Sam Houston State University

Department of Economics and International Business
Working Paper No. 16-05
February 2016

The Effect of Birth Order on Schooling in India

Santosh Kumar
Department of Economics and International Business
Sam Houston State University
skumar@shsu.edu

Abstract

Using large nationally representative data, I estimate the effect of birth order on educational outcomes of children in India. To establish causality, endogeneity of family size is addressed by approaching an instrumental variable method. Employing a district fixed effects model and proportion of boys in the family as the instrument for number of children, I show that later-born children attain higher education compared to earlier-born children. Results are robust to inclusion of child, parents, and household characteristics.

***Keywords*:** Birth order; family size; education; India

***JEL Codes*:** D1, I2, J1

1. Introduction

Low human capital is considered an important cause for the underdevelopment and poor economic performance in developing countries. While exploring the process of human capital accumulation, a large body of research has examined the relationship between birth order and the educational outcomes of children and found that birth order has significant impact on the educational attainment of the children (Black, Devereux, and Salvanes, 2005). Existing evidence, however, is far from being settled because the birth order effect on education is confounded by the family size effect, since higher birth order children live in larger households. For example, Black et al. (2005) shows that the negative effect of family size on education disappears when birth order of children is taken into account. Thus, it is important to address the endogeneity of family size while estimating the causal impact of birth order on education of the children.

The empirical evidence on the effect of birth order on education is mixed. Studies from developed countries show negative birth order effect and found that first-born children attain higher education and have better labor market outcomes compared to later-born children (Black et al., 2005; Booth and Kee, 2008; Conley and Glauber, 2006). In contrast, studies from developing countries document a positive birth order effect on education and show that first-born children attain less schooling than the later-born children (De Haan et al., 2014; Ejrnaes and Pörtner, 2004; Emerson and Souza, 2008).

Given the mixed evidence on the birth order effect on education of children and scarcity of credible birth order studies in low-income countries and none in India, this study takes the Instrumental Variable (IV) approach and estimate a Two Stage Least Square (2SLS) model to investigate the causal effect of birth order on educational

attainment of children in India. A positive birth order effect documented in low-income countries is plausible if resource-constrained parents send their older children to work and those extra earnings are spent on the schooling of their younger children. In addition, older children in developing countries are more likely to work because they are more experienced and productive compared to younger children (Booth and Kee, 2008). Maternal experience is another channel that may generate positive birth order effect, since experienced mothers may provide a better nurturing environment for their later-born children.

2. Data

I use the third round of the District Level Household Survey (DLHS). DLHS sampled about 1000 households in each district of India and is representative at the district level, the lowest tier of administration. I restrict the sample to women with two and more children, whose eldest child was 20 years old at the time of the survey. The resulting sample consists of 393,597 children from 162,619 households. The dependent variables are current school enrollment, primary school completion, and years of schooling.

Means, standard deviations, and number of observations for each variable are reported in Table 1. The average years of schooling attained by the children is 3.07 years and about 29% of the children have completed primary schooling. The percentage of first-born is 41% while 32% of the children are of second birth order. On average, the sampled children are 10 years of age. The family size variable, which is denoted as number of 0-20 years old children in the household, has a mean value of 3.54 children. The majority of the children is Hindu and lives in rural areas.

3. Empirical Estimation

The main objective of this study is to establish the causal effect of birth order on schooling in India. However, the birth order effect on educational outcomes is complicated, due to positive correlation between birth order and family size. Since high birth order children are observed in households with a larger family size, it is plausible that the birth order effect is confounded by the effect of family size on educational outcomes. I address this endogeneity by estimating the following 2SLS model:

First stage:

$$Family\_Size_{hd} = \alpha_0 + \alpha_1 Share\_Boys_{hd} + \alpha_2 \sum_{i=1}^{5}(Birth\_Order_{ihd}) + \alpha_3 X_{ihd} + \gamma_d + \epsilon_{ihd} \quad (1)$$

Second stage:

$$Schooling_{ihd} = \beta_0 + \beta_i \sum_{i=1}^{5}(Birth\_Order_{ihd}) + \beta_6 Family\_Size_{hd} + \beta_7 X_{ihd} + \eta_d + \mu_{ihd} \quad (2)$$

where $Schooling_{ihd}$ is the educational outcomes of the child *i* living in household *h* in district *d*. $Family\_Size_{hd}$ refers to the number of children below 21 years of age in the household, $Share\_Boys_{hd}$ is the proportion of boys in the family, $Birth\_Order_{ihd}$ denotes the birth order of child *i;* X_{ihd} represents vector of observed control variables such as, child's age and gender, rural location, caste, father's age, parents education, and religion of the household; γ_d and η_d are district fixed effects to control for time invariant characteristics of the districts, and ϵ_{ihd} and μ_{ihd} are the random error terms.

The main coefficients of interest are $\beta_1, \beta_2 \ldots., \beta_5$ in equation (2). These coefficients show the birth order effects after controlling for family size and other confounding variables. $Share\_Boys_{hd}$ denotes the proportion of boys in the family and is the instrument for the endogenous variable, family size. The instrument exploits the widely documented "son preference" culture in India. Indian households tend to prefer sons to daughters because sons are considered to carry the family legacy forward and are expected to provide care for the parents in their old age. If parents prefer sons, then siblings' sex composition in the household would be favorable to sons; one would expect a negative correlation between proportion of boys and family size, and therefore expect β_1 to be negative in equation (1). Exclusion restriction implies that $Share\_Boys_{hd}$ must not have independent effect on schooling. It is difficult to test the exclusion restriction but one can imagine a situation in which exclusion restriction may not be accurate. If wealthier households have more boys due to sex-selective abortions and also invest more in their children's education, then the 2SLS estimates will be upward biased.

4. Results

Consistent with the findings in the previous studies in developing countries, the OLS results reported in Table 2 show evidence of positive birth order effects on schooling in India. Relative to the first-born, later-born children are more likely to be enrolled in school, complete primary school, and attain more years of schooling. However, given the limitation of OLS coefficients interpretation as causal, I proceed to estimate a 2SLS model in Table 3.

Panel A in Table 3 presents the first-stage results, which show that households with a higher proportion of boys have a smaller family size (1.32 fewer children in

Column 1). The first-stage results in columns 2-3 also report a significantly negative impact of proportion of boys on family size. This satisfies the instrument relevance condition.[2] Regarding the effects of birth order on educational outcomes, the 2SLS estimates in panel B, show that first-born children are significantly less likely to be attain higher education compared to their later-born siblings. The 2SLS results are smaller than the OLS estimates, suggesting that had the endogeneity in the family size not been controlled, the OLS coefficients would have overestimated the true birth-order effects. The OLS estimates may be upward biased if there are omitted variables that are positively correlated with the family size as well as the schooling outcomes. If richer households have a preference for bigger family size and also care more about education, then the OLS estimates will be higher than the IV estimates.

Relative to the first-born, children of second birth order are 0.2 percentage points more likely to be enrolled, 0.4 percentage points more likely to complete primary education, and attain 0.02 extra years of schooling. I also find that the size of estimated birth order effects increases with the birth order. For example, relative to a first-born child, children with fifth birth order gain 0.44 extra years of schooling, while the second-born children gains only 0.02 years of schooling.

5. Conclusion

To the best of my knowledge, this is the first study to estimate the effect of birth order on educational attainment of children in India. I find evidence of positive birth order effects implying that later-born children have significantly better educational outcomes compared to early-born children; however, there are several other channels that

[2] Panel A also reports F-statistics for weak identification tests. The null hypothesis of weak instrument is rejected: the F-statistics are greater than 10.

may impact education in addition to the birth order. For example, education of later-born children may be affected by higher household income or the improvement in educational infrastructure over time. However, inclusion of household's wealth at the time of the survey and district fixed effects in the regression model should rule out these mechanisms.

Furthermore, if parents engage in sex-selective abortions or mortality rate differs by gender, then the proportion of boys may not satisfy the exclusion restrictions; therefore, the results should be interpreted carefully. However, sex-selective abortions may not be a concern in this study because India banned pre-natal diagnostic of a fetus's gender in 1995 that made it impossible for parents to know the gender of the child at fetal stage.[3] Lack of mortality information in the data would not allow us to check if differential mortality by gender may be biasing the results.

From a policy perspective, I conclude by recommending that nutritional and financial supplementation policies targeted at improving children's education should take into account the birth order effect and more resources should be targeted to earlier-born children in low-income countries.

[3] Kugler and Kumar (2015) show no evidence of sex-selective abortions at first parity in India.

References:

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

Booth, A. L. and Kee, H. J. 2008. "Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment." *Journal of Population Economics* 22:367–97.

Conley, D. and Glauber, R. 2006. "Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility." *The Journal of Human Resources* 41:722–37.

De Haan, M., Plug Erik, Rosero, Jose. 2014. "Birth Order and Human Capital Development: Evidence from Ecuador." *Journal of Human Resources* 49(2):359-392.

Ejrnaes, M. and Pörtner, C. C. 2004. "Birth Order and the Intrahousehold Allocation of Time and Education." *Review of Economics and Statistics* 86:1008–19.

Emerson, P. M. and Souza, A. P. 2008. "Birth Order, Child Labor, and School Attendance in Brazil." *World Development* 36:1647–64.

Kugler, Adriana D. and Kumar, Santosh. 2015. "Preference for Boys, Family Size, and Educational Attainment in India NBER Working Paper No. 21138.

Table 1: Descriptive statistics

Variables	Mean	Standard deviations	N
Outcome variables			
Currently enrolled in school	0.95	0.21	346034
Completed primary schooling	0.29	0.46	393597
Years of schooling	3.07	2.91	393597
Child's characteristics			
First birth order	0.41		393597
Second birth order	0.32		393597
Third birth order	0.17		393597
Fourth birth order	0.07		393597
Fifth birth order	0.03		393597
Child age (in years)	9.60	3.45	393597
Male child	0.52	0.50	393597
Proportion of boys	0.45	0.28	393597
Parent's characteristics			
Father is primary schooled	0.20	0.40	393597
Mother is primary schooled	0.16	0.37	393597
Father's age	36.48	4.80	393597
Household characteristics			
Family size	3.54	1.33	393597
Low caste	0.41	0.49	393597
Religion (Hindu)	0.77	0.42	393597
Rural	0.82	0.39	393597
Poor household	0.49	0.50	393597

Table 2: OLS effect of birth order on schooling outcomes

	Current enrollment	Primary school completion	Years of schooling
	(1)	(2)	(3)
Second birth order	0.006 ***	0.010***	0.052***
	(0.0008)	(0.001)	(0.006)
Third birth order	0.021***	0.019***	0.171***
	(0.001)	(0.002)	(0.011)
Fourth birth order	0.039***	0.052***	0.385***
	(0.002)	(0.003)	(0.018)
Fifth birth order	0.061***	0.127***	0.752***
	0.003	(0.004)	(0.028)
District fixed effect	Yes	Yes	Yes
R-squared	0.15	0.52	0.70
Observations	345985	393597	393597

Notes: Robust standard errors clustered at district level are reported in the parentheses. Control variables: child's age, age square, and gender; household size, rural dummy, lowcaste dummy, religion, father's age, father's education, mother's education, and household wealth quintile.

* significant at 10 percent

** significant at 5 percent

*** significant at 1 percent level

Table 3: 2SLS effect of birth order on schooling outcomes

	Current enrollment	Primary school completion	Years of schooling
	(1)	(2)	(3)
Panel A: First stage results			
Share of boys	-1.32***	-1.34***	-1.34***
	(0.020)	(0.020)	(0.020)
F-test of excluded instrument	4135.04	4469.37	4469.37
Weak identification test			
Cragg-Donald Wald F statistic	37394.77	40969.75	40969.75
Kleibergen-Paap Wald rk F statistic	4135.04	4469.37	4469.37
Panel B: 2SLS results			
Second birth order	0.002*	0.004***	0.016**
	(0.0009)	(0.002)	(0.007)
Third birth order	0.008***	0.002	0.066***
	(0.001)	(0.002)	(0.014)
Fourth birth order	0.015***	0.021***	0.193***
	(0.002)	(0.004)	(0.022)
Fifth birth order	0.023***	0.078***	0.438***
	(0.004)	(0.007)	(0.036)
District fixed effect	Yes	Yes	Yes
R-squared	0.13	0.49	0.68
Observations	345985	393597	393597

Notes: Robust standard errors clustered at district level are reported in the parentheses. Control variables: child's age, age square, and gender; household size, rural dummy, lowcaste dummy, religion, father's age, father's education, mother's education, and household wealth quintile.

* significant at 10 percent

** significant at 5 percent

*** significant at 1 percent level