and 25^{th} percentiles. The results of these sub-samples replicate the direction of the effects obtained for the whole distribution in the expected manner. In terms of morbidity, the results suggest that a shift in the number of children is coupled with a higher probability of children getting diarrhea

[5] In fact, only 683 (458) first-born (first- and second-born) observations live in families with at least two (three) children and are four years old or younger.

(between 1% and 2%) in the two weeks preceding the interview. Nevertheless, additional siblings in the household do not have a statistically significant influence on the likelihood of a child getting fever. Finally, I find that an extra child appears to be connected with a drop (between 1.4% and 4.6%) in the probability of children being immunized against tuberculosis, diphtheria and tetanus, poliomyelitis and measles.

V. Conclusions

This paper is an attempt to examine the influence of family size on average child quality using exogenous variation in fertility induced by parental preferences for sibling sex-mix composition. Empirical verification of this relationship is a challenging task because large samples with detailed socio-demographic and pregnancy retrospective data are not often found in household surveys in developing countries. I use a rich dataset to find evidence pointing to a detrimental effect of family size on children's quality, namely that children on average may be better off if their families had not been larger. These empirical results have been drawn from a wide range of outcomes that have been rarely explored and provide information on education, household resources, labor participation, use of time, health care utilization and domestic violence. Other less ambitious but illustrative exercises using data on nutritional status, morbidity and vaccination records also seem to fit well with the main results of the quasi-experimental research design.

The findings of this work clearly depart from evidence obtained in studies applied to rich nations in which the number of births has been found to have little effect on children's outcomes. At this point it is worth mentioning some conditions in which the results of this paper can be reconciled within the context of a developing country. My argument is that poor families in less developed societies have relatively less room to reallocate different types of quality inputs in response to exogenous changes in fertility, and thus, children's well-being is at risk of being cut back. For instance, in the face of larger families, parents may try to adjust their labor supply and consume less leisure. But this sort of adjustment is not very feasible for parents with low levels of education that in several cases have either access to low-wage informal jobs or deal with very high levels of unemployment and underemployment. Actually, in some previous calculations using Colombian data I find that husband's labor supply is not very responsive to childbearing whereas mother's labor supply is negatively affected. On the other hand, families that are already under low standards of living are unable to hold quality constant and have less scope to substitute away from parent's consumption to children's consumption.

The evidence presented here also give the impression to go along with the idea that children are "cheap" and economically useful for poor families in developing communities and supports the idea of an existing taste for the number instead of quality, regardless of the interdependence in prices between each other. Indeed, children from these backgrounds happen to be cash earners and a source of extra labor at home at very young ages. Many of them are not enrolled in school, are not covered by health insurance and are inadequately fed. Besides, the cost of home care is relatively trivial in the sense that female wages are low and employment opportunities quite scarce. That is, children can be very inexpensive to raise in these settings.

Although the findings of this study apply exclusively to Colombia, to some extent they might reflect the effects of fertility on economic circumstances of people in countries confronted to similar constraints. Policies aimed to discourage large families, forbid child labor, increase public access to child care, school and health services can be helpful in switching the emphasis from number to quality in low-income families. I plan to tackle some shortcomings of this study and investigate some of these hypotheses and policy options in more depth in future work with the aid of alternative datasets and evidence from other children’s outcomes.

References

Angrist, J., Lavy, Victor and Analia Schlosser, "New Evidence on The Causal Link between the Quanity and Quality of Children", *National Bureau of Economic Research*, Working paper 11835, 2005.

Angrist, J. and W. Evans, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *The American Economic Review* 88(3), 1998, pp. 450-477.

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

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

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

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

Bronars, S. and J. Grogger, "The Economic Consequences of Unwed Motherhood: Using Twins as a Natural Experiment", *American Economic Revi*ew, 84, 1994.

Caceres, J., "Impact of Family Size On Investment in Child Quality: Multiple Births as Natural Experiment," mimeo, Department of Economics, University of Maryland, 2004.

Conley, D. and R. Glauber, "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility," *National Bureau of Economic Research*, Working paper 11302, 2005.

Conley, D. and R. Glauber, "Sibling Similarity and Difference in Socioeconomic Status: Life Course and Family Resource Effects," *National Bureau of Economic Research*, Working paper 11320, 2005.

Duflo, E., "Evaluating the Effect of Birth-Spacing on Child Mortality," mimeo, Department of Economics, Massachusetts Institute of Technology, 1998.

Gelbach, J., "Public Schooling for Young Children and Maternal Labor Supply," *American EconomicReview* 92(1), 2002, 307-322.

Gronau, R., "The Effect of Children on the Housewife's Value of Time", *Journal of Political Economy*, 81, 1973.

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

Imbens, G.W. and J. Angrist, “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62(2), 1994, 467-75.

Lee, J., “Sibling Size and Investment in Children’s Education: An Asian Instrument,” *IZA*, Discussion Paper 1323, 2004.

Moav, O., “Cheap Children and the Persistence of Poverty,” *The Economic Journal*, 115 (January), 2005, pp. 88-110.

Qian, N., “Quantity-Quality and the One Child Policy: The Positive Effect of Family Size on School, Enrollment in China,” Thesis chapter, Department of Economics, Massachusetts Institute of Technology, 2004.

Rosenzweig, M. and K. Wolpin, “Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment,” *Econometrica*, 48(1), 1980, pp. 227-240.

Rosenzweig, M. and K. Wolpin, "Natural 'Natural Experiments' in Economics", *Journal of Economic Literature*, 38(4), 2000, pp. 827-74.

Rosenzweig, M. and Junsen Zhang, "Do Population Policies Induce More Human Capital Investments? Twins, Birthweight and China’s “One Child” Policy", IZA Discussion Paper No. 2082, April,.2006.

Willis, R., “What Have We Learned from the Economics of the Family?”, *American Economic Review*, 77, 1987.

Table 1. Means Differences of Some Relevant Characteristics by Treatment Status

Variable	Families with at least two births: Same Sex (=1)	Same Sex (=0)	Diff	Families with at least three births: Same Sex (=1)	Same Sex (=0)	Diff
Wealth index	2.770	2.800	-0.030	2.537	2.495	0.042
	[0.021]	[0.021]	[0.030]	[0.029]	[0.017]	[0.033]
Household living in a rural area	0.252	0.260	0.256	0.290	0.308	-0.018
	[0.007]	[0.007]	[0.004]	[0.010]	[0.006]	[0.012]
Percentage of married mothers	0.870	0.874	-0.004	0.865	0.873	-0.008
	[0.005]	[0.005]	[0.007]	[0.007]	[0.004]	[0.008]
Age of the mother	32.4	32.4	0.000	33.7	33.7	0.000
	[0.100]	[0.100]	[0.141]	[0.127]	[0.076]	[0.065]
Age of the household head	36.5	36.5	0.026	37.8	37.7	0.100
	[0.121]	[0.123]	[0.173]	[0.156]	[0.097]	[0.184]
Mother's school attainment	7.72	7.86	-0.140	6.86	6.70	0.158
	[0.067]	[0.065]	[0.094]	[0.090]	[0.052]	[0.104]
Age of mother at first birth	21.5	21.5	0.020	20.6	20.6	-0.040
	[0.069]	[0.069]	[0.098]	[0.084]	[0.050]	[0.098]
Single headed households	0.172	0.172	0.000	0.177	0.177	0.000
	[0.006]	[0.006]	[0.008]	[0.008]	[0.005]	[0.009]
Age of mother at second birth	25.1	25.2	-0.040	23.5	23.7	-0.111
	[0.081]	[0.081]	[0.115]	[0.095]	[0.057]	[0.111]
Age of mother at third birth	----	----	----	27.2	27.2	-0.005
				[0.112]	[0.066]	[0.057]
Age of first born	10.9	10.9	-0.030	13.1	13.0	0.048
	[0.070]	[0.069]	[0.028]	[0.090]	[0.055]	[0.106]
Age of second born	7.3	7.2	0.030	10.1	10.0	0.120
	[0.072]	[0.073]	[0.051]	[0.088]	[0.054]	[0.104]
Age of third born	----	----	----	6.4	6.4	0.011
				[0.094]	[0.056]	[0.110]
Birth spacing between first and second births	3.61	3.68	-0.069	2.97	3.04	-0.070
	[0.040]	[0.039]	[0.056]	[0.044]	[0.026]	[0.051]
Percentage with girl at first birth	0.491	0.491	0.000	0.489	0.487	0.002
	[0.008]	[0.008]	[0.011]	[0.012]	[0.006]	[0.013]
Percentage with girl at second birth	0.494	0.510	-0.016	0.489	0.495	-0.006
	[0.008]	[0.008]	[0.011]	[0.012]	[0.006]	[0.013]
Percentage with girl at third birth	----	----	----	0.478	0.497	-0.019
				[0.011]	[0.006]	[0.013]
Percentage with boys at first two births	0.257	----	----	----	----	----
	[0.438]					
Percentage with girls at first two births	0.244	----	----	----	----	----
	[0.427]					
Percentage with boys at first three births	----	----	----	0.139	----	----
				[0.346]		
Percentage with girls at first three births	----	----	----	0.123	----	----
				[0.328]		
Percentage with first two children having same sex	0.500	----	----	----	----	----
	[0.500]					
Percentage with first three children having same sex	----	----	----	0.259	----	----
				[0.440]		
Number of observations	3,829	3,842	7,671	1,992	5,585	7,577

Notes: Standard errors (clustered by household for the three columns in the second panel) presented in square brackets. The first (second) panel includes estimates from households with at least two (three) children that are 18 years old or younger, are matched to their biological mothers and are living in the same dwelling. The wealth index is a measure included in the survey related to asset holdings that ranges between zero and five with three decimal positions. See text for definitions of treatment and control families.

Table 2. Sibling-Sex-Composition First Stages by Type of Family

Outcomes	Families with at least two children (First-born)				Families with at least three children (First- and Second-born)			
	Coefficient (S-S-C=1)		F	N	Coefficient (S-S-C=1)		F	N
Total number of children								
Unconditional	0.073	*	10.01	7,671	0.084	**	5.46	7,577
	[0.023]		[0.001]		[0.036]		[0.019]	
Conditional	0.064	*	64.01	7,346	0.063	*	44.92	7,240
	[0.017]		[0.000]		[0.019]		[0.000]	
More than two kids								
Unconditional	0.045	*	16.31	7,671	----		----	----
	[0.011]		[0.000]					
Conditional	0.047	*	48.46	7,432	----		----	----
	[0.010]		[0.000]					
More than three kids								
Unconditional	----		----	----	0.046	**	6.95	7,577
					[0.017]		[0.008]	
Conditional	----		----	----	0.043	*	39.85	7,240
					[0.011]		[0.000]	

Notes: Robust standard errors (clustered by household for the three columns in the second panel) presented in square brackets. The symbols * and (**) stand for significance at the 1%, (5%) levels, respectively. The first (second) panel includes estimates from households with at least two (three) children that are 18 years old or younger, are matched to their biological mothers and are living in the same dwelling. The conditional regressions include other covariates such as the age and gender of the child, mother's marital status and age at first birth, spacing between first and second birth, mother's years of schooling, a wealth index (as described in Table 1), and a set of dummies to identify single headed families, rural households, state and municipality effects and other characteristics of the households. See text for definitions of treatment and control families.

Table 3. OLS and 2SLS Reduced Form Estimates of the Effect of Family Size on Measures of Children's Quality (Endogenous variable: Total number of children)

	Families with at least two children								Families with at least three children							
Outcomes	Means	OLS		2SLS		IV Probit		N	Means	OLS		2SLS		IV Probit		N
School Performance																
Children's school attainment	4.478	-0.356	***	-1.429	*	__		7,346	4.443	-0.354		-1.176		__		7,261
	[3.829]	[0.036]		[0.792]					[3.367]	[0.034]		[0.787]				
Children attending school	0.793	-0.021	***	-0.244	**	-1.504	**	5,968	0.769	-0.039	**	-0.437	**	-2.248	**	6,189
	[0.431]	[0.005]		[0.116]		[0.791]			[0.421]	[0.006]		[0.220]		[0.970]		
Children held back in school	0.188	0.020	***	0.187		0.699		5,954	0.219	0.038	**	0.347	**	1.238	*	6,153
	[0.390]	[0.006]		[0.136]		[0.582]			[0.414]	[0.007]		[0.210]		[0.717]		
Household Resources																
Children sharing a room	0.797	0.107	***	0.469	***	2.738	**	7,346	0.920	0.029	***	0.250	*	3.875	***	7,261
	[0.401]	[0.004]		[0.183]		[0.800]			[0.269]	[0.003]		[0.150]		[1.518]		
Children living in household with access to sewer or septic well	0.844	-0.036	***	-0.196	*	-2.182	**	7,346	0.803	-0.039	***	-0.394	*	-3.021	**	7,261
	[0.362]	[0.005]		[0.109]		[1.089]			[0.397]	[0.007]		[0.235]		[1.248]		
Children having access to clean water	0.825	-0.006		-0.252	**	-1.331	**	7,346	0.816	-0.003	***	-0.453	*	-2.184	**	7,261
	[0.379]	[0.005]		[0.128]		[0.701]			[0.387]	[0.006]		[0.259]		[0.941]		
Children living in a household with walls of mud	0.216	0.056	***	0.157		1.147	*	6,655	0.270	0.044		0.314		1.262	*	7,261
	[0.411]	[0.005]		[0.130]		[0.700]			[0.444]	[0.008]		[0.223]		[0.767]		
Children living in a household with earth floor	0.099	0.040	***	0.165	**	2.017	**	7,346	0.135	0.041	***	0.387	*	2.126	**	7,261
	[0.299]	[0.004]		[0.083]		[1.043]			[0.342]	[0.006]		[0.212]		[0.812]		
Attachment to the Labor Market																
Mother participating in the labor market	0.529	-0.037	***	-0.301	*	-0.834	*	6,655	0.537	-0.067	***	-0.469	*	-0.335		7,261
	[0.499]	0.006		[0.167]		[0.479]			[0.498]	[0.009]		0.287		[0.234]		
Children working in the household or in the labor market	0.400	0.035	***	0.460	**	1.256	**	7,346	0.391	0.017	***	0.450	*	1.342	**	7,261
	[0.489]	[0.008]		[0.232]		[0.623]			[0.488]	[0.007]		[0.280]		[0.701]		
Hours per week worked by teenagers (conditioned on being working)	7.986	1.136	***	3.351		__		5,460	8.264	0.366	**	4.733	*	__		6,290
	[7.323]	[0.185]		[2.255]					[7.923]	[0.155]		[2.867]				
Health Care Utilization																
Children taken to the doctor (conditioned on being sick)	0.928	-0.014	**	-0.263		-0.803		3,412	0.915	-0.020		-0.259		-1.582		3,422
	[0.257]	[0.006]		[0.309]		[1.437]			[0.278]	[0.008]		[0.185]		[1.504]		
Domestic Violence																
Children affected by domestic violence	0.364	0.055	***	0.198	**	0.761	**	5,874	0.385	0.039	***	0.203		0.966	**	5,258
	[0.481]	[0.005]		[0.095]		[0.339]			[0.486]	[0.007]		[0.153]		[0.485]		

Notes: 2SLS estimators calculated by IV linear probability model and IV Probit Newey's efficient two-step minimum chi-squared procedure. Robust standard errors (clustered by household for OLS and 2SLS in the two panels) presented in square brackets. The symbols ***, (**) and [*] stand for significance at the 1, (5%) and [10%] levels, respectively. The first (second) panel includes estimates from households with at least two (three) children that are 18 years old or younger, are matched to their biological mothers and are living in the same dwelling. The variables *children's school attainment*, *children attending school* and *children held back in school* are restricted to children between 6 and 18 years old. The latter is a dummy variable that takes the value one for children that are in grades behind their reference cohorts due to school repetition, temporary and permanent dropout or late enrollment, and zero otherwise. The variable c*hildren having access to clean water* takes the value one for families receiving either piped water from utility company or bottled water, and zero otherwise. The variable *children working in the household or in the labor market* takes the value one for children engaged in domestic chores, family businesses, other household activities, self-employed or in remunerated and non-remunerated jobs, and zero otherwise. The variable *children affected by domestic violence* is equal to one if the children were pushed, deprived from food, hit with objects, assigned non-appropriate work, left out of the household for some time, thrown water, withdrawn economic support or witnessed actions of violence between their parents, and zero otherwise. The regressions include covariates such as the age, gender and schooling of the child, mother's marital status and age at first birth, spacing between first and second birth, mother's years of schooling, a wealth index (as described in Table 1), and a set of dummies to identify single headed families, rural households, state and municipality effects and other characteristics of the households. See text for definitions of treatment and control families.

Table 4. OLS and 2SLS Reduced Form Estimates of the Effect of Family Size on Measures of Children's Quality (Endogenous variable: Probability of having more than two (three) children in families with at least two (three) children)

	Families with at least two children					Families with at least three children				
Outcomes	**Means**	**OLS**	**2SLS**	**IV Probit**	**N**	**Means**	**OLS**	**2SLS**	**IV Probit**	**N**
School Performance										
Children's school attainment	4.478	-0.429 ***	-1.310 *	__	7,346	4.443	-0.433 ***	-1.740	__	7,261
	[3.829]	[0.060]	[0.800]			[3.367]	[0.050]	[1.166]		
Children attending school	0.793	-0.019 ***	-0.352 **	-2.207 **	5,968	0.769	-0.049 ***	-0.656 *	-3.945 **	6,189
	[0.431]	[0.007]	[0.162]	[1.101]		[0.421]	[0.010]	[0.350]	[1.634]	
Children held back in school	0.188	0.014	0.272	1.030	5,954	0.219	0.039 ***	0.519	1.989 *	6,153
	[0.390]	[0.010]	[0.195]	[0.837]		[0.414]	[0.012]	[0.351]	[1.123]	
Household Resources										
Children sharing a room	0.797	0.238 ***	0.613 ***	3.347 ***	7,346	0.920	0.066 ***	0.364 *	5.606 ***	7,261
	[0.401]	[0.009]	[0.202]	[0.922]		[0.269]	[0.006]	[0.213]	[2.120]	
Children living in household with access to sewer or septic well	0.844	-0.028 ***	-0.262 **	-2.922 **	6,655	0.803	-0.057 ***	-0.581 *	4.463 **	6,468
	[0.362]	[0.007]	[0.143]	[1.339]		[0.397]	[0.012]	[0.335]	[1.692]	
Children having access to clean water	0.825	0.006	-0.349 **	-1.873 **	7,346	0.816	0.004	-0.668 *	-3.235 **	7,261
	[0.379]	[0.008]	[0.166]	[0.922]		[0.387]	[0.012]	[0.366]	[1.308]	
Children living in a household with walls of mud	0.216	0.072 ***	0.200	1.630 *	7,346	0.270	0.077 ***	0.467	1.886 *	7,261
	[0.411]	[0.009]	[0.157]	[0.983]		[0.444]	[0.015]	[0.329]	[1.111]	
Children living in a household with earth floor	0.099	0.040 ***	0.234 **	2.334 **	7,346	0.135	0.076 ***	0.575 **	3.906 ***	7,261
	[0.299]	[0.006]	[0.116]	[1.144]		[0.342]	[0.011]	[0.299]	[1.428]	
Attachment to the Labor Market										
Mother participating in the labor market	0.529	-0.039 ***	-0.419 *	-1.164 *	6,655	0.537	-0.102 ***	-0.702 *	-1.892 **	6,468
	[0.499]	[0.012]	[0.225]	[0.643]		[0.498]	[0.018]	0.428	[0.897]	
Children working in the household or in the labor market	0.400	0.047 ***	0.607 **	1.656 **	7,346	0.391	0.702 ***	0.643 *	1.919 **	7,261
	[0.489]	[0.012]	[0.274]	[0.738]		[0.488]	[0.228]	[0.383]	[0.958]	
Hours per week worked by teenagers (conditioned on being working)	7.986	1.313 ***	5.094	__	5,460	8.264	0.749 ***	9.658 *	__	6,290
	[7.323]	[0.304]	[3.429]			[7.923]	[0.245]	[5.912]		
Health Care Utilization										
Children taken to the doctor (conditioned on being sick)	0.928	-0.010	-0.145	-1.197	3,956	0.915	-0.016	-0.558	-2.842	4,096
	[0.257]	[0.008]	[0.252]	[2.025]		[0.278]	[0.013]	[0.413]	[2.352]	
Domestic Violence										
Children affected by domestic violence	0.364	0.092 ***	0.359 **	1.322 **	5,874	0.385	0.073 ***	0.386	2.360	6,000
	[0.481]	[0.011]	[0.172]	[0.601]		[0.486]	[0.015]	[0.984]	[3.023]	

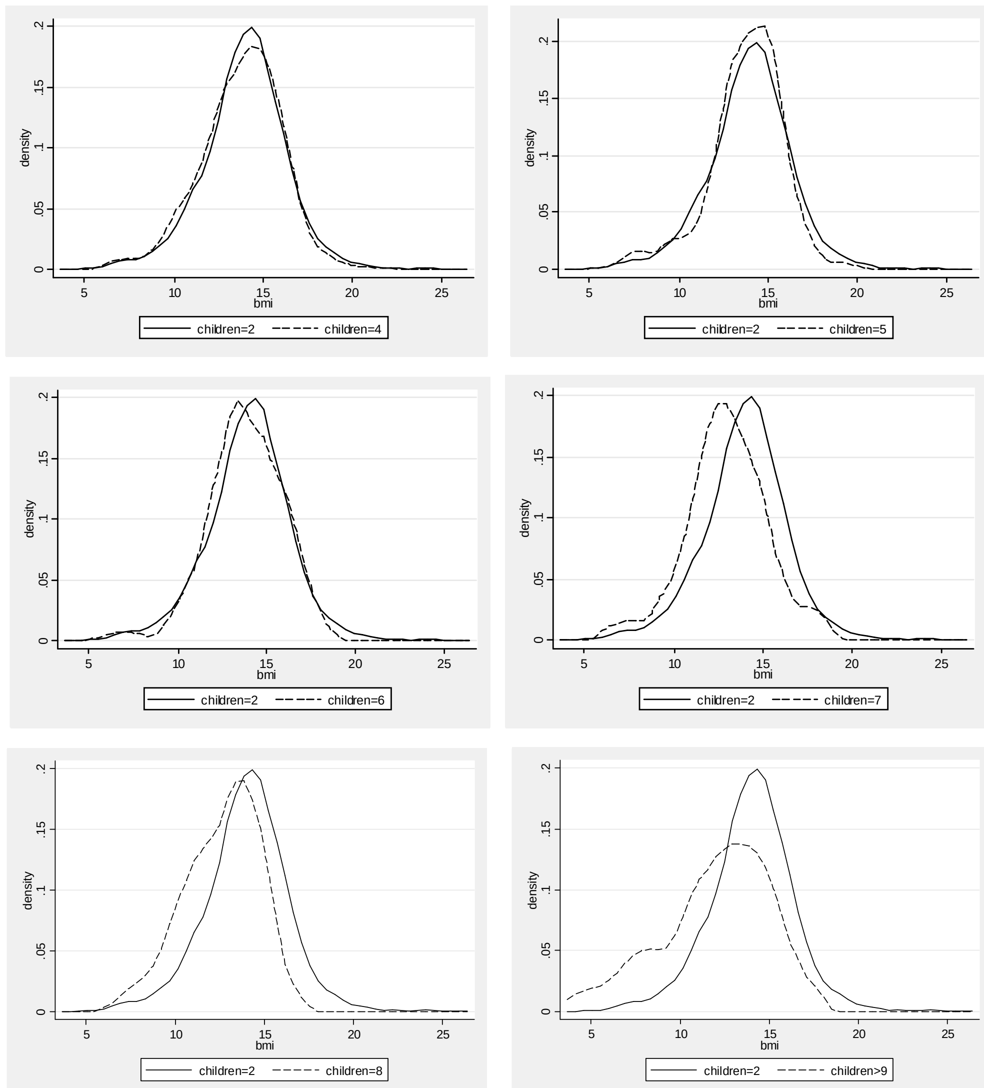
Notes: 2SLS estimators calculated by IV linear probability model and IV Probit Newey's efficient two-step minimum chi-squared procedure. Robust standard errors (clustered by household for OLS and 2SLS in the two panels) presented in square brackets. The symbols ***, (**) and [*] stand for significance at the 1, (5%) and [10%] *levels*, respectively. The first (second) panel includes estimates from households with at least two (three) children that are 18 years old or younger, are matched to their biological mothers and are living in the same dwelling. The variables *children's school attainment*, *children attending school* and *children held back in school* are restricted to children between 6 and 18 years old. The latter is a dummy variable that takes the value one for children that are in grades behind their reference cohorts due to school repetition, temporary and permanent dropout or late enrollment, and zero otherwise. The variable c*hildren having access to clean water* takes the value one for families receiving either piped water from utility company or bottled water, and zero otherwise. The variable *children working in the household or in the labor market* takes the value one for children engaged in domestic chores, family businesses, other household activities, self-employed or in remunerated and non-remunerated jobs, and zero otherwise. The variable *children affected by domestic violence* is equal to one if the children were pushed, deprived from food, hit with objects, assigned non-appropriate work, left out of the household for some time, thrown water, withdrawn economic support or witnessed actions of violence between their parents, and zero otherwise. The regressions include covariates such as the age, gender and schooling of the child, mother's marital status and age at first birth, spacing between first and second birth, mother's years of schooling, a wealth index (as described in Table 1), and a set of dummies to identify single headed families, rural households, state and municipality effects and other characteristics of the households. See text for definitions of treatment and control families.

Table 5. OLS and Probit Estimates of the Reduced Form Relationship between Family Size and Children's Anthropometric Measures, Morbidity and Vaccination

	OLS						Probit				
Outcomes	**(1)**		**(2)**		**N**	**Outcomes**	**(1)**		**(2)**		**N**
Anthropometrics											
Body Mass Index	-0.090	***	-0.072	***	5,251	**Illness Prevalence**					
	[0.019]		[0.019]								
Body Mass Index (Percentil: 0th - 50th)	-0.062	***	-0.044	***	2,626	Did children get diarrhea in the last two weeks?	0.018	***	0.010	***	5,768
	[0.022]		[0.022]				[0.006]		[0.005]		
Body Mass Index (Percentil: 0th - 25th)	-0.072	**	-0.071	**	1,312	Did children get fever in the last two weeks?	0.003		0.000		5,768
	[0.034]		[0.032]				[0.004]		[0.004]		
Weight for Age Index	0.003		0.003		5,231						
	[0.005]		[0.005]			**Immmunization**					
Weight for Age Index (Percentil: 0th - 50th)	-0.002	***	-0.002	***	2,615	Children immunized against tuberculosis?	-0.019	***	-0.014	***	5,764
	[0.000]		[0.000]				[0.003]		[0.003]		
Weight for Age Index (Percentil: 0th - 25th)	-0.002	***	-0.002	***	1,292	Children immunized against diphteria and tetanus?	-0.046	***	-0.033	***	5,399
	[0.000]		[0.000]				[0.008]		[0.008]		
Height for Age Index	0.122		0.101		5,220	Children immunized against poliomyelitis?	-0.025	***	-0.021	**	5,439
	[0.078]		[0.076]				[0.010]		[0.010]		
Height for Age Index (Percentil: 0th - 50th)	-0.011	***	-0.010	***	2,610	Children immunized against measles?	-0.024	***	-0.015	**	5,725
	[0.002]		[-0.010]				[0.006]		[0.006]		
Height for Age Index (Percentil: 0th - 25th)	-0.009	***	-0.007	***	1,305						
	[0.002]		[0.002]								

Notes: Probit coefficients reported correspond to marginal effects. Robust standard errors clustered by household presented in square brackets. The symbols *** and ** stand for significance at the 1% and 5% levels, respectively. The sample includes only children that are 4 years old or younger and belong to households with at least two children that are 18 years old or younger, are matched to their biological mothers and are living in the same dwelling. The first specification of each model includes covariates such as the age, gender and schooling of the child, birth order, mother's marital status and age at first birth, spacing between first and second birth, mother's years of schooling, a three-digit wealth index (as described in Table 1) and a dummy for single headed families. The second specification expands the first one by including a set of dummies to control for rural households, municipality effects and other characteristics of the households. The Body Mass Index (BMI) was calculated as the ratio of a child's weight in kilograms divided by height in meters squared. The Weight for Age and Height for Age indexes were computed as the ratio of child's weight (in kilograms) and height (in centimeters) divided by child's age in months, respectively. Dummies for immunization against diphtheria, tetanus and poliomyelitis take value one for children under four that have received the three doses, and zero otherwise.

Figure 1. Body Mass Index Density Estimates by Family Size
(Families with More than Two Children vs. Families with Two Children)



Notes: All density estimates based on a standard Epanechnikov kernel function and a width that minimizes the mean integrated squared error of the data. The Body Mass Index (BMI) was defined as the ratio of a child's weight in kilograms divided by his/her height in meters squared.

.