New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country

Vladimir Ponczek
Andre Portela Souza

ABSTRACT

This paper presents new evidence of the causal effect of family size on child quality in a developing-country context. We estimate the impact of family size on child labor and educational outcomes among Brazilian children and young adults by exploring the exogenous variation of family size driven by the presence of twins in the family. Using the Brazilian Census data for 1991, we find that the exogenous increase in family size is positively related to labor force participation for boys and girls and to household chores for young women. We also find negative effects on educational outcomes for boys and girls and negative impacts on human capital formation for young female adults. Moreover, we obtain suggestive evidence that credit and time constraints faced by poor families may explain the findings.

I. Introduction

This paper presents new evidence of the causal effect of family size on child quality in a developing-country context. We explore dimensions of child quality that are prevalent in developing countries and have not previously been examined by the literature, which has focused on the developed world. The literature in economics has long discussed the relationship between family size and child quality. It has been argued that a tradeoff exists between the quantity and quality of children (Becker and Lewis 1973; Becker and Tomes 1976; Hanushek 1992). In

Vladimir Ponczek is an associate professor at the São Paulo School of Economics—FGV. André Portela Souza is an associate professor at the São Paulo School of Economics—FGV Rua Itapeva, 474 12 andar-01332-000-Sao Paulo, SP-Brazil. *They thank four anonimous referees for the helpful comments. They also thank participants at EESP-FGV, PUC-RIO, LACEA (2007, Bogotá) and SBE congresses and seminars for their comments and advice. The usual disclaimers apply. The data used in this article can be obtained beginning beginning August 2012 through July 2015 at vladimir.ponczek@fgv.br.*
[Submitted June 2009; accepted March 2011]
ISSN 022-166X E-ISSN 1548-8004

general, child quality is understood as any child outcome that is valued by the parents. Authors have considered the well-being of children, or their accumulation of human capital. Becker and Lewis (1973) developed a model that introduces a theoretical framework to analyze this issue. They assume that the cost of an additional child, when holding quality constant, increases with the number of children. Similarly, the cost of increasing the average quality of a child, when holding quantity constant, rises as quality increases. An important implication of such models is that family size becomes an input in the production of child quality.

Any empirical exercise that attempts to estimate the causal effect of family size on child quality must take into consideration the endogeneity of fertility. The empirical literature that is concerned with the endogeneity problem focuses on educational outcomes as child quality measures. The results are mixed for industrialized countries. Black et al. (2005) find no impact of family size on individual educational achievement in Norway. Haan (2005) finds no significant effect of the number of children on educational achievement in the United States and the Netherlands. Angrist et al. (2005), Angrist et al. (2006) do not find any causal impact of family size on completed educational achievement and earnings in Israel. Conley and Glauber (2005) use the 1990 US PUMS to estimate that children living in larger families are more likely not to attend private school and are more likely to be held back in school. Caceres-Delpiano (2006) finds a negative impact of family size on the likelihood that older children attend private school, but he finds no significant results on grading retention in the United States. Additionally, Goux and Maurin (2005) show that children living in France in larger families perform worse in school than children in smaller families.

For developing countries, using data from India between 1969 and 1971, Rosenzweig and Wolpin (1980) estimate that households with higher fertility rates have lower levels of children's schooling. Lee (2004) finds negative impacts of family size on per-child investment in education for South Korean households. Qian (2009) explores relaxations in China's One Child Policy if the first child is a girl to show that an additional sibling in the family of Chinese first-borns can result in an increase in school enrollment. She finds a greater impact if the second child is a girl. In contrast, Li et al. (2007) find a negative effect of family size on Chinese children's education.

Differently from previous studies, we extend the analysis to a richer set of child-quality measures. In addition to educational outcomes, we investigate the impact of family size on labor market participation and domestic work among Brazilian children and young adults. The literature has emphasized the potential detrimental effects of child labor on the individual's well-being. Moreover, we analyze possible channels through which family size may impact child quality. In particular, we investigate the effectiveness of the credit and time constraint channels that are more pervasive in a developing-country context.

In principle, the impact of family size on child quality can be harmful or beneficial. A situation can be imagined in which a larger family will have more diluted resources. For instance, in an environment where credit markets are imperfect, families with many children would invest less in each child than if they had fewer children. However, it is possible that having more children decreases maternal labor supply (Angrist and Evans 1998). As argued by Blau and Grossberg (1992), this reduction may increase the probability of the mother spending more time parenting, which may improve child quality. Therefore, measuring the impact of family size on child-quality outcomes is an empirical endeavor.

Child quality and quantity are jointly determined. To consistently estimate the causal effects of family size on child quality, we make use of the instrumental variable technique. We explore the exogenous variation of family size driven by the presence of twins in the family. We believe that this a plausible instrument based both on the prior literature and the fact that assisted fertilization techniques were neither widely practiced nor available in the time period we study. Using the Brazilian Census data for 1991, we find that this exogenous increase in family size is positively related to labor force participation for boys and girls and to household chores for young women. Moreover, we find negative effects on educational outcomes for boys and girls and negative impacts on human capital formation for young female adults.

It is possible that the quantity and quality tradeoff is more acute in environments where credit constraints are more pervasive. In developing countries, where credit markets are imperfect, parents cannot easily smooth out family consumption and resource allocation over time.[1] Therefore, the resource dilution induced by an extra child in the family may alter the time allocation of the children. This phenomenon may not occur in developed countries because credit markets make consumption-smoothing over time possible.

Correctly estimating the causal effect of family size on child-quality outcomes is important for a developing country's public-policy perspective. The majority of large families are poor, and our results suggest that family size has a direct impact on important outcomes for children. This discussion can better inform the public debate about how to understand and address poverty, education, and child labor in developing countries.

The paper proceeds as follows. Section II discusses the dimensions of child quality explored in this paper. Section III describes the data set and the sample selection. The identification strategy is presented in Section IV. The results are discussed in section V. Section VI concludes.

II. Measures of Child Quality

Child quality is multidimensional and related to a child's well-being. In practice, empirical studies often restrict their attention to indicators of human capital accumulation, usually measured by educational outcomes. However, in environments where the quantity and quality tradeoff is more acute, other dimensions of child quality may be as relevant as the child's formal education. One dimension that is frequently neglected is child labor. Child labor is a common phenomenon in developing countries in both market and domestic work. It is often used by families to complement their total resources (see Basu 1999; Edmonds and Pavcnik 2005,

1. For instance, Beegle et al. (2006) show that negative income shocks are associated with child labor incidence in Tanzania but that these effects can be mitigated by household asset holdings.

2006; Edmonds 2008 for surveys on child labor) and is typically associated with lower human-capital formation. The theoretical literature emphasizes the tradeoff between child labor and human-capital accumulation. The main channels are time constraints (the child has less time to acquire education) as well as physical and psychological constraints (the child is less capable of learning after hours of work) (see Baland and Robinson 2000). Moreover, child labor can harm short and long run health outcomes. Working children may face health threats by the nature of the work such as insalubrity and hazardous conditions. Therefore, child labor can be characterized as an important determinant of child well-being.

There is evidence that child labor is harmful for the individual's well-being, especially for young children. For instance, Beegle et al. (2005) find damaging effects of child labor on schooling outcomes, but no effects on health outcomes in Vietnam, and Emerson and Souza (2011) find a detrimental impact of child labor on adult earnings over and above the benefit of education in Brazil. For older children, the literature has no clear evidence on the harmful effects of labor force participation. This lack of evidence may be due to the fact that older children might benefit from the productivity gains caused by on-the-job training. For health outcomes, Cigno and Rosati (2005) and Orazem and Lee (2010) find detrimental effects of child labor in Guatemala and Brazil, respectively. Therefore, our findings reveal the importance of not overlooking other dimensions of child quality because they corroborate the idea that an unexpected additional child may harm the human capital formation of the child and its siblings, thus perpetuating intergenerational poverty traps.

The advantage of exploring the quantity-quality tradeoff in a developing-country context is that credit rationing is more pervasive. An extra child may impose a larger resource dilution without the possibility of consumption-smoothing over time in a family. Fewer resources may be allocated to other siblings. Moreover, the time endowment of family members may be reallocated to domestic work or labor market activities rather than to the human capital accumulation of the children. This effect may be more easily detected in environments such as Brazil, where it is possible to combine child labor and schooling. In fact, in Brazil, formal schooling occurs for four hours a day. Indeed, 12.65 percent of all children aged between ten and 15 years old worked in Brazil in 1991, and 41.34 percent of those also attended schools.

Our measures of child quality include a set of child labor and educational indicators. Specifically, they are: (i) labor force participation, (ii) household chores, (iii) school attendance, (iv) school progression, (v) literacy, (vi) high school completion, (vii) college attendance and (viii) completed years of schooling for those who are not currently attending school.

III. Data

The data used were obtained from the 1991 Brazilian Census micro database, collected decennially by the Brazilian Census Bureau (IBGE). For each Census, IBGE draws a random sample of the households that contains extensive information on personal and household characteristics. For each person, information about, for example, age, schooling attendance, literacy, years of completed schooling, migration, labor participation, retirement, and income sources (including values) is available. The 1991 random sample contains 10 percent of the households in municipalities with more than 15,000 inhabitants and 20 percent of those in the smaller municipalities, totaling around five million households. We choose to use the Census database for two main reasons. First, a twin birth is a rare event, and we need a large sample size to obtain a sufficient number of observations of twins. Second, as further discussed below, we choose the 1991 Census to avoid twin births that arose from in vitro fertilization treatments (IVF), which became generally available later on in Brazil.

We follow the literature of (Angrist et al. 2005, Angrist et al. 2006, Black et al. 2005 and Caceres-Delpiano 2006) to create two samples based on the birth order of twin occurrence. The first sample consists of children in families with two or more births (2+ sample). The instrument is a binary variable that indicates whether the second birth is a multiple birth. We restrict this sample to first-born children. We examine children born before the twin occurrence for two reasons. The first reason is to avoid including the twins themselves because twins have special characteristics that might directly affect the outcomes of interest other than the family size. It is well-known that multiple-birth children are more likely to have low birth weights and higher morbidity rates (see Behrman and Rosenzweig 2004). Second, we do not include children born after the occurrence of twins to avoid possible post-treatment effects correlated with the outcome of interest.[2]

The second sample includes children in families with three or more births (3+ sample). In this case, the instrument indicates whether the third birth is a multiple birth. We look at two subsamples. One subsample includes first- and second-born children. The other subsample is restricted to first-born children only (3+ sample—first-borns).

Both the 2+ and 3+ samples include individuals of two age groups: One group includes children between ten and 15 years of age only, and the second includes young adults between 18 and 20 years of age. The reason for the restriction in the first group is that the Census does not have the information about labor market participation for children under the age of ten. We restrict the children's age to be at most 15 because at that age an individual is expected to have completed the primary- and middle-school cycle (ensino fundamental) and thus completed mandatory schooling. The second group was created to analyze the existence of lasting effects of family size (quantity) on the individual's human capital formation (quality). We restrict the sample to individuals who are old enough to complete high school but likely to still live with their parents. As we explain below, the Census only keeps information on the sibling composition of those individuals who co-reside with their relatives. Thus, we restrict the second group to the 18–20 year age range because 70 percent have not moved out of their parents' homes yet.

All samples are restricted to families with two adults (the mother and her husband), whose eldest child is younger than 16 years old (first group) or 21 years old (second group). These selection criteria are intended to avoid the potential endogenous decision about the number of adults living together in the same family. There-

2. We exclude families with more than one twin birth occurrence and families with twin births other than the nth birth. We select our sample in this way to avoid twin-birth effects in the control group.

fore, we are only looking at families composed of two adults and their children. Because we are interested in the tradeoff between quality and quantity of children, we want to explore the variation in family size due to variation in the number of children only. Hence, we fix the number of adults in the family. The natural choice is to select families with the presence of the mother and her husband. First, the presence of the mother is essential to the identification of siblings. Second, those families represent the majority of families in Brazil. Another possibility would be to select single-mother families. In fact, we selected a sample of single-mother families, but we ended up with very few twin birth observations. For instance, in the 2+ sample, we found only 254 families with twin births. In comparison, there are 3,599 families with twin births in the two-adult-family 2+ sample. Not surprisingly, we did not find a significant impact of family size for most of the single-mother sample regressions. The results are available upon request.[3] For the first age group, the family-size variation comes from the number of children in the family between zero and 15 years old, and its impacts are estimated on 10–15-year-old first-born children in the case of the 2+ sample and on 10–15-year-old first- and second-born children in the case of the 3+ sample. For the second age group, the family-size variation comes from the number of children in the family between zero and 20 years of age, and its impacts are estimated for 18–20-year-old individuals.[4]

The Brazilian Census allows identification of the mother of the child as long as the mother lives in the same household. Therefore, we classify individuals as twins if the children are (i) living in the same household, (ii) from the same mother and (iii) the same age.[5]

Ideally, we want to have the exact date of birth to classify twins. Unfortunately, the Census does not provide information on the exact date of birth. Therefore, it is possible that two nontwin siblings replied as having the same age, leading to a birth interval of between nine and 11 months to be misclassified as twins in our sample. It is possible that families with more closely spaced children have lower socioeconomic status. In this case, the measurement error would be correlated with the outcomes of interest. However, we do not believe that this measurement error is a major problem. We calculated the proportion of twin births using other Brazilian household surveys that contain the exact birth date for the same period and compared to the proportions obtained in the Census.[6] They are quite similar. For this reason, we believe that this type of measurement error is not very severe.

3. This sample selection also avoids the problem of having an extra adult member (for exemple, other relatives) that is related to past fertility decisions and to current children's time allocation decisions. However, we are aware that these selection criteria might bring other potential selection biases that might jeopardize the external validity of our findings. Nevertheless, we do not believe that this is a serious problem in our analysis because of all of the families with children aged 15 or younger, only 15.1 percent are not two-adult families, and of all families with children aged 20 or younger, only 15.2 percent are not two-adult families.

4. Hereafter, the expressions family size and number of children will be used interchangeably to designate the same variable.

5. We exclude from the samples all families in which the mother is dead or not present. We also exclude families with, for example, triplets, quadruplets, and quintuplets.

6. We use the Brazilian National Household Survey (PNAD) of 1992 from IBGE

We search for the presence of twins among all children younger than 16 and younger than 21 years of age for the first and second group, respectively. Our instrumental variable for the number of children is the occurrence of twin births in the *n*th birth.

Finally, to avoid measurement error problems due to children who have died or who live in other households, we exclude families in which the mother has children living outside her household or has had children who are no longer alive. In the first age group, 24 percent of the families are excluded by these circumstances. This figure increases to 30 percent in the second group. This exclusion may create selection biases such that the results cannot be generalized to the entire population. We compare socio-economic characteristics of both samples (excluded and nonexcluded families). We find that older and less-educated parents comprise the excluded families. In fact, these differences are not surprising because the children living outside the parent's household are more likely to be older, and consequently, their parents are older as well. In addition, individuals from older cohorts have lower educational attainment in Brazil. The results are presented in the Appendix.[7] Table A2 compares the characteristics (parents' age and education, child's age and race, and urban and metropolitan area indicators) of families with and without the presence of twins for all samples separately. We find similar figures for both types of families for most of the characteristics. However, there are differences in the parents' education and race. Families with twins have parents who are slightly less educated and a greater proportion of nonwhite children. Although the differences are present in all samples, they are relatively small compared to standard deviations.[8]

We cannot test the exogeneity assumption of the instrument. However, the randomness of twin births, the sample selection of nontwins born before the birth of twins, the conditioning of the twin occurrence in a particular birth, and the use of the family characteristics as additional controls in all regressions make the validity of the instrument reliable.

IV. Empirical Strategy

An important aspect to be considered is that child quality and quantity are jointly determined by the parents. For instance, in the Becker and Tomes (1976) model, families decide how many children to have and how much to invest in each child. Given the nonlinear constraints, an exogenous increase in the number of children raises the per-child cost of quality. Consequently, the model implies that there is a negative relationship between quantity and quality. Moreover, both decisions can be correlated with unobserved variables. For example, parents' endowments (such as ability, wealth, education and cultural factors) affect child quality via

7. Notice that we do not have information on the characteristics of the children living outside their parents' household. In particular, we do not know their age and, therefore, whether they are twins or not. Hence, we cannot compare the probability of the presence of twins in a family among excluded and nonexcluded families.

8. Caceres-Delpiano (2006) finds differences in the mothers' education and age and children's race among families with and without twins in the United State.

intergenerational transmission mechanisms. Poorly endowed parents may produce poorly endowed children who benefit less from an extra investment in their quality compared to highly endowed children. If this is the case, parents with low endowments would optimally decide to have more children and lower quality per child compared to highly endowed parents. It is also possible that ability and taste factors not captured by the controls may exert an influence on both quantity and quality, then simple Ordinary Least Square (OLS) estimators of the effects of family size on child quality will be biased. By the same token, wealth and ability are determinants of being able to afford and correctly understand the use of birth-control methods. If these determinants are also correlated with the children's educational and labor outcomes, the OLS estimator will not capture the causal impact of family size on these quality indicators. Depending on the correlation between unobservables and family size and between unobservables and the dependent variables, the OLS estimator might be upward- or downward-biased. If the above example about the relationship between ability and the value of education and birth-control measures is true, then we would expect that a naive approach would overestimate the actual impact of family size on education. However, it can be imagined that the parents' decision to have another child is made after a positive income shock or expected increase in future resources, which can offset part of the extra burden. In this case, the OLS estimator would underestimate the effect of having the extra child in the family.

Analogously, child labor and fertility are ambiguously related. The Baland and Robinson (2000) model shows that when fertility is exogenous, an increase in family size decreases the amount of work performed by each child. This phenomenon may occur because having more children may increase total family income, reducing the required labor intensity per child. However, if fertility and child labor are jointly determined, then the direction of causality is not clear. On the one hand, increased child labor reduces the net cost of a child, raising the demand for children. On the other hand, increasing fertility increases the total cost of all children and requires more child labor to generate the extra income. In addition, Cigno and Rosati (2005) illustrate a model for which wealth and fertility are negatively correlated through birth-control costs. In this case, the relationship between fertility and child quality would be driven by a third factor. The causal effect of the former on the latter would not necessarily be present.

To the best of our knowledge, the literature lacks studies on the determinants of child labor that correctly take into account the endogeneity of family size. For instance, Psacharopoulos and Patrinos (1997) find that having more young siblings is associated with less schooling, greater age-grade distortion, and less child labor among Peruvian children in 1991. Cigno and Rosati (2002), studying the determinants of child labor and education in rural India, find a significantly positive effect of the number of children aged 6–16 on the time used to work and a negative effect on the time devoted to attending school. Although these works are an important step for understanding the determinants of child labor, the potential endogeneity of fertility can bias their results and generate misleading conclusions. The only attempt to address the endogeneity problem of the relationship between child labor and fertility is in Deb and Rosati (2004). They use the gender of the first-born child, the ages of the parents, and the village-level mortality rate as instruments of fertility. They find a positive effect of number of children on the probability of work when endogeneity is taken into account. This result is different from the case where fertility is assumed to be exogenous. In this case, they find a negative and insignificant effect on child labor. Although their study attempts to deal with the endogeneity of fertility, we doubt that the instruments have the indispensable characteristic of being orthogonal to the unobservables. It is very likely that the instruments, especially parents' ages and the village mortality rate, are correlated to wealth, ability, and other unobservable variables that might be jointly related to child labor and fertility, jeopardizing their results and conclusions.

We therefore need a source of variation of family size orthogonal to any unobserved characteristics of the households and, at the same time, related to the dependent variables. The Instrumental Variable (IV) approach is able to generate a consistent estimator as long as the excluded instrument is not correlated with the unobserved characteristics but plays an important role in explaining the endogenous variable.

The presence of twins in a family has the two characteristics expected of a good IV. It is clearly correlated with family size and, because it is very likely to be a random occurrence, tends to be orthogonal to the error term in the main regression. A potential flaw of our strategy arises if there is any independent effect of the presence of twins on quality that does not operate through quantity. The presence of twins is directly associated with narrower spacing among sibling births. If average spacing is correlated with child-quality outcomes, then our instrument will not be valid. For instance, it is possible that breastfeeding twins may physically exhaust the mother, which may affect the raising of other children in the family. If this is the case, then the impact of family size on quality will be overestimated.

Our benchmark strategy consists of a Two Stage Lesat Square (2SLS) regression where, in the first step, we regress number of children (N_{ij}) on the presence of twins on the *n*th birth indicator variable (*PT*) and other predetermined variables (*W*):

$$(1) \quad N_{ij} = \alpha + \hat{\beta} PT_j + \gamma' W_{ij} + \varepsilon_{ij}$$

The second step follows[9]:

$$(2) \quad Y_{ij} = \alpha + \hat{\beta} N_j + \gamma' W_{ij} + \nu_{ij}$$

where Y_{ij} is the outcome of interest of children *i* living with family *j*.[10] The outcomes of interest for the 10–15-year-old age group are as follows: (i) labor force participation, defined as a binary variable that indicates whether the child participates in the labor market. Individuals are considered to be in the labor market if they have regularly or occasionally worked during the last 12 months or if they are currently searching for a job;[11] (ii) household chores, defined as a binary variable that indicates

9. When calculating the variance and covariance matrices of ε and ↑, we allow for correlation of residuals within the family unit for the 3+ sample.

10. First-born children in the 2+ sample and first-and second-born children in the 3+ sample.

11. The Census questions whether an individual has worked for part or all of the last 12 months, and the answer options are (i) regularly, (ii) occasionally or (iii) not worked. For those who answered that they have not worked, it also asks if they were searching for a job. Alternatively, we also tried a different definition whereby those who answered "occasionally" were not considered participants. The results are not sensitive to the definition of participation in the labor market.

whether the individual does household chores as a main activity, conditional on the event that the individual is not participating in the labor market;[12] (iii) attendance, defined as a binary variable that indicates whether the individual is currently attending school; (iv) school progression, defined as age-grade distortion (*years of schooling*)/(*age*-6); (v) literacy, defined as a binary variable that indicates whether the individual knows how to read and write. Because we want to have other measures of quality that include completed educational attainments, we use the 18–20-year-old age group to investigate the impact of family size on additional outcomes; (vi) high school completion; (vii) college attendance; and (viii) completed years of schooling for those who are not currently attending school. Finally, *W* is a vector of control variables used for both age groups that includes the following: child's gender as well as age and its square; indicator variables for whether the child is white, lives in a urban area, or lives in a metropolitan area; family head's years of schooling, age, and gender; mother's years of schooling and age; and year indicator variables capturing any ongoing trend on the dependent variables. We run separate regressions for each gender.

The interpretation of the labor force participation as one dimension of child quality is less clear for young adults. Although for children, the literature considers child labor detrimental to human capital formation, this may not be so for young adults. It is possible that it helps the human capital accumulation through labor market experience and a learning-by-doing process. We retain this outcome for the young adults for two reasons. First, we want to compare the results with the 10–15-year-old age group results. Second, we may conjecture about the detrimental effect of labor force participation when considering the results together with those for schooling outcomes.

V. Results

Our approach consists of using the presence of twins in the family as an instrument for family size. Although the birth of twins seems to be a random event, important endogeneity issues must be addressed. A possible reason why twins may not be random is the choice of fertility treatments by the parents, such as in vitro fertilization pre-embryo transfer (IVF). The medical literature estimates that 25 percent of pregnancies with IVF are twins when multiple preembryos are transferred.[13] In general, families that make use of IVF treatments are more likely to have natural pregnancy difficulties and stronger preferences for having children. Moreover, these treatments are costly or provided by private health insurance. These services were not generally offered by the public health service in Brazil until 2005. Those willing to undergo treatment had to afford it privately. Therefore, it is likely that the characteristics of the families that are correlated with the occurrence of twins (via IVF) are also related to the children-quality outcomes. Therefore, if the

12. The Census only asks whether an individual does household chores to those who have not studied or worked in the last 12 months. Note that job search and household chores are also exclusive options.
13. See http://www.ivf.com

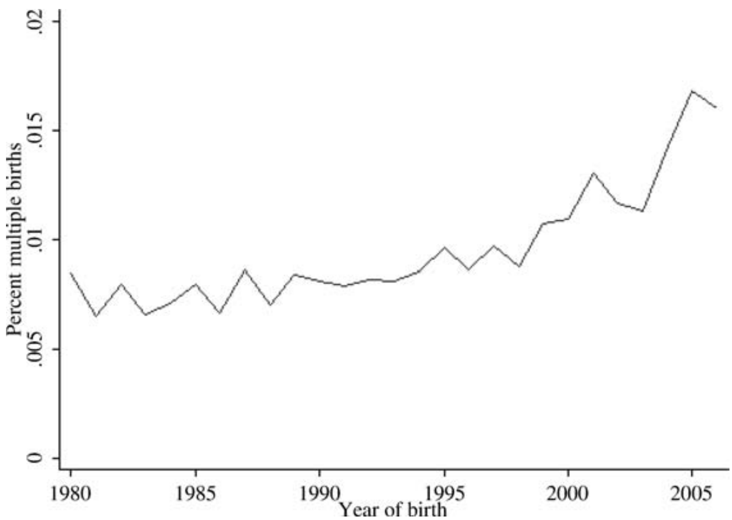

Figure 1
Evolution of Multiple Births

IVF treatment is pervasive, the instrument might not be orthogonal to the error term of the second-stage regression.

Does the surge in fertility treatments jeopardize the randomness of a twin's birth in our analysis? IVF treatments became available in the mid-1980s in Brazil but became generally accessible privately after the mid-1990s (Borlot and Trindade 2004)[14]. Our sample does not observe whether a birth is due to an IVF treatment. To check the influence of the availability of this procedure on the occurrence of twins, we analyze the evolution of the occurrence of twins across time. A relative increase in the proportion of families with twins in the period after the IVF became broadly available would suggest an influence of the fertility treatments on the instrument. Using the annual data from the Brazilian Household Survey (PNAD) also collected by IBGE, we show the evolution of these figures in Brazil. Figure 1 depicts this trend. Indeed, there is an increase in the proportion of twins after the 1990s. For instance, on average, of all births occurring from 1980 to 1991, 0.72 percent were twin births. However, of all births after that, 1.04 percent were multiple births. In the 1991 Census, the proportion of twins among births after 1980 was 0.87 percent. The small positive difference in the Census data may be due to our definition of twins, which only uses the siblings' age and not the exact date of birth.

Because IVF is a costly procedure, we check whether twin occurrence before and after the mid-1990s is correlated with some observed measure of family wealth. We calculate the correlation coefficient between the proportion of multiple births and

14. The world's first IVF case was in 1978 in England. The first case in Brazil only occurred in 1984 The first private human reproduction clinic opened only in 1989

average years of schooling of the mothers of twins. The value is 0.16 (and not significant at 5 percent) before 1994 and 0.90 (and significant at 5 percent) in 1994 and thereafter.

These findings suggest that some twin births that occurred after the IVF treatments became accessible may not be random events. Although in our regressions we control for parental education and other characteristics that are proxies for permanent income and wealth, it is still possible that unobserved characteristics (for exemple, tastes) may be correlated with the occurrence of twins and child-quality outcomes. To avoid this potential problem, we decided to use the 1991 Census, in which all the births occurred before the surge of IVF treatments in Brazil.

A. Impacts on the 10–15-Year-Old Age Group

Table 1 presents the descriptive statistics of the 2+, 3+, and 3+—first-born samples for the 10–15-year-old age group for boys and girls separately. We notice that labor-force participation for boys is almost twice as large as that of girls. For instance, in the 2+ sample, 14 percent of all boys participate in the labor market, whereas 6 percent of all girls are participants in the labor market. Conversely, the incidence of household-chore activities is greater among girls compared to boys. Of all first-born girls who are not participating in the labor market, 7 percent report doing some chores in the 2+ sample. This figure is 1.5 percent for boys. Regarding educational outcomes, 85 percent of the boys and 87 percent of the girls are attending school in the 2+ sample. As is commonly observed in developing countries, children experience some school delays. This phenomenon occurs for both boys and girls and is slightly more frequent among boys. The average school progression indices are 0.71 and 0.76 for boys and girls in the 2+ sample, respectively. Finally, 91 percent of the boys and 94 percent of the girls in the 2+ sample are literate. Of all families with two or more births, 1.3 percent have twins in the second birth. The same figure appears in the 3+ sample. The average family size is 3.1 for the 2+ sample and 3.8 for the 3+ sample. Approximately 74 percent live in urban areas in the 2+ sample and 69 percent in the 3+ sample.

Table 3 presents the OLS regression results for all three samples of the 10–15-year-old age group. They are qualitatively the same for boys and girls. All show positive and significant coefficients of family size on labor-force participation and household chores. Negative and significant coefficients are also observed for attendance, school progression and literacy.

Overall, the OLS results indicate that individuals in larger families are more prone to engage in labor-market activities and chores (especially girls) and present lower educational outcomes. These figures suggest a strong detrimental effect of family size on child-quality outcomes.

Are these results reliable? First, we check whether the IV has a strong correlation with the potential endogenous variable. Table 5 displays the first-stage results of the IV regressions for the 10–15-year-old age group. All coefficients of the instrument are positive and statistically significant. We also notice that families whose child is white, whose parents are more educated, and who live in urban and metropolitan areas have fewer children (not shown in the tables). The same results are obtained in the first-stage regressions of the 18–20-year-old age group and are shown in Table 6.

Table 1
Summary statistics—10–15-years-old

Variable	2+ Sample			3+ Sample			3+ Sample First Born		
	Mean	Standard deviation	N	Mean	Standard deviation	N	Mean	Standard deviation	N
Boys									
Labor force participation	0.14	0.35	148,242	0.15	0.36	145,296	0.17	0.38	92,272
Household chore activities	0.02	0.12	129,844	0.02	0.13	126,484	0.02	0.13	78,190
School progression	0.71	0.31	148,155	0.64	0.32	145,231	0.64	0.32	92,214
Literacy	0.91	0.29	148,242	0.87	0.34	145,296	0.87	0.34	92,272
Attendance	0.86	0.35	148,292	0.83	0.38	145,296	0.82	0.39	92,272
Family size	3.12	1.28	148,242	3.83	1.22	145,296	3.79	1.19	92,272
Presence of twins	0.01	0.11	148,242	0.01	0.11	145,296	0.01	0.11	92,272
Head's schooling	5.35	4.40	148,201	4.57	4.15	145,256	4.59	4.16	92,244
Head's gender	0.83	0.37	148,242	0.86	0.34	145,296	0.86	0.33	92,272
Head's age	38.68	6.16	148,242	38.60	6.06	145,296	38.37	6.09	92,272
Mother's schooling	5.41	4.10	148,143	4.67	3.87	145,192	4.87	3.88	92,207
Mother's age	34.65	4.90	148,242	34.82	4.63	145,296	34.10	4.66	92,272
Child's age	12.35	1.70	148,242	12.11	1.63	145,296	12.50	1.70	92,272
White	0.60	0.49	148,242	0.53	0.50	145,296	0.53	0.50	92,272
Urban area	0.74	0.44	148,242	0.68	0.47	145,296	0.75	0.43	92,272
Metropolitan area	0.27	0.44	148,242	0.23	0.42	145,296	0.23	0.42	92,272
Second child	—	—	—	0.33	0.40	145,296	—	—	—

Girls

Labor force participation	0.06	0.23	139,360	0.06	0.23	138,815	0.07	0.25	87,635
Household chore activities	0.07	0.25	133,195	0.09	0.28	132,449	0.09	0.29	82,859
School progression	0.76	0.30	139,302	0.71	0.31	138,768	0.71	0.31	87,597
Literacy	0.94	0.24	139,360	0.91	0.29	138,815	0.92	0.28	87,635
Attendance	0.88	0.33	139,360	0.85	0.36	138,815	0.84	0.36	87,635
Family size	3.13	1.27	139,360	3.83	1.21	138,815	3.79	1.17	87,635
Presence of twins	0.01	0.11	139,360	0.01	0.12	138,815	0.01	0.11	87,635
Head’s schooling	5.37	4.38	139,311	4.59	4.14	138,772	4.63	4.14	87,606
Head’s gender	0.84	0.37	139,360	0.85	0.36	138,815	0.85	0.35	87,635
Head’s age	38.65	6.16	139,360	38.58	6.07	138,815	38.32	6.07	87,635
Mother’s schooling	5.43	4.09	139,260	4.69	3.86	138,728	4.74	3.86	87,583
Mother’s age	34.61	4.89	139,360	34.29	4.63	138,815	34.03	4.64	87,635
Child’s age	12.33	1.70	139,360	12.10	1.62	138,815	12.46	1.69	87,635
White	0.62	0.49	139,360	0.55	0.50	138,815	0.55	0.50	87,635
Urban area	0.74	0.44	139,360	0.69	0.46	138,815	0.69	0.46	87,635
Metropolitan area	0.27	0.45	139,360	0.23	0.42	138,815	0.23	0.42	87,635
Second child	—	—	—	0.34	0.41	138,815	—	—	—

Table 2
Summary statistics—18–20-years-old

Variable	2+ Sample			3+ Sample			3+ Sample First Born		
	Mean	Standard deviation	N	Mean	Standard deviation	N	Mean	Standard deviation	N
Boys									
Labor force participation	0.70	0.46	38,948	0.72	0.45	34,847	0.72	0.45	28,430
Household chore activities	0.05	0.21	14,122	0.04	0.21	12,034	0.04	0.20	9,681
School progression	0.57	0.28	38,881	0.53	0.29	34,784	0.53	0.29	28,381
Literacy	0.93	0.26	38,948	0.91	0.28	34,847	0.91	0.28	28,430
Attendance	0.40	0.49	38,948	0.37	0.48	34,847	0.37	0.48	28,430
Completed years of schooling	6.02	3.56	23,265	5.49	3.47	21,740	5.59	3.51	17,850
High school	0.17	0.38	38,948	0.13	0.33	34,847	0.14	0.34	28,430
College attendance	0.08	0.27	38,948	0.06	0.23	34,847	0.06	0.24	28,430
Family size	3.68	1.76	38,948	4.32	1.70	34,847	4.29	1.68	28,430
Presence of twins	0.01	0.12	38,948	0.01	0.12	34,847	0.01	0.11	28,430
Head's schooling	4.80	4.33	38,943	4.32	4.11	34,840	4.31	4.11	28,425
Head's gender	0.81	0.35	38,948	0.85	0.34	34,847	0.84	0.31	28,430
Head's age	45.60	6.12	38,948	45.29	5.96	34,847	45.21	5.96	28,430
Mother's schooling	4.62	3.96	38,932	4.18	3.78	34,833	4.18	3.79	28,419
Mother's age	41.36	3.96	38,932	40.88	4.50	34,847	40.79	4.52	28,430
Child's age	18.88	0.81	38,948	18.78	0.79	34,847	18.89	0.81	28,430
White	0.63	0.48	38,948	0.57	0.50	34,847	0.58	0.49	28,430
Urban area	0.74	0.44	38,948	0.70	0.46	34,847	0.70	0.46	28,430
Metropolitan area	0.30	0.46	38,948	0.26	0.44	34,847	0.26	0.44	28,430
Second child	—	—	—	0.14	0.30	34,847	—	—	—

Girls

Labor force participation	0.46	0.50	28,633	0.45	0.50	25,709	0.46	0.50	20,603
Household chore activities	0.24	0.42	16,747	0.26	0.44	15,115	0.27	0.44	11,996
School progression	0.68	0.27	28,585	0.64	0.28	25,656	0.64	0.28	20,562
Literacy	0.96	0.19	28,633	0.96	0.20	25,709	0.96	0.21	20,603
Attendance	0.53	0.50	28,633	0.51	0.50	25,709	0.51	0.50	20,603
Completed years of schooling	7.26	3.75	13,370	6.71	3.72	12,477	6.82	3.76	10,809
High school	0.28	0.45	28,633	0.22	0.42	25,709	0.24	0.43	20,603
College attendance	0.15	0.35	28,633	0.11	0.31	25,709	0.12	0.32	20,603
Family size	3.56	1.65	28,633	4.22	1.61	25,709	4.16	1.57	20,603
Presence of twins	0.01	0.12	28,633	0.01	0.12	25,709	0.01	0.12	20,603
Head’s schooling	5.33	4.47	28,635	4.79	4.29	25,703	4.85	4.30	20,597
Head’s gender	0.84	0.32	28,633	0.86	0.33	25,709	0.84	0.33	20,603
Head’s age	45.60	6.06	28,633	45.25	5.85	25,709	45.12	5.87	20,603
Mother’s schooling	5.06	4.07	28,617	4.59	3.91	25,692	4.65	3.92	20,591
Mother’s age	41.50	4.73	28,617	40.99	4.42	25,709	40.89	4.43	20,603
Child’s age	18.85	0.81	28,633	18.74	0.78	25,709	18.85	0.81	20,603
White	0.67	0.47	28,633	0.61	0.49	25,709	0.62	0.49	20,603
Urban area	0.78	0.41	28,633	0.74	0.44	25,709	0.75	0.44	20,603
Metropolitan area	0.33	0.47	28,633	0.30	0.46	25,709	0.30	0.46	20,603
Second child	—	—	—	0.17	0.34	25,709	—	—	—

Table 3
OLS Regressions (10–15-year-old age group)

	2+ Sample Boys	2+ Sample Girls	3+ Sample Boys	3+ Sample Girls	3+ Sample First-borns Boys	3+ Sample First-borns Girls
Labor force participation						
Family size	0.024***	0.011***	0.023***	0.009***	0.025***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Second child			0.000	0.003***	0.000	0.000
			(0.002)	(0.001)	(0.000)	(0.000)
N	148,106	139,214	145,156	138,689	92,182	87,556
Household chore activities						
Family size	0.001***	0.018***	0.002***	0.020***	0.002***	0.022***
	0.000	(0.001)	0.000	(0.001)	(0.001)	(0.001)
Second child			−0.002***	−0.002	0.000	0.000
			(0.002)	(0.002)	(0.000)	(0.000)
N	129,716	133,054	126,351	132,328	78,106	82,784
Attendance						
Family size	−0.033***	−0.026***	−0.036***	−0.029***	−0.037***	−0.031***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Second child			0.016***	0.005***	0.000	0.000
			(0.002)	(0.002)	(0.000)	(0.000)
N	148,106	139,214	145,156	138,689	92,182	87,556

School progression						
Family size	−0.036***	−0.036***	−0.037***	−0.040***	−0.037***	−0.039***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Second child			0.000	0.004***	0.000	0.000
			(0.001)	(0.001)	(0.000)	(0.000)
N	148,019	139,158	145,091	138,643	92,124	87,519
Literacy						
Family size	−0.029***	−0.020***	−0.037***	−0.029***	−0.033***	−0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Second child			0.018***	0.014***	0.000	0.000
			(0.002)	(0.002)	(0.000)	(0.000)
N	148,106	139,214	145,156	138,689	92,182	87,556

Robust standard errors in parenthesis. School Progression ≡ *education*/(*age*-6).
Sample: Children aged in ∈ [10,15] living with two adults (the mother and her husband). Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; child's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas.

Table 4
OLS Regressions (18–20-year-old age group)

	2+ Sample		3+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls
Labor Force Participation						
Family size	0.014***	0.009***	0.015***	0.006***	0.014***	0.006**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Second child			0.005	0.015**		
			(0.006)	(0.008)		
N	38,928	28,609	34,827	25,686	28,415	20,585
Household Chore Activities						
Family size	0.000	0.021***	0.000	0.024***	0.000	0.025***
	−0.001	(0.002)	−0.001	(0.003)	(0.002)	(0.003)
Second child			0.007	−0.007		
			(0.005)	(0.008)		
N	14,115	16,735	12,023	15,102	9,674	11,987
Attendance						
Family size	−0.019***	−0.019***	−0.020***	−0.018***	−0.019***	−0.019***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Second child			−0.004	0.004		
			(0.007)	(0.008)		
N	38,928	28,609	34,827	25,686	28,415	20,585
School Progression						
Family size	−0.023***	−0.025***	−0.012	−0.049***	−0.022***	−0.025***
	(0.001)	(0.001)	(0.010)	(0.012)	(0.001)	(0.001)
Second child			−0.011*	0.006		
			(0.006)	(0.006)		
N	34,765	25,634	34,765	25,634	28,366	20,545

Literacy						
Family size	−0.013***	−0.008***	−0.014***	−0.009***	−0.014***	−0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Second child			0.012***	0.012***		
			(0.004)	(0.003)		
N	38,928	28,609	34,827	25,686	28,415	20,585
Completed years of schooling						
Family size	−0.233***	−0.277***	−0.230***	−0.283***	−0.233***	−0.296***
	(0.010)	(0.015)	(0.010)	(0.015)	(0.011)	(0.018)
Second child			−0.112**	−0.061		
			(0.045)	(0.062)		
N	23,256	13,361	21,730	12,469	17,842	10,082
High school						
Family size	−0.008***	−0.017***	−0.005***	−0.015***	−0.006***	−0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Second child			−0.021***	−0.037***		
			(0.004)	(0.006)		
N	38,928	28,609	34,827	25,686	28,415	20,585
College attendance						
Family size	0.001**	−0.003***	0.003***	0.000	0.002***	−0.001
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Second child			−0.011***	−0.026***		
			(0.003)	(0.004)		
N	38,928	28,609	34,827	25,686	28,415	20,585

Robust standard errors in parenthesis. School Progression ≡ *education*/(*age*-6).
Sample: Children aged in ∈ [18,20] living with two adults (the mother and her husband).
Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; individual's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas.

Table 5
First-Stage Regressions (10–15-year-old age group)

	Dependent Variable: # of children in the family					
	2 + Sample		3 + Sample		3 + Sample (First-Borns)	
	Boys	Girls	Boys	Girls	Boys	Girls
Presence of twins	0.852***	0.862***	0.924***	0.986***	0.950***	0.989***
	(0.028)	(0.028)	(0.026)	(0.027)	(0.032)	(0.033)
Second child			0.351***	0.355***		
			(0.007)	(0.006)		
F-test	1070.76***	1137.98***	1123.65***	1188.20***	648.57***	691.94***
N	139,214	148,106	138,689	145,156	87,556	92,182

Robust standard error in parenthesis.
Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age.
Sample: Children aged in ∈ [10,15] living with two adults (the mother and her husband).
2 + Sample: 1st borns with at least 2 births in the families.
3 + Sample: 1st and 2nd borns with at least 3 births in the families.

Table 6
First-Stage Regressions (18–20-year-old age group)

Dependent Variable: # of children in the family						
	2+ Sample		3+ Sample		3+Sample (First-Borns)	
	Boys	Girls	Boys	Girls	Boys	Girls
Presence of twins	0.919***	0.935***	0.910***	0.950***	0.937***	1.018***
	(0.076)	(0.075)	(0.076)	(0.071)	(0.082)	(0.080)
Second child			0.429***	0.530***		
			(0.024)	(0.023)		
F-test	215.41***	332.48***	201.40***	305.45***	157.38***	241.50***
N	28,609	38,928	25,686	34,827	20,585	28,415

Robust standard error in parenthesis.
Year and state dummies as control variables.
Sample: Individuals aged in $\in$ [18,20] living with two adults (the mother and her husband).
2+ Sample: first borns with at least two births in the families.
3+ Sample: first and second borns with at least three births in the families.

Some observations are worth noting. First, the measure of the number of siblings is accurate in our samples because we exclude all families that have some children who are deceased or living outside the household. In other words, we include complete families only. Second, our measurement of the presence of twins may be overestimated because we identify twins by the age of the individuals. Therefore, it is possible that some nontwin siblings are misclassified as twins, which might generate measurement errors in our instrumental variable. However, as long as they are not correlated with the error in the second-stage equation and the measure of presence of twins is correlated to the actual variable, the second-stage coefficients are consistent even if they bias the first-stage coefficients.

The IV second-stage regressions for the 10–15-year-old age group displayed in Table 7 show the impacts of family size on child-quality outcomes for boys and girls separately. The results for labor-force participation show a positive and significant impact of family size in the three different samples for girls. For boys, only the coefficient in the 2+ sample is statistically significant. The IV point estimates imply that an extra sibling increases the likelihood of a girl (boy) joining the labor market by 1.4 (2.0) percentage points in the 2+ sample. This increase is sizeable. The child labor incidence among 10–15-year-old girls and boys is approximately 6 percent and 14 percent, respectively, in Brazil, which implies that an extra sibling increases the probability of a child working by 1.4/6 = 23 percent for girls and 2.0/14 = 14 percent for boys. The results for the 3+ samples are also positive and significant, with similar magnitudes for girls. However, they are not statistically significant for boys.[15]

The results for household chores are not statically significant, except for the values for girls in the 3+ firstborn sample. It is important to notice that the definition of household chores is narrow and consists of a full-time activity for those are who not participating in the labor market. [16] It is possible that individuals engage in household chores combined with other activities and that this is more likely to occur in larger families. However, we cannot capture this impact due to the definition of household chores used in the Census. Nevertheless, it is interesting to notice that there is a positive and significant impact of family size for first-born girls in the 3+ sample, which may suggest a process in the intra-household allocation decisions in which older females in larger families dedicate more time to taking care of younger siblings and/or household duties. The fact that we did not find this effect in the 2+ sample indicates that the impact may be nonlinear with respect to family size.

The results for school attendance suggest a detrimental effect of family size. The coefficient is statistically different from zero at a 5 percent level of significance for girls in the 3+ firstborn sample only. However, all of the coefficients are negative, and point estimates range from −1.3 to −2.3 percentage points. For the first-born girls in the 3+ sample, an extra sibling reduces the probability of attending school

15. We also investigate the impact of family size on work intensity. We observe the intensive margin by including weekly working hours as an additional outcome. We did not find any impact of family size on this margin. The results are available upon request.

16. For those who are not participating in the labor market, the Census classifies their time allocation among the following exclusive options: (i) doing chores, (ii) studying, (iii) job-searching or (iv) retired or not occupied.

by 2.3 percentage points, which corresponds to an increase of 2.3/14.81 = 15.5 percent for the probability of not attending school.

For school progression, the results indicate a detrimental effect of family size. The coefficients are negative and significant for boys and girls and for all samples, and they imply a considerable effect on educational attainment. For instance, the point estimate for girls in the 2+ sample is approximately 0.03. The average school progression is approximately 0.9. If a 15-year-old girl is in the correct grade, she must have nine years of schooling. The marginal impact of family size would decrease this figure to $9 \times 0.03 = 0.27$ years of schooling, that is, one-third of a year of schooling. Indeed, there is evidence that a school delay is associated with lower final educational attainments (see, for instance, Meisels and Liaw 1993 and Bedi and Marshall 2002).

Finally, for literacy, we find negative point estimates of family size. For girls, the results are significant in the 3+ samples only. For boys, family size is not significant in the 3+ firstborn sample only. For instance, boys in the 2+ sample are impacted by a decrease of two percentage points in the probability of being literate. This magnitude implies that an extra sibling increases the chances of being illiterate by 2/9 = 22 percent.

When comparing the IV with the OLS regressions, the OLS coefficients seem to overestimate (in absolute terms) the actual impact of family size on household chores, school attendance, school progression, and literacy outcomes for both boys and girls. However, in the case of labor-force participation, the OLS biases are different for boys and girls. For boys, the OLS bias is positive. The opposite occurs for girls, even though the point estimates of OLS and IV are similar. These comparisons suggest that families with a preference for child quantity care less about child quality or are more likely to be credit-constrained. The same bias direction for both boys and girls for most of the outcomes suggests that unobservable tastes, wealth, and ability are similarly correlated with girls' and boys' outcomes. The twins IV approach accounts for these biases.

B. Impacts on the 18–20-Year-Old Age Group

The tradeoff between the quantity and quality of children may cause lasting effects on individual human capital accumulation. Ideally, to measure these completed impacts, we want to observe the adult siblings after completion of their human capital formation. Although no such data are available in Brazil, we can observe young adults still living with their parents and siblings. Therefore, it is possible to build a sample of young adult siblings. The choice of the age range of the young adult sample involves a tradeoff. An older young adult is more likely to have completed the human capital formation but is less likely to still live with the parents and siblings. To maximize the number of individuals who still live with their parents and have completed their formal education process, we choose the 18–20-year-old age range. In fact, approximately 70 percent of the 18–20-year-old age-group individuals live in the same household as their parents. Moreover, of all 18–20-year-old individuals, 70 percent were not attending school in 1991. It is important to note that females generally move out of their parents' household earlier than males to

Table 7
Second-Stage Regressions (10–15-year-old age group)

	2+ Sample		3+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls
Labor force participation						
Family size	0.020**	0.014**	0.003	0.012*	−0.001	0.017*
	(0.009)	(0.007)	(0.009)	(0.007)	(0.011)	(0.009)
Second child			0.007**	0.002		
			(0.004)	(0.003)		
N	148,106	139,214	145,156	138,689	92,182	87,556
Household chore activities						
Family size	0.003	0.009	0.001	0.011	0.002	0.019*
	(0.004)	(0.008)	(0.004)	(0.008)	(0.005)	(0.010)
Second child			−0.002	0.001		
			(0.001)	(0.003)		
N	129,716	133,054	126,351	132,328	78,106	82,784
Attendance						
Family size	−0.013	−0.016*	−0.018*	−0.013	−0.016	−0.023**
	(0.009)	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)
Second child			0.009**	−0.001	0.000	0.000
			(0.004)	(0.004)	(0.000)	(0.000)
N	148,106	139,214	145,156	138,689	92,182	87,556

School Progression						
Family size	−0.026***	−0.029***	−0.026***	−0.025***	−0.029***	−0.024***
	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)
Second child			−0.004	−0.002		
			(0.003)	(0.003)		
N	148,019	139,158	145,091	138,643	92,124	87,519
Literacy						
Family size	−0.020**	−0.010	−0.016*	−0.020**	−0.013	−0.023**
	(0.008)	(0.007)	(0.009)	(0.008)	(0.010)	(0.009)
Second child			0.011***	0.011***		
			(0.004)	(0.003)		
N	148,106	139,214	145,156	138,689	92,182	87,556

Robust standard errors in parenthesis. School Progression ≡ education/(age-6). Instrumental variable: occurrence of twins in the 2nd (2 + Sample) or in the 3rd (3 + Sample) births. Sample: Children aged in ∈ [10,15] living with two adults (the mother and her husband). Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; child's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas.

Table 8
Second-Stage Regressions (18–20-year-old age group)

	2+ Sample		3+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls
Labor force participation						
Family size	−0.029	0.019	0.004	0.045*	0.003	0.047
	(0.021)	(0.027)	(0.021)	(0.027)	(0.022)	(0.030)
Second child			0.011	−0.001		
			(0.013)	(0.014)		
N	38,928	28,609	34,827	25,686	28,415	20,585
Household chore activities						
Family size	0.028	−0.008	0.000	0.086**	−0.004	0.108**
	(0.020)	(0.028)	(0.015)	(0.037)	(0.014)	(0.043)
Second child			0.007	−0.035*		
			(0.008)	(0.019)		
N	14,115	16,735	12,023	15,102	9,674	11,987
Attendance						
Family size	0.013	−0.016	−0.009	−0.060**	0.004	−0.060**
	(0.021)	(0.026)	(0.021)	(0.027)	(0.023)	(0.030)
Second child			−0.010	0.022		
			(0.013)	(0.014)		
N	38,928	28,609	34,827	25,686	28,415	20,585
School progression						
Family size	−0.005	−0.044***	−0.012	−0.049***	−0.017	−0.049***
	(0.010)	(0.011)	(0.010)	(0.012)	(0.011)	(0.013)
Second child			−0.011*	0.006		
			(0.006)	(0.006)		
N	38,861	28,562	34,765	25,634	28,366	20,545

Literacy						
Family size	0.006	−0.004	−0.002	−0.011	−0.012	−0.003
	(0.013)	(0.011)	(0.014)	(0.013)	(0.015)	(0.013)
Second child			0.005	0.012*		
			(0.008)	(0.006)		
N	38,928	28,609	34,827	25,686	28,415	20,585
Completed years of schooling						
Family size	−0.137	−0.372*	−0.060	−0.634***	−0.221	−0.598***
	(0.146)	(0.198)	(0.164)	(0.194)	(0.173)	(0.207)
Second child			−0.217*	0.089		
			(0.111)	(0.104)		
N	23,256	13,361	21,730	12,469	17,842	10,082
High school						
Family size	−0.014	−0.046**	−0.018	−0.056***	−0.020	−0.064***
	(0.014)	(0.020)	(0.012)	(0.019)	(0.012)	(0.021)
Second child			−0.014*	−0.020**		
			(0.007)	(0.010)		
N	38,928	28,609	34,827	25,686	28,415	20,585
College attendance						
Family size	0.003	−0.010	0.001	−0.030**	−0.003	−0.041***
	(0.010)	(0.015)	(0.008)	(0.014)	(0.008)	(0.015)
Second child			−0.011**	−0.013*		
			(0.005)	(0.007)		
N	38,928	28,609	34,827	25,686	28,415	20,585

Robust standard errors in parenthesis. School Progression ≡ education/(age-6).
Instrumental variable: occurrence of twins in the second (2 + Sample) or in the third (3 + Sample) births. Sample: Individuals aged in ∈ [18,20] living with two adults (the mother and her husband).
Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; individual's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas.

become spouses. Of all individuals who still live with their parents, 58 percent are males, and 42 percent are females.

Similar to the 10–15-year-old age group, we construct three different samples for the 18–20-year-old age group. Table 2 presents the descriptive statistics of the 2+, 3+ and 3+—firstborn samples for boys and girls separately. In the 2+ sample, 70 percent of all males and 46 percent of the females participate in the labor market. Of all females who are not participating in the labor market, roughly 24 percent report doing some chores in the 2+ sample. This figure is 4.5 percent for males. In total, 40 percent of the males and 53 percent of the females are attending school in the 2+ sample.[17] The average school progression indices are 0.57 and 0.68 for males and females in the 2+ sample, respectively. In total, 93 percent (96 percent) of the males (females) in the 2+ sample are literate. We construct three additional outcomes of this age group. The first one is the an indicator variable if the individual completed high school.[18] The second is an indicator variable if the individual has attended college. Finally, for those who are not attending school anymore, we obtain the completed years of schooling. The goal is to capture the completed human capital formation. For high school completion, 17 percent (28 percent) of the males (females) have completed high school in the 2+ samples. For the 3+ sample, the figures for high school completion are 13 percent and 22 percent for males and females, respectively. For college attendance, 8 percent (15 percent) and 13 percent (22 percent) of the males (females) attend college in the 2+ and 3+ samples, respectively. For those who are not attending school, the average completed years of schooling is 6.4 (7.4) in the 2+ sample. Of all families in the 2+ sample, 1.4 percent have twin births in the second birth. A similar number is obtained for the 3+ sample. The average family size is approximately 3.6 for the 2+ sample and 4.3 for the 3+ sample. Roughly 74 percent live in urban areas in the 2+ sample, and 70 percent live in urban areas in the 3+ sample.

Table 4 presents the OLS regression results for all three samples in the 18–20-year-old age group. The outcomes of child labor and education remain qualitatively similar to those found for the 10–15-year-old age group, except for household chores, which present statistically insignificant coefficients for males. Particularly important is that an extra sibling is associated with a lower probability of completing high school. This result holds for boys and girls for all samples. For college attendance, we find that family size is positively correlated with the probability of a male attending college. However, this correlation is negative for females in the 2+ sample. Finally, an extra sibling is associated with −0.23 (−0.28) completed years of schooling for males (females).

Tables 6 and 8 present the first-and second-stage results. The first-stage results depict a strong positive correlation between the presence of twins and family size for all samples. The second-stage regressions show no impact of family size on

17. These figures are different from the population averages because the sample is restricted to sons and daughters living with their parents. We observe that they are more likely to attend school compared to those who live outside the parents' household.

18. Notice that this the high school completion variable is related to the school progression variable, since the individuals that have not yet completed the high school are not attending or are delayed in completing high school.

labor-force participation and household chores for boys. For girls, they show a positive impact on these outcomes for the 3+ sample. Indeed, the regressions are statistically significant for household chores. The impact is considerable. For instance, the coefficient in the 3+ firstborn sample is 0.108, which implies an increase in the probability of doing chores by 0.108/0.265 = 40.7 percent.

The same pattern is observed for school attendance and progression, high school completion and college attendance. There are no impacts of family size for boys, but there are significant effects for girls. For school and college attendance, the results are significant for the 3+ sample, and the estimated coefficients are 0.06 and 0.04, respectively, for the firstborn females, which implies a decrease of 0.06/0.51 = 11.8 percent and of 0.04/0.12 = 30 percent, respectively, in the probability of attending school and college. There is a significant impact on school progression and high school completion for females in all three samples. For literacy, the results show no significant impact for males or females.

Finally, the results reveal a negative and significant effect of family size on the completed years of schooling for females. Although the value is negative, there is no significant impact for males. For females in the 3+ firstborn sample, the estimated coefficient is −0.598, showing that an extra sibling reduces the expected completed formal education by more than half a year.

Compared to the OLS results, the IV estimations generate higher point estimates for girls (in absolute values) and lower point estimates, in most cases, for boys in the young adults age group. One possible explanation is that there is a household specialization among siblings, in particular for credit-constrained families. An unexpected extra sibling forces them to gender-specialize even more. The results suggest that those families invest more in the human capital formation of the oldest male, while the females specialize in doing household chores to the detriment of schooling investments. The IV may be capturing this specialization forced by the unexpected child.

Taken together, these results indicate that the tradeoff between quantity and quality implies time-allocation choices that are detrimental for females, especially for those in larger families. In general, it seems that they have a lower probability maintaining the process of human capital accumulation, and those who have completed this process end up completing fewer years of schooling. Moreover, they are more likely to do household chores as their main time-allocation activity.

C. Mechanisms

There are two plausible conjectures about the mechanisms through which the family size impacts human capital formation: credit and time constraints. The credit-constraint channel operates through the dilution of income resources due to the birth of the extra sibling. The impossibility of bringing the family's future income to the present may reduce the amount of resources available to each family member. Therefore, to maintain current consumption, this constraint may force the family to underinvest in the human capital formation of the children and to push them to the labor market or chore activities earlier.

The time-constraint channel works through the dilution of time devoted by parents to each child after the birth of the extra sibling. The birth of a new sibling may restrict the time that the parents allocate to the raising of each child. This restriction may cause a reallocation of the time of the parents and of the older siblings to take care of the younger siblings. This restriction may impact the children's human capital formation in two ways. First, parents may dedicate less time to the care and education of each child. Second, the family may require the use of time of some children, especially the older ones, in activities that might be detrimental to their own human capital formation, such as labor market activities, taking care of the younger siblings, or other chores.

This subsection sheds some light on these possible channels. First, to investigate the existence of the credit-constraint channel, we divide the sample between families with low- and high-educated mothers. We define low-educated mothers as those with three or less years of schooling and high-educated mothers as those with 11 or more years of schooling. The assumption is that the mother's education is correlated with the family's wealth. If wealthy families are less credit-constrained, then the mother's education may be a good indicator of how likely a family is to be credit-constrained.[19]

The results are presented in Tables 9 and 10 for the ten to 15 and 18–20 age groups, respectively. We find evidence that suggests that the credit-constraint channel is, indeed, a relevant mechanism through which the quantity and quality tradeoff operates. We find that the detrimental effects on schooling are more pervasive among low-educated mother families.

Second, we analyze the presence of the time-constraint channel by exploring the differences in the exposure time of the firstborn sibling to the occurrence of an unexpected extra birth in the family. This mechanism may operate in different ways. For instance, consider the cases of two firstborn children that are both exposed to an extra sibling. In the first case, the child is ten years old, and in the second case, the child is two years old. On the one hand, the parents of the ten-year-old can devote more years to the sole care of the firstborn child compared to the parents of the two-year-old child. This can have different accumulative effects in the human capital formation of the children. On the other hand, the larger age gap may make the firstborn child more likely to engage in activities that steal time from her own human capital accumulation, such as taking care of the younger sibling and doing household chores. Therefore, the net effect of an extra sibling through this channel is ambiguous. Moreover, these effects may be nonlinear. The age gap between the two siblings might be sufficiently wide that the human capital accumulation process is already completed.

We investigate the presence of the time-constraint channel by dividing the sample into narrow and wide-birth spacing groups. The narrow-spacing group is formed by firstborn children living in families in which the difference between the first birth and the births of twins is equal to or less than two years. The wide-spacing group is formed by firstborn in families in which the difference between the first birth and the births of twins is greater than two years. The time-constraint channel is inferred by comparing the impacts of family size in the narrow and wide-spacing samples, controlling for the firstborn child's age. Therefore, the difference in spacing captures

19. In this subsection, we estimate the impacts for the 2+ sample and 3+ sample firstborns.

Table 9
Second-Stage Regressions—Credit-Constraint Mechanism (10 to 15)

	Low-educated mother				High-educated mother			
	2+ Sample		3+ Sample First-borns		2+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Labor force participation								
Family size	0.016	0.016**	0.010	0.004	−0.000	−0.008	−0.020	0.009
	(0.010)	(0.007)	(0.011)	(0.008)	(0.014)	(0.010)	(0.015)	(0.014)
N	144,855	135,033	103,486	96,719	30,101	28,546	13,921	13,365
Household chore activities								
Family size	0.004	0.012	−0.001	0.016	−0.004	0.013	0.017	−0.005
	(0.005)	(0.009)	(0.005)	(0.011)	(0.004)	(0.011)	(0.013)	(0.008)
N	114,863	125,111	79,615	88,853	29,082	28,260	13,299	13,193
Attendance								
Family size	−0.018*	−0.007	−0.017	−0.035***	−0.003	−0.029*	0.004	0.015**
	(0.010)	(0.010)	(0.012)	(0.012)	(0.012)	(0.016)	(0.013)	(0.008)
N	144,855	135,033	103,486	96,719	30,101	28,546	13,921	13,365
School progression								
Family size	−0.016***	−0.024***	−0.015**	−0.026***	−0.013	−0.027*	−0.017	0.003
	(0.006)	(0.006)	(0.006)	(0.007)	(0.014)	(0.016)	(0.016)	(0.016)
N	144,761	134,964	103,418	96,673	30,086	28,535	13,912	13,358
Literacy								
Family size	−0.009	−0.008	−0.005	−0.037***	−0.007	0.008***	−0.005	−0.001
	(0.009)	(0.008)	(0.010)	(0.010)	(0.008)	(0.001)	(0.011)	(0.007)
N	144,855	135,033	103,486	96,719	30,101	28,546	13,921	13,365

Robust standard errors in parenthesis. School Progression ≡ education/(age-6). Instrumental variable: occurrence of twins in the 2nd (2 + Sample) or in the 3rd (3 + Sample) births. Sample: Children aged in ∈ [10,15] living with two adults (the mother and her husband). Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; child's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas. Low-educated mothers: three or less years of schooling High-educated mothers: 11 or more years of schooling.

Table 10
Second-Stage Regressions—Credit-Constraint Mechanism (18 to 20)

	Low-educated mother				High-educated mother			
	2+ Sample		3+ Sample First-borns		2+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Labor force participation								
Family size	−0.020	0.053*	−0.000	0.043	0.064	−0.130*	0.141**	0.044
	(0.023)	(0.031)	(0.023)	(0.036)	(0.090)	(0.077)	(0.069)	(0.082)
N	27,340	18,957	21,091	14,430	5,486	4,805	3,255	2,917
Household chore activities								
Family size	0.013	−0.022	0.004	0.138**	0.051	0.086	−0.048***	0.048
	(0.022)	(0.034)	(0.016)	(0.059)	(0.059)	(0.072)	(0.011)	(0.067)
N	8,049	10,595	6,006	8,143	3,436	3,397	1,985	2,026
Attendance								
Family size	0.007	−0.020	0.014	−0.048	0.006	0.065	−0.076	−0.048
	(0.024)	(0.030)	(0.024)	(0.035)	(0.084)	(0.075)	(0.066)	(0.078)
N	27,340	18,957	21,091	14,430	5,486	4,805	3,255	2,917
School progression								
Family size	0.010	−0.036***	−0.008	−0.052***	0.031	−0.042	−0.029	0.016
	(0.012)	(0.013)	(0.011)	(0.017)	(0.024)	(0.030)	(0.032)	(0.017)
N	27,298	18,919	21,063	14,397	5,481	4,800	3,252	2,913

Literacy								
Family size	0.020	−0.003	−0.001	0.000	0.008**	0.002*	0.005*	0.003**
	(0.016)	(0.015)	(0.018)	(0.018)	(0.003)	(0.001)	(0.003)	(0.001)
N	27,340	18,957	21,091	14,430	5,486	4,805	3,255	2,917
Completed years of schooling								
Family size	−0.042	−0.306	−0.073	−0.550**	0.889**	−0.298	−0.269	0.028
	(0.155)	(0.220)	(0.181)	(0.257)	(0.387)	(0.639)	(0.524)	(0.124)
N	18,563	10,244	14,748	8,034	1,797	1,309	1,103	823
High school								
Family size	−0.011	−0.045**	−0.018*	−0.079***	−0.022	0.001	−0.051	0.012
	(0.013)	(0.020)	(0.011)	(0.020)	(0.086)	(0.078)	(0.071)	(0.076)
N	27,340	18,957	21,091	14,430	5,486	4,805	3,255	2,917
College attendance								
Family size	−0.003	−0.014	−0.000	−0.030***	−0.042	0.032	−0.065	−0.094
	(0.007)	(0.012)	(0.006)	(0.009)	(0.077)	(0.075)	(0.050)	(0.076)
N	27,340	18,957	21,091	14,430	5,486	4,805	3,255	2,917

Robust standard errors in parenthesis. School Progression ≡ education/(age-6). Instrumental variable: occurrence of twins in the 2nd (2 + Sample) or in the 3rd (3 + Sample) births. Sample: Individuals aged in ∈ [18,20] living with two adults (the mother and her husband). Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; individual's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas. Low-educated mothers: three or fewer years of schooling. High-educated mothers: 11 or more years of schooling.

the difference in the age of exposure to the increase in the family size between both samples. The results are presented in Tables 11 and 12. Overall, we find an increase in the family size tends to affect more firstborn children in the wide-spacing group. Family size is detrimental to school progression for the children aged 10–15 years in wide-spacing families. For the 18–20 age group, the family size negatively impacts most of the educational outcomes of firstborn girls in wide-spacing families in the 3+ sample. Moreover, there is also a positive impact on the likelihood of doing household chores for firstborn girls in the 3+ sample.

These results suggest that both mechanisms (credit and time constraints) may be acting to explain the impacts of family size in child quality outcomes. In the presence of gender specialization, the credit- and time-constraint channels seem to negatively impact the educational attainment of females that stay at home engaged in household chores. Moreover, because wide-spacing females are more affected by an extra sibling, it seems that the time-constraint channel occurs through the time reallocation of the older sibling.

VI. Conclusion

In this paper, we estimate the causal impact of family size on new dimensions of child quality. Specifically, we gauge the effect of an extra sibling on child labor, an often-neglected outcome that is closely related to children's well-being. We use two distinct indicators: labor-force participation and household chores. We also investigate the effects on more traditional educational outcomes, such as school attendance, school progression, and literacy outcomes. We observe two age groups. The first outcome encompasses children aged between ten and 15 years. To check possible lasting effects of family size, we also analyze the impacts on young adults aged 18–20 years. For the second age group, we additionally estimate the impact on high school completion, college attendance and the completed years of schooling for those who have completed their formal human capital accumulation.

The main empirical problem in measuring such an effect is the potential endogeneity of fertility, as it is a choice variable, and unobservables might influence both family-size choice and child-quality outcomes. To overcome this problem, we use the instrumental variable estimation approach. We use the presence of twins as the instrument for family size. We show that the presence of twins is strongly correlated with the number of children. Because the birth of twins is very likely to be a random event and orthogonal to the unobservables that may affect children and family characteristics, we believe that the presence of twins has the required properties to be a good instrumental variable.

A simple OLS approach shows a strong detrimental relationship between family size and child-quality outcomes. The IV estimators reveal that the OLS coefficients are upward-biased. They also show that the exogenous increase in family size is positively related to labor force participation for boys and girls and to household chores for young women. Moreover, we find negative effects on educational outcomes for boys and girls and negative impacts on human capital formation for young female adults. Indeed, the results show that young women suffer most of the detrimental impact of an extra sibling, particularly in larger families. They are more

Table 11
Second-Stage Regressions—Time-Constraint Mechanism (10–15-years-old)

	Narrow spacing				Wide spacing			
	2+ Sample		3+ Sample First-borns		2+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Labor force participation								
Family size	0.040***	0.025**	−0.078	0.032	0.008	0.007	0.002	0.019
	(0.014)	(0.012)	(0.049)	(0.052)	(0.011)	(0.007)	(0.016)	(0.014)
N	70,003	65,976	4,731	4,598	78,103	73,238	36,760	35,017
Household chore activities								
Family size	−0.000	0.015	0.002	−0.016	0.004	0.007	0.000	0.012
	(0.006)	(0.013)	(0.020)	(0.053)	(0.005)	(0.008)	(0.007)	(0.016)
N	60,073	62,527	3,779	4,261	69,643	70,527	31,594	33,333
Attendance								
Family size	−0.024*	−0.021	−0.046	0.012	−0.006	−0.012	−0.030*	0.002
	(0.015)	(0.015)	(0.067)	(0.063)	(0.011)	(0.010)	(0.018)	(0.018)
N	70,003	65,976	4,731	4,598	78,103	73,238	36,760	35,017
School progression								
Family size	−0.016	−0.022**	−0.048	0.027	−0.032***	−0.031***	−0.029***	−0.028**
	(0.010)	(0.011)	(0.038)	(0.036)	(0.007)	(0.007)	(0.011)	(0.011)
N	69,964	65,952	4,728	4,593	78,055	73,206	36,736	35,002
Literacy								
Family size	−0.034**	−0.007	−0.046	−0.002	−0.009	−0.010	−0.016	−0.018
	(0.013)	(0.012)	(0.053)	(0.052)	(0.009)	(0.008)	(0.015)	(0.015)
N	70,003	65,976	4,731	4,598	78,103	73,238	36,760	35,017

Robust standard errors in parenthesis. School Progression ≡ education/(age-6). Instrumental variable: occurrence of twins in the second (2 + Sample) or in the third (3 + Sample) births. Sample: Children aged in ∈ [10,15] living with two adults (the mother and her husband)Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; child's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas. Narrow spacing: difference between the first birth and the births of twins is equal or less than two years. Wide spacing: difference between the first birth and the births of twins is greater than two years.

Table 12
Second-Stage Regressions—Time-Constraint Mechanism (18–20-year-olds)

	Narrow spacing				Wide spacing			
	2+ Sample		3+ Sample First-borns		2+ Sample		3+ Sample First-borns	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Labor force participation								
Family size	0.013	0.014	−0.168	0.016	−0.044*	0.025	0.005	0.045
	(0.030)	(0.039)	(0.217)	(0.130)	(0.025)	(0.031)	(0.022)	(0.032)
N	18,868	13,657	1,346	1,000	20,060	14,952	27,069	19,585
Household chore activities								
Family size	0.023	−0.041	−0.022	0.610	0.028	0.019	−0.001	0.092**
	(0.031)	(0.044)	(0.017)	(0.586)	(0.022)	(0.031)	(0.015)	(0.046)
N	6,369	7,854	415	568	7,746	8,881	9,259	11,419
Attendance								
Family size	0.015	−0.036	0.121	−0.259*	0.007	−0.006	0.001	−0.048
	(0.034)	(0.038)	(0.261)	(0.156)	(0.024)	(0.030)	(0.023)	(0.031)
N	18,868	13,657	1,346	1,000	20,060	14,952	27,069	19,585
School progression								
Family size	−0.006	−0.057***	0.066	−0.172*	−0.004	−0.028**	−0.023	−0.038**
	(0.017)	(0.016)	(0.130)	(0.096)	(0.012)	(0.013)	(0.014)	(0.018)
N	18,836	13,627	1,345	998	20,025	14,935	12,240	9,004

Literacy								
Family size	0.007	−0.012	−0.112	−0.121	0.008	0.004	0.019	−0.001
	(0.020)	(0.017)	(0.174)	(0.116)	(0.014)	(0.012)	(0.020)	(0.019)
N	18,868	13,657	1,346	1,000	20,060	14,952	12,259	9,016
Completed years of school								
Family size	0.062	−0.590**	−0.931	−1.333	−0.253	−0.120	−0.198	−0.574***
	(0.237)	(0.259)	(1.966)	(0.947)	(0.163)	(0.245)	(0.171)	(0.219)
N	11,771	6,723	914	540	11,485	6,638	16,928	9,542
High school								
Family size	−0.029	−0.059**	−0.160	−0.024	−0.004	−0.029	−0.017	−0.067***
	(0.021)	(0.029)	(0.128)	(0.077)	(0.016)	(0.023)	(0.013)	(0.021)
N	18,868	13,657	1,346	1,000	20,060	14,952	27,069	19,585
College attendance								
Family size	−0.007	−0.019	−0.054	−0.049	0.009	−0.002	−0.002	−0.041***
	(0.014)	(0.021)	(0.045)	(0.081)	(0.012)	(0.018)	(0.009)	(0.015)
N	18,868	13,657	1,346	1,000	20,060	14,952	27,069	19,585

Robust standard errors in parenthesis. School Progression ≡ education/(age-6). Instrumental variable: occurrence of twins in the 2nd (2 + Sample) or in the 3rd (3 + Sample) births.Sample: Individuals aged in ∈ [18,20] living with two adults (the mother and her husband). Control variables: Year and state dummies; head's schooling, gender and age; mother's schooling and age; individual's age, squared age and gender; dummy variable if the family lives in urban areas and if the family lives in metropolitan areas. Narrow spacing: difference between the first birth and the births of twins is equal or less than two years. Wide spacing: difference between the first birth and the births of twins is greater than two years.

likely to work domestically, less likely to attend school, more likely to lag further behind in school, and more likely to complete fewer years of schooling.

Unlike studies that focus on developed countries, our study shows detrimental effects on educational outcomes for both boys and girls as well as young females. The difference in the findings may be related to developing-country contexts, where credit rationing is more pervasive. Therefore, the quantity-quality tradeoff is measurable. An extra child may impose a larger resource dilution without the possibility of consumption-smoothing over time in a family. In this situation, fewer resources may be allocated to other siblings. In particular, the time endowment of young women may be reallocated to domestic work rather than to their own human capital accumulation. Indeed, we find evidence that suggests that both credit and time constraints are important mechanisms of the family-size effects on child quality.

These results can help to improve the design of anti-poverty programs in developing countries. For instance, conditional cash transfer programs are widespread across developing countries. Such a program usually consists of cash transfers that are conditional upon school attendance and/or visits to health care centers. The value of the transfers generally depends on the number of children in the family. A family with more children is eligible to receive greater cash transfers. If our results are correct, then the positive effect of the program on child-quality outcomes can be offset by the potential incentive of higher fertility.

Finally, the results suggest that there is gender specialization inside the family when a newborn arrives. Older female children in larger families seem to dedicate more time to taking care of younger siblings and/or to household duties, thereby jeopardizing their long-term human capital accumulation when the family size increases.

Appendix 1

Summary Statistics

Table A1
Summary Statistics—Excluded and Nonexcluded Families

	2+ Sample				3+ Sample			
	Excluded		Nonexcluded		Excluded		Nonexcluded	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Panel A: 10–15-years-old								
Father's age	41.24	8.43	38.68	6.18	41.11	8.07	38.79	6.06
Mother's age	36.49	6.43	34.63	4.89	36.10	5.92	34.47	4.62
Father's education	2.82	3.30	5.35	4.39	2.28	2.88	4.57	4.16
Mother's education	2.94	3.18	5.42	4.09	2.46	2.81	4.68	3.88
Young adult age	12.77	1.70	12.34	1.70	12.38	1.65	12.11	1.62
White young adult	0.41	0.49	0.61	0.49	0.35	0.48	0.54	0.50
Urban area	0.58	0.49	0.74	0.44	0.52	0.50	0.68	0.47
Metro area	0.18	0.38	0.27	0.44	0.14	0.35	0.23	0.42
Panel B: 18–20-years-old								
Father's age	41.24	8.43	38.68	6.18	41.11	8.07	38.79	6.06
Mother's age	36.49	6.43	34.63	4.89	36.10	5.92	34.47	4.62
Father's education	2.82	3.30	5.35	4.39	2.28	2.88	4.57	4.16
Mother's education	2.94	3.18	5.42	4.09	2.46	2.81	4.68	3.88
Young adult age	12.77	1.70	12.34	1.70	12.38	1.65	12.11	1.62
White young adult	0.41	0.49	0.61	0.49	0.35	0.48	0.54	0.50
Urban area	0.58	0.49	0.74	0.44	0.52	0.50	0.68	0.47
Metro area	0.18	0.38	0.27	0.44	0.14	0.35	0.23	0.42

Table A2
Summary Statistics—Families with and without Twins

	2+ Sample				3+ Sample			
	No Twins		With Twins		No Twins		With Twins	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Panel A: 10–15-years-old								
Father's age	38.68	6.17	38.75	6.47	38.78	6.05	39.16	6.83
Mother's age	34.63	4.89	34.81	5.13	34.47	4.62	34.45	4.79
Father's education	5.36	4.39	4.70	4.23	4.58	4.16	3.77	3.89
Mother's education	5.43	4.09	4.70	3.96	4.69	3.88	3.92	3.61
Child age	12.34	1.70	12.40	1.69	12.11	1.62	12.18	1.61
White child	0.61	0.49	0.54	0.50	0.54	0.50	0.45	0.50
Urban area	0.74	0.44	0.68	0.47	0.68	0.47	0.61	0.49
Metro area	0.27	0.44	0.25	0.44	0.23	0.42	0.18	0.39
Panel B: 18–20-years-old								
Father's age	45.62	6.13	45.77	6.29	45.40	5.95	45.27	5.68
Mother's age	41.42	4.78	41.71	5.12	41.03	4.47	40.73	4.76
Father's education	5.03	4.39	4.23	4.29	4.50	4.19	3.95	4.00
Mother's education	4.82	4.01	4.05	4.01	4.34	3.83	3.72	3.79
Young adult age	18.87	0.81	18.90	0.81	18.76	0.79	18.71	0.78
White young adult	0.64	0.48	0.56	0.50	0.59	0.49	0.51	0.50
Urban area	0.76	0.43	0.68	0.47	0.72	0.45	0.67	0.47
Metro area	0.31	0.46	0.27	0.45	0.28	0.45	0.24	0.43

References

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

Angrist, Joshua, Victor Lavy and Analia Schlosser. 2005. "New Evidence on the Causal Link Between the Quantity and Quality of Children." NBER Working Papers 11835, Cambridge: National Bureau of Economic Research.

———. 2006. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children." Technical Report, MIT Department of Economics Working Paper Series, Cambridge: Massachusetts Institute of Technology.

Baland, Jean-Marie and James A. Robinson. 2000. "Is Child Labor Inefficient*?" Journal of Political Economy* 108(4):663–79.

Basu, Kaushik. 1999. "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards." *Journal of Economic Literature* 37(3):1083–1119.

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

Becker, Gary and Nigel Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy* 84(4):S143–62.

Bedi, Arjun and Jeffery Marshall. 2002. "Primary School Attendance in Honduras." *Journal of Development Economics* 69(1):129–53.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti. 2005. "Why Should We Care about Child Labor? The Education, Labor Market, and Health Consequences of Child Labor." Policy Research Working Paper Series 3479, Washington: The World Bank.

———. 2006. "Child Labor and Agricultural Shocks." *Journal of Development Economics* 81(1):80–96.

Behrman, Jere and Mark Rosenzweig. 2004 "Returns to Birthweight." *The Review of Economics and Statistics* 86(2):586–601.

Black, Sandra, Paul Devereux and Kjell 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.

Blau, Francine and Adam Grossberg. 1992. "Maternal Labor Supply and Children's Cognitive Development." *The Review of Economics and Statistics* 74(3):474–81.

Borlot, Ana Maria Monteiro and Zeidi Araújo Trindade. 2004. "As Tecnologias de Reprodução Assistida e as Representações Sociais de Filho Biológico." *Estudos de Psicologia* 9(1):63–70.

Caceres-Delpiano, Julio. 2006 "The Impacts of Family Size on Investment in Child Quality." *Journal of Human Resources* XLI(4):738–54.

Cigno, Alessandor and Camillo Rosati. 2002. "Child Labour Education and Nutrition in Rural India." *Pacific Economic Review* 7:1–19.

———. 2005. "The Economics of Child Labor." Oxford: Oxford University Press.

Conley, Dalton and Rebecca Glauber. 2005. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variations in Fertility." NBER Working Papers 11302, Cambridge: National Bureau of Economic Research.

Deb, Partha and Furio Rosati. 2004. "Estimating the Effect of Fertility Decisions on Child Labour." UCW Working Paper. Rome: Understanding Children's Work Project.

Edmonds, Eric. 2008. "Child Labor." In *Handbook of Development Economics*, eds. Theodore Paul Schultz and John Strauss, 3607–3709. Amsterdam: North-Holland.

Edmonds, Eric and Nina Pavcnik. 2005. "Child Labor in the Global Economy." *Journal of Economic Perspectives* 19(1):199–220.

———. 2006. “International Trade and Child Labor: Cross-country Evidence.” *Journal of International Economics* 68(1):115–40.

Emerson, Patrick and Andre Portela Souza. 2011. “Is Child Labor Harmful? The Impact of Working Earlier in Life on Adult Earnings.” *Economic Development and Cultural Change*. Forthcoming.

Goux, Dominique and Eric Maurin. 2005. “The Effect of Overcrowded Housing on Children’s Performance at School.” *Journal of Public Economics* 89(5–6):797–819.

Haan, Monique De. 2005. “Birth Order, Family Size and Educational Attainment.” Tinbergen Institute Discussion Papers 05-116/3. Tinbergen: Tinbergen Institute.

Hanushek, Eric A. 1992. “The Tradeoff between Child Quantity and Quality.” *Journal of Political Economy* 100(1):84–117.

Lee, Jungmin. 2004. “Sibling Size and Investment in Children’s Education: An Asian Instrument.” IZA Discussion Paper 1323. Bonn: Institute for the Study of Labor.

Li, Hongbin, Junsen Zhang and Yi Zhu. 2007. “The Quantity-Quality Tradeoff of Children in a Developing Country: Identification Using Chinese Twins.” IZA Discussion Paper 3012. Bonn: Institute for the Study of Labor.

Meisels, Samuel and Fong-Ruey Liaw. 1993. “Failure in Grade: Do Retained Students Catch Up?” *The Journal of Educational Research* 87(2):69–77.

Orazem, Peter and Chanyoung Lee. 2010. “Child Labor and the Transition Between School and Work.” *Research in Labor Economics*, 31: pp.iii.

Psacharopoulos, George and Harry Anthony Patrinos. 1997. “Family Size, Schooling and Child Labor in Peru—An Empirical Analysis.” *Journal of Population Economics* 10(4):387–405.

Qian, Nancy. 2009. “Quantity-Quality and the One Child Policy: The Only-Child Disadvantage on School Enrollment in Rural China.” NBER Working Papers 14973, Cambridge: National Bureau of Economic Research.

Rosenzweig, Mark and Kenneth Wolpin. 1980. “Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment.” *Econometrica* 48(1):227–40.