Munich Personal RePEc Archive

Testing the Children Quantity-Quality Trade-Off in India

Kumar, Santosh and Kugler, Adriana

Harvard University

1 September 2011

Online at https://mpra.ub.uni-muenchen.de/42487/
MPRA Paper No. 42487, posted 07 Nov 2012 15:47 UTC

Testing the Quantity-Quality Trade-Off in India

Adriana Kugler* Santosh Kumar†

Draft: Sep 2011

Abstract

Using data from a representative sample from India, we test the empirical validity of Quantity-Quality tarde-off model of Becker and Lewis (1973). To address the endogeneity arising from the joint determination of quantity and quality of children by parents, we instrument the family size by sex of the first child. We find a negative relationship between family size and children's educational attainment, even after controlling for parent's characteristics and birth order of children. The effects are heterogeneous. The trade-off is more pronounced in rural areas, for low-caste children, for illiterate mothers, and for children belonging to low wealth category. Overall, the findings support the quantity-quality trade-off in a resource poor setting such as India. Given that for long-run economic development, the quality of human capital is equally important, policymakers should invest more in education and other welfare programs in order to mitigate the adverse impacts of the trade-off.

JEL classification: N75, O18, O20.
Keywords: Family size, Education, India.

*Professor, Georgetown Public Policy Institute, Georgetown University, Washington DC. (email: ak659@georgetown.edu)
†Lecturer, University of Washington, Seattle, WA. (e-mail: skumar3@uw.edu). Please do not cite without author's prior permission.

1 Introduction

Policymakers in developing countries lay much emphasis on family planning programmes to reduce the family size in order to speed up the economic development. The policy intervention is based on the idea that a resource-constrained household with smaller family size will have more resources for investments in human capital of their children. In economics, this notion has been first modeled by Becker and Lewis (1973), which suggests that decrease in the quantity of children in a resource-constrained population will free up more resources to be invested on each child, leading to an increase in the average quality of child, quality often being measured by education or health status of children. In this paper, we test the empirical validity of Q-Q model in India, instrumenting family size by sex of the first child.

Testing the Q-Q model in Indian context is interesting and is of tremendous policy relevance. Today, over 35% of Indian population is below the age of 20. By 2020, it is expected that 325 million people in India will reach working age, which will be the largest in the world. This will come at a time when the rest of the developed world will be faced with an ageing population. It is estimated that by 2020, US will be short of 17 million people of working age, China by 10 million, Japan by 9 million and Russia by 6 million. At the same time, India will have a surplus of 47 million working people.[1]

The question that policymakers in India are facing today is whether this young Indian population is a bane or a boon. Economic growth require not just a large working population (quantity), but people with high human capital, skilled enough to enter the labor market. If the quality of working population is not high enough to complement the the sheer size of the working age population, the projected growth path may not be realized.[2] The working population size is a necessary but not sufficient condition for eco-

[1]Even when compared to developing countries, Brazil's working population is set to grow by 12%, ChinaŠs by 1%, Russia's will decline by 18%, while India will grow by 30%.

[2]India's adult illiteracy levels are a big concern, which stands at 39%. 25 million children are out of school in India, out of a total of 100 million out of school children in the world.

nomic growth. Quality of population (such as education and health) is equally important to make them productive when they join the labor force. The population growth could "transform into a demographic dividend" if every child is born healthy and is educated. To test this in the Becker and Lewis's Q-Q framework, we embark upon to examine the *causal* relationship between family size (quantity) and education attainment of children (quality) in India.

However, testing the existence of Q-Q trade-off empirically is challenging because child quantity and quality are endogenous variables, thereby confounding the causal interpretation. Fertility decisions and investment on child are jointly determined by parents (Browning, 1992; Haveman and Wolfe, 1995), which means that they are both affected by unobservable heterogeneity in parental preferences and household characteristics. Parents who are more concerned about the quality of their children may choose to have fewer children to educate each better.

The key method of addressing this endogeneity has been to take advantage of exogenous variation in family size that are either caused due to some policy experiment (one-child policy in China, forced sterilization in India) or due to natural occurrences (such as twins birth). The study by Rosenzweig and Wolpin (1980) is one of the early studies that exploits the birth of twins to isolate the causal effect of family size on child quality. The study found an inverse relationship between family size and children's educational attainment in a small sample of 25 twins in 1,600 children in India. Qian (2009) exploits the exogenous variation in the family size due to relaxation of one-child policy norm in China, and surprisingly found a positive correlation between number of children and school enrollment. She argues that a positive correlation is plausible if there are economies of scale in raising the children.

In addition to twins and one-child policy, another plausible instrument for family size can be sex of the first child or the gender composition of the first two children. The former instrument is based on the prevailing preference for sons that is observed in Asian countries, and the idea behind the latter instrument is that parents of same-gender siblings are more likely to go on to have an additional child. Using sex of the first child Lee (2008) finds negative impacts of family size on per-child investment in education for South Korean households. While Angrist el al. (2005) that used both twin births and gender composition as the instruments found no evidence for a quantity-quality trade-off in Israel. Similarly, Millimet and Wang (2010) fails to find statistically meaningful evidence of the quantity-quality trade over the whole distribution of the sample in Indonesia. The latter used gender composition of first two children as an instrument and examined the effect of number of children on child health outcomes (BMI and height).

Following this strand of literature, this paper uses the sex of the first child to instrument the family size as measured by number of children. We believe that IV estimation method will yield consistent results under the reasonable assumption that this instrument is a random event. The preference for sons in East and South East Asian countries are widely documented. The son preferences are in countries like India, Chin, and Korea is deeply rooted in social, cultural, and economic factors. Preference for sons can alter the family configuration by affecting fertility. This is because the couple who have all the sons they want will stop bearing additional child, while couples who are hoping to for a son will continue having children. Therefore, in a son-biased society, the first child's sex should be a good predictor for the probability of having a second child or the total number of children (Lee 2008).

The empirical evidence on the validity of Q-Q model is mixed at least in industrialized countries. Black et al. (2005) find no impact of family size on individual educational achievement in Norway once birth order is controlled. Similarly, Haan (2005) finds no significant effect of the number of children on educational attainment of the oldest child in US and the Netherlands. Angrist et al. (2005, 2010) do not find any causal impact of family size on completed educational achievement and earnings in Israel. Caceres-Delpiano (2006) finds a negative impact of family size on the likelihood that older children attend private school, but he finds no impact of an additional child on education attainment in US. In contrast, 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. Glick et al. (2007) use the natural experiment of twins at first birth to estimate the effects of unplanned fertility on the nutritional status and school enrolment of children in Romania. They find that a first-birth twins shock has negative impacts on childrenŠs human capital investments, particularly for later-born siblings. Additionally, Goux and Maurin (2005) show that children living in larger families perform worse in school than children in smaller families in France.[3]

The validity check of Q-Q model also gathered momentum in developing countries, since resource constraint argument inherent in Q-Q model seems more logical in a resource poor setting. The earliest study can be traced back to 1980, when 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. Recently a number of studies (Li et al., 2008; Qian, 2009; Rosenzweig and Zhang, 2009) have found mixed evidence of the Q-Q trade-off in China. Qian (2008) found positive effect, while Li et al., (2008) found some evidence of a Q-Q trade-off in rural areas of China, however the results were weaker in urban areas. Rosenzweig and Zhang (2009) indicate that an extra child significantly decreases the schooling progress, the expected college enrolment, grades in school and the assessed health of all children in the family.

Lee (2008) finds negative impacts of family size on per-child investment in education for South Korean households. Using twinning Ponczek and Souzay (2011) also finds negative effects on educational outcomes for Brazilian children, while Agueroy and Marksz (2008) find evidence of Q-Q trade-off for health indicators but not for education.[4]. Similarly,

[3]Using marital fecundability- as measured by the time interval from the marriage to the first birth- as a source of exogenous variation in family size, Klemp and Weisdorf (2011) documents a large and significantly negative effect of family size on children's literacy.

[4]The study uses data from Demographic and Health Surveys in Latin America, namely, Bolivia (conducted in 1994 and 1998), Brazil (1996), Colombia (1995 and 2000), the Dominican Republic (1996), Guatemala (1998), Nicaragua (1998), and Peru (1996), and instrument the family size by a mother's infertility status.

Razak et al. (2010) found no evidence in support of Q-Q trade-off in Malaysia.

Using a representative sample from India, we find evidence in support of the Q-Q trade-off. Our results show a negative effect of larger family size on educational attainment of children. In the 2SLS estimation, an extra child in the family reduces the probability of primary school completion by 5 percentage points and years of schooling by 0.36 years. Moreover, we find that the negative effects on educational outcomes are more pronounced for children in rural and low-caste households. We also find the trade-off is more severe for illiterate mother and low-wealth households.

Our paper adds to the existing literature on quantity-quality trade-off in an important ways. Most importantly, compared to most of the previous studies, this paper looks at a poor country, where extent of trade-off can be acute due to scarcity of resources.[5] We are the first to test the validity of Q-Q trade-off in India where almost one-sixth of world's population reside.[6] It is quite likely that the extent and nature of trade-off among Indian households are different compared to developed countries. Indian households face a quantity-quality trade-off due to financial constraint and lack of credit.

It is possible that the quantity and quality trade-off 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. 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.

The paper is organized as follows. The next section discuss the instrument and its validity. Section 3 explains the empirical framework followed by Section 4 that introduces the District Level Health Surveys (DLHS) and discuss the data used in this paper. The

[5] Qian (2009), Rosenzweig and Zhang (2009) and Li, Zhang, and Zhu (2008) focus on China while Lee (2008) focuses on South Korea.

[6] Incidently, the first paper on Q-Q in a developing country is on India by Rosenzweig and Wolpin (1980), however due to very small sample size, results can be assumed to be true in other parts of India given the heterogeneity in Indian culture and institutions.

main results are presented in section 5, while section 6 presents the heterogeneity in the results. Finally, Section 7 closes the paper summarizing the findings and outlining the agenda for future research and policy considerations.

2 Exogeneity of gender of first birth

Is the gender of first-birth a valid instrument for family size? The question of whether the exclusion restriction is satisfied translates into a question of whether gender of first birth (henceforth FB) is indeed exogenous, and whether it is correlated with any omitted variables which would affect educational outcomes of children. In this section, we discuss the validity of sex of first child as the instrument. The instrument, sex of the first child, is a valid instrument if it satisfies the following two conditions:

$$Relevance : Corr(Z_i, X_i) \neq 0 \tag{1}$$

$$Exogeneity : Corr(X_i, Z_i) = 0 \tag{2}$$

Condition (1) implies that the instrument, sex of first child, should be highly correlated with the endogeneous variable, family size, and condition (2) implies that the instrument should not affect the child outcomes except through family size. In a country like India where son preferences persist, the sex of first child is an important source of exogenous variation in fertility. The sex of the first birth has been used in the previous research by Lee (2008). However, the existence of sex-selective abortion may undermine the validity of the instrument because the access to ultrasound use and abortion services allows parents to choose the sex of their children. However, this does not seem to be a big concern given that ultrasound technology is not widely available in rural areas and due to the passing of Pre-natal Diagnostic Technique (PNDT) Act made the fetal-sex determination illegal. Retherford and Roy (2003) using the first two rounds of the National Family and Health

Survey finds little or no evidence of sex selection on the first birth.

In the absence of any interventions the probability of having a son is approximately 0.512 and this probability is independent of genetic factors (Ben-Porath and Welch 1976; Jacobsen, Møller and Mouritsen 1999).

3 Empirical Framework

Using sex of the first birth as an instrument we investigate the effect of number of children on their educational outcomes. We employ OLS and 2SLS regression analyses on the sample of 311,942 children as described in Table 1 above. Formally, we estimate the following equation:

$$Y_i = \beta_0 + \beta_1 FAMILYSIZE_i + \beta_2 X_i + \epsilon_i \tag{3}$$

Where, Y_i is the educational attainment of the child as measured by primary school completion and years of schooling. The variable FAMILYSIZE is the number of children in the family; X is a vector of covariates, and ϵ is an error term. Covariates comprise of child characteristics, including age, gender, ethnic group, birth order and place of residence. In addition to these child-level characteristics, we also include a set of parental attributes, including age and education level of children's father and mother. We use the same covariates in both the OLS and the 2SLS regression analyses. The main coefficient of interest is $beta_1$ that will provide the evidence on Q-Q trade-off.

The coefficient β_1 as estimated by the ordinary least squares (OLS) method may be biased and may not have causal interpretation (Angrist et al. 2010). The OLS estimates will be either downward or upward biased depending on nature of endogeneity either due to reverse causality or omitted variable. For example, when education is negatively correlated with fertility, and fertility in turn is negatively correlated with education, then the OLS estimate of β_1 will be downward biased. If education and fertility, on the other hand, are both positively correlated with an omitted variable, then the estimate of β_1 will be upward biased. Following the existing literature we attempt to tackle these potential problems of endogeneity by use of 2SLS analysis. In the first stage, family size is predicted by using the sex of the FB, as well as covariates. The second stage then predicts educational outcomes using the equation described above.

The first stage of the two-stage least squares (2SLS) estimation is given by

$$FAMILYSIZE = \alpha_0 + \alpha_1 SEXFIRSTBORN_i + \alpha_2 X_i + \epsilon_i \tag{4}$$

and equation (1) becomes the second stage. In equation (2), SEXFIRSTBORN is a dummy variable that equals 1 if the first-born is a female and 0 otherwise, and all of the other variables are the same as specified in equation (1).

4 Data and Sample Statistics

The data used in this study are taken from the third round of Indian District Level Household Survey (DLHS) collected in 2007-08. The sample is representative at district level, the lowest level of administration and policy-making. The DLHS covers 601 districts and on average draws a random sample of 1000 households in each district. The DLHS is similar to standard Demographic and Household Survey (DHS) that gathered fertility histories in addition to the household module. We use the information on relation to the household head in the household roster to construct our analytical sample. The household roster contains extensive information on personal and household characteristics. For each person in the household, information about, e.g., age, gender, schooling attendance, literacy, years of completed schooling, is available. We identify individuals who are labeled "sons/daughters" as the primary observation, and then obtain the family size by counting the number of children in the household. We then attach the data of parents, those who are labeled as "household head" or "spouse" to all the sons & daughters in the household.

For simplification, we trim the sample in the following ways. First, we restrict the sample to individuals who are either parents (head of the household, and spouse) or own children (who are either sons/daughters of the head of the household).[7] Second, we restrict the sample to households with atleast one child to use the first child's sex as an IV. Third, we restrict the sample to young mother who are not older than 35 at the time of survey, and the children to school going age of 6-20 years. Finally, we exclude households with missing or unreliable information on any of the variables used in the analysis. Finally, the analytical sample include 311,942 children from 145,962 households.

For all children and adults older than six the household roster collects information about their education. The main outcome variable this paper analyzes are years of schooling and primary school completion.

Table 1 reports the summary statistics of the analytical sample. The average age of children in the sample is 10 years with 3.72 as average years of schooling. 52 percent of first born children are female. Around 36 percent of all children in the sample have completed primary schooling. Fathers are relatively older than mothers. The average age of mothers is 31 years while fathers are 37 years old. We kept young parents in the sample to makes it fairly certain that no adult children have moved out of a household. We impose such a restriction because we are unable to track children who had already left the household by the time of the survey. Mothers have less education than fathers. The evareg years of schooling for mothers is 3.1 years, while fathers have 5.7 years of schooling.

The average family size is 3.5 with majority of the children in residing in rural areas. About 81 percent of children in the sample lives in rural areas. About two-fifth of the children belongs to low caste and one-fifth to high caste. Finally, around 46 percent of children belong to low-wealth household; slightly less (41 percent) belong to middle-wealth category; while the rest (13 percent) belong to high-wealth category.

[TABLE 1]

[7] We drop individuals who are son or daughter-in-law, grandchildren, parent, parent-in-law, brother/sister, brother or sister-in-law, niece or nephew, other relative etc.

5 Main Results

We consider two educational outcomes as measure of child quality: primary school completion and years of schooling. The main independent variable is the total number of 7-20 years old children in the family at the time os survey. The OLS results on primary school completion as well as years of schooling are reported in Table 2. Column (1-3) shows the results with primary school completion as the dependent variable, while col (4-6) show results with years of schooling as the dependent variable. In the most parsimonious model, we only include children's characteristics in addition to the district fixed-effect (col 1 & col 4). However, recognizing that parents characteristics may be an important factor in diluting the trade-off, in columns (2) & (4) we include parents' characteristics as well. Finally, following Black (2005), column (3) & col (6) examines if the inclusion of birth order changes the observed relationship in previous columns.

We observe a significantly negative correlation between family size and children's education. The results in table 2 implies that after controlling for children's characteristics, on average, an extra child in the family reduces the probability of completing primary schooling by 1.8 percentage points (col 1). We observe similar results when we include parents' characteristics (col 2) and birth order (col 3). It is difficult to find comparable estimates for India as this is the only study that have looked at trade-off in child quantity and quality in India. While a similar studies were done in other asian countries, such as China, Korea, Indonesia, but none of them have analyzed primary school completion as an outcome variable.

The last three columns in Table 2 reports the trade-off between family size and years of schooling. Similar to the pattern observed for primary school completion, we find significantly negative relationship between family size and years of schooling. The coefficient -0.122 in col (4) means that, on average, addition of an extra child reduces the average educational attainment of the children by 0.122 years. Inclusion of parents' characteristics and birth order, results do not change much- the estimate is close to 0.12 years. Black (2005), though used a twins as an instrument, reports coefficients of about -0.13 for the Unites States. Li, Zhang, and Zhu (2008), the only study available in Asia that looked at educational outcomes, report a coefficient of 0.03 for school enrollment.[8]

[Insert Table 2]

Recognizing the limitation of interpreting the OLS estimates as causal, we then proceed to instrument the main endogeneous variable, family size, by sex of the first child, and estimate the same relationship in 2SLS framework. The first three columns in Table 3 report the coefficients for primary school completion as the dependent variable, while the last three columns report the coefficients for years of schooling. The 2SLS estimates confirm the negative correlation observed in OLS estimation. The 2SLS coefficients are statistically significant at 1% level of significance and slightly bigger in magnitude than OLS coefficients. It is almost three times bigger than the OLS estimate. The IV estimates suggest that OLS coefficients underestimate the true trade-off and are downward biased.

From the first-stage regression, it follows that birth of first child as female increases the family size by 0.20 children (Column 2 of Table 3). This effect is significant at the one-percent level. Column 3 and 6 in Table 3 reports the 2SLS estimates of trade-off. The second-stage estimate on educational outcomes support the finding of a trade-off from the OLS analysis. According to 2SLS estimates, each additional child reduces the probability of completing primary school among all children in the family by 5 percentage points and schooling by 0.36 years.

[Insert Table 3]

[8]Lee (2008) is analyzes the trade-off using sex of the first child as an instrument in Korea, however, the paper uses educational expenditure as the dependent variable rather than educational outcomes.

6 Heterogeneity in the Trade-off

6.1 Caste and Rural-Urban Differences

Given that there is considerable rural-urban gap in the family size and educational attainment in India, we hypothesize the effect of family size on children's education to be different in rural and urban areas. For example, for our sample children, the primary school completion rate is 35% in rural areas while it is 41% in urban areas. Similarly, to capture the heterogeneity across different caste categories we also compare the effect of family size across low-vs-middle-high caste children.The results are presented in Table 4. The upper panel in table 4 shows the effect of family size on primary school completion while the lower panel displays the results for years of schooling. The results for caste categories in the first three columns show negative effects on primary school completion of 5.7 percentage points for low-caste children and of 5.0 percentage points for high-caste children. Surprisingly, the effect in rural sample is not much different from urban sample, and they are very close (0.052 vs 0.054) in magnitudes, though both coefficients are significant.

[Insert Table 4]

The quantity-quality trade-off is severe in lower panel when quality is measured by years of schooling. The 2SLS estimate for low-caste children is 0.42 years and 0.29 years for high-caste children. We find caste gradient in the trade-off- the effect size decreases in magnitude as the children move into higher caste category. We observe a similar pattern in rural areas. The trade-off coefficient is 0.40 and highly significant in the rural areas while in the urban sample the trade-off coefficient is insignificant, suggesting that quantity-quality trade-off exists mainly in rural areas. This finding is similar to Li, Zhang, and Zhu (2008), who also found that trade-off was more evident in rural parts of China and was negligible in urban areas.

6.2 Mother's Education and Wealth Differences

In this section, we allow the effect to vary by mother's education level and households' wealth level. Mother's education is categorized as illiterate, less than primary, and primary and more, and we estimate the main model on these three samples separately. Similarly, we also categorize the sample by households' wealth level to see whether the trade-off differs among low, median, and high and run the IV model separately for these categories. Results on the stratified sample is presented in Table 5. In the first three columns, sample is broken down by mother's education while by wealth in the last three columns.

The picture across OLS and 2SLS estimates are very consistent with our expectations. We expect to see a larger effect for illiterate and less-educated mothers due to lack of financial resources. Based on the same argument, we also expect to see a larger coefficient for the trade-off in households who are less well-off. In the 2SLS model, the trade-off coefficient decreases in size with the level of wealth for both the educational outcomes. However, for the high-wealth sample, the 2SLS coefficients are statistically insignificant.

[Insert Table 5]

To a large extent, the story remains the same when we stratify the sample by mother's education level. For primary school outcome, there is a negative effect of family size for illiterate and primary school completed mothers. Fir illiterate mothers, the trade-off coefficient is very severe- arrival of an extra child reduces the probability of completing primary school by 8.7 percentage points and reduces years of schooling by 0.7 years. We do not find any tangible evidence of trade-off in less than primary and primary completed mothers when the educational outcome analyzed is years of schooling. Though the coefficients monotonically decreases with the level of mother's education, the estimates are not significant in col 2 and 3.

7 Conclusions and Discussions

Testing the theoretical trade-off between the quantity and quality of children has been on the research agenda for a long time, however, the empirical evidence supporting the prediction of Beckerian model is limited. The empirical evidence has been mixed so far. A few studies find a negative effect of family size on the quality of children, measured by either education or health status (Ref). In contrast, others find no empirical support for the child quantity-quality trade-off (Ref). A variety of instruments such as twinning, sex of first child, sex of first two child, infertility etc. are used to address the endogeneity concern.

In this paper, we has used household data from India to test the empirical validity of child quantity-quality trade-off. A strong preference for sons over daughters in Indian societies allows us the use of a novel instrumental variable, namely sex of first birth, to test the Q-Q trade-off. Testing this model has important policy implications. From policy point of view, it is important to know the extent to which a policy formulated to control population improves the human capital of the country and quality of the labor force. Not only the quantity of human capital plays a role, rather quality of human capital is equally important for economic development.

We find that Beckerian theory of child quantity-quality trade-off holds in India. Family size has significant negative causal impact on educational outcomes of children. After controlling for potential endogeneity, an additional child in the family reduces the probability of completing primary school for all children by 5 percentage points and years of schooling by 0.36 years, hence a strong support to Becker's trade-off hypothesis. The observed trade-off exists after including child and parents characteristics. We find non-uniformity in the existence of trade-off between rural and urban India. The negative relationship between family size and children's education is more pronounced and evident among rural households who are severely budget-constrained. Urban children are less likely to face the trade-off.

The effect also differs by caste, mother's education level, and household wealth. For children belonging to low and middle caste, the trade-off is severe compared to high-caste children. More educated mothers are also able to mitigate the trade-off as the trade-off is only evident for illiterate mothers. Similarly, we observe a wealth-gradient in the trade-off across wealth groups, with trade-off being more pronounced in low-wealth households with an extra children reducing the years of schooling by as high as 0.6 years.

Our findings are also supportive of theoretical work by Galor and Moav (2002), who were the first to argue that the quantity-quality trade-off was decisive to economic advancement, not just from the onset of the demographic transition, but throughout human history.

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.

Our results suggest that policymakers in developing countries should invest more in education in areas and households for whom the trade-off is severe in order to mitigate the adverse impacts of larger family size.

References

Angrist JD, Evans W (1998) Children and their parents' labor supply: evidence from exogenous variation in family size. *American Economic Review* 88(3):450-477

Angrist JD, Lavy V, Schlosser A (2005) New evidence on the causal link between the quantity and quality of children. NBER Work Pap no.11835

Angrist, J. D., V. Lavy, and A. Schlosser (2010.) "Multiple Experiments for the Causal Link between the Quantity and Quality of Children," J*ournal of Labour Economics*, 28:773-824.

Black, Sandra E, Devereux, Paul J, Salvanes, Kjell G. (2005) "The more the merrier? The effect of family size and birth order on children's education," *Quarterly Journal of Economics*, 120(2):669-700.

Caceres-Delpiano, Julio (2006) The Impacts of Family Size on Investment in Child Quality, Journal of Human Resources, XLI (4), 738754.

Conley D., Glauber R. (2005) Parental Educational Investment and Children's Academic Risk: Estimates of the Effects of Sibship Size and Birth Order from Exogenous Variation in Fertility. New York: University of New York.

Galor, O. and O. Moav (2002.) "Natural selection and the origin of economic growth," *Quarterly Journal of Economics*, 117:1133-1191.

Glick, P., A. Marini, and D.E. Sahn (2007) Estimating the Consequences of Unintended Fertility for Child Health and Education in Romania: An Analysis Using Twins Data. *Oxford Bulletin of Economics & Statistics*, 69, 667-691.

Goux, Dominique and Eric Maurin (2005) The effect of overcrowded housing on children's performance at school," Journal of Public Economics, 89 (5-6), 797819.

Haan, Monique De (2005) , Birth Order, Family Size and Educational Attainment," Tinbergen Institute Discussion Papers 05-116/3, Tinbergen Institute December 2005. available at http://ideas.repec.org/p/dgr/uvatin/20050116.html.

Klemp P.B., Mark and Weisdorf L. Jacob (2011) The Child Quantity-Quality Trade-Off during the Industrial Revolution in England, Working Paper, department of Economics, University of Copenhagen.

Lee, Jungmin (2008) "Sibling size and investment in children's education: an asian instrument", *Journal of Population Economics*, vol. 21, issue 4, pages 855-875.

Li, Hongbin; Zhang, Junsen, and Zhu Yi (2008) "The Quantity-Quality Trade-Off of Children in a Developing Country: Identification using Chinese Twins", *Demography*, vol. 45, number 1, pages 223-243.

Ponczek, Vladimir, Souzay P. Andre (2011) New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country, C-Micro Working Paper Series, Sao Paulo School of Economics, Getulio Vargas Foundation.

Qian, Nancy (2009) Quantity-Quality and the One Child Policy: The Only Child Disadvantage in School Enrollment in China. Working Paper, Department of Economics, Yale university.

Retherford, Robert D and T K Roy (2003) Factors Affecting Sex-Selective Abortion in India and 17 Major States, National Family Health Survey Subject Reports 21, International Institute for Population Sciences, Mumbai, India.

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

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

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

Wang, Lee, Millimet, Daniel (*Forthcoming*) "Is the Quantity-Quality Trade-off a Trade-off for All, None, or Some?" *Economic Development & Cultural Change.*

Table 1: Descriptive Statistics of the sample

	Mean	Standard Deviation
	(1)	(2)
Age (6-20 years old)	10.37	3.22
Gender of first child (female=1)	0.52	*binary*
Years of schooling	3.72	2.83
Primary school	0.36	*binary*
Mother's age	31.33	3.15
Father's age	36.99	4.71
Mother's years of schooling	3.11	4.07
Father's years of schooling	5.68	4.70
Family size	3.53	1.31
Rural	0.81	*binary*
Low caste (SC & ST)	0.40	*binary*
Middle caste (OBC)	0.39	*binary*
High caste	0.21	*binary*
Low wealth	0.46	*binary*
Median wealth	0.41	*binary*
High wealth	0.13	*binary*
No of households	145,962	
No of observation	311,942	
No of district	601	

Notes: Standard deviations are shown in parentheses. All sampled children were 6-20 years old at the time of survey (2007-08) Mother's sample is restricted to 22-35 years.

Table 2: OLS estimates of the effect of family size on children's educational outcome

	Primary School Completion			Years of Schooling		
	(1)	(2)	(3)	(4)	(5)	(6)
Family size	-0.018***	-0.017***	-0.019***	-0.122**	-0.108***	-0.121***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)
Birth order	no	no	yes	no	no	yes
Children's control	yes	yes	yes	yes	yes	yes
Parents' controls	no	yes	yes	no	yes	yes
District fixed-effect	yes	yes	yes	yes	yes	yes
N	311,942	311,942	311,942	311,942	311,942	311,942
r2	0.57	0.57	0.57	0.79	0.79	0.79

Notes: *, **, and *** represent significance levels of 10, 5, and 1 percent. Robust standard error, clustered by district, are shown in parentheses. Children's controls include age, age square, religion, caste, SES and rural dummies. Parent controls include age, age square, and education levels of father and mother. Family size is total number of children in the family at the time of survey.

Table 3: IV estimates of the effect of family size on children's educational outcome

	Primary School Completion			Years of Schooling		
	OLS (1)	First Stage (2)	2SLS (3)	OLS (4)	First Stage (5)	2SLS (6)
Family size	-0.019 *** (0.001)	0.203*** (0.008)	-0.050*** (0.010)	-0.121*** (0.004)	0.203*** (0.008)	-0.363*** (0.061)
Birth order	yes		yes	yes		yes
Children's control	yes		yes	yes		yes
Parents' controls	yes		yes	yes		yes
District fixed-effect	yes		yes	yes		yes
N	311,942		311,942	311,942		311,942
r2	0.57		0.28	0.79		0.39

Notes: *, **, and *** represent significance levels of 10, 5, and 1 percent. Robust standard error, clustered by district, are shown in parentheses. Children's controls include age, age square, religion, caste, SES and rural dummies. Parent controls include age, age square, and education levels of father and mother. Family size is total number of children in the family at the time of survey.

Table 4: IV estimates of the effect of family size on children's educational outcome

	Low Caste	Middle Caste	High Caste	Rural	Urban
	(1)	(2)	(3)	(4)	(5)
Primary School Completion					
OLS	-0.017***	-0.020***	-0.019***	-0.019***	-0.018***
	(0.001)	(0.001)	(0.002)	(0.0008)	(0.002)
2SLS	-0.057 ***	-0.048***	-0.050**	-0.052***	-0.054**
	(0.016)	(0.015)	(0.020)	(0.010)	(0.027)
Years of Schooling					
OLS	-0.102***	-0.120***	-0.134***	-0.116***	-0.131***
	(0.006)	(0.006)	(0.009)	(0.005)	(0.009)
2SLS	-0.424***	-0.374***	-0.299	-0.403***	-0.252
	(0.091)	(0.098)	(0.122)	(0.064)	(0.163)
No of observations	122,073	119,584	63,768	252,775	59,167

Notes: *, **, and *** represent significance levels of 10, 5, and 1 percent. Robust standard error, clustered by district, are shown in parentheses. Children's controls include age, age square, gender, religion, caste, SES and rural dummies. Parent controls include age, age square, and education levels of father and mother. Family size is total number of children in the family at the time of survey. Instrument is the a dummy variable indicating if the first-born is a female child. Low caste includes scheduled caste (SC) and scheduled tribe (ST). Middle caste is the other backward caste (OBC).

Table 5: IV estimates of the effect of family size on children's educational outcome

	Mother's Education			Wealth Category		
	Illiterate	Less than Primary	Primary & above	Low	Median	High
	(1)	(2)	(3)	(4)	(5)	
Primary School Completion						
OLS	-0.020***	-0.018***	-0.020***	-0.017***	-0.017***	0.011***
	(0.0009)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
2SLS	-0.087***	-0.020	-0.039**	-0.070***	-0.057***	0.004
	(0.014)	(0.020)	(0.019)	(0.016)	(0.013)	(0.028)
Years of Schooling						
OLS	-0.122***	-0.119***	-0.091***	-0.098***	-0.120***	0.099***
	(0.005)	(0.009)	(0.007)	(0.006)	(0.006)	(0.012)
2SLS	-0.669***	-0.183	-0.173	-0.601***	-0.332***	0.019
	(0.088)	(0.117)	(0.115)	(0.099)	(0.081)	(0.169)
No of observations	170,426	52,644	82,355	139,772	125,567	40,045

Notes: *, **, and *** represent significance levels of 10, 5, and 1 percent. Robust standard error, clustered by district, are shown in parentheses. Children's controls include age, age square, gender, religion, caste, SES and rural dummies. Parent controls include age, age square, and education levels of father and mother. Family size is total number of children in the family at the time of survey. Instrument is the a dummy variable indicating if the first-born is a female child.