Journal of Population Economics (2019) 32:1353–1399
https://doi.org/10.1007/s00148-019-00730-z

ORIGINAL PAPER

Family size and schooling in sub-Saharan Africa: testing the quantity-quality trade-off

Sahawal Alidou[1,2] · Marijke Verpoorten[1,2]

Received: 2 February 2018 / Accepted: 1 February 2019 / Published online: 5 April 2019

Abstract
Many family planning programs are based on the idea that small families lead to improved development outcomes, such as more schooling for children. Because of endogeneity issues, this idea is however difficult to verify. A handful of studies have made use of twin birth to deal with the endogeneity of family size. We do so for sub-Saharan African countries. In a compilation of 86 survey rounds from 34 countries, we exploit the birth of twins to study the effect of a quasi-exogenous increase in family size on the schooling of children at the first, second and third birth order. Our findings do not support the generally assumed negative effect of family size on schooling.

Keywords Family size · Schooling · Quantity-quality trade-off · Sub-Saharan Africa

JEL classification D1 · O1

1 Introduction

Countries worldwide have devoted much effort and resources to family planning programs (Bongaarts 2009). Most of these programs have been voluntary, but some have left little choice to parents, such as China's one-child program and India's sterilization camps. A major assumption underlying these programs is that "a small family is a happy family",[1] or that a reduction in family size enables families to raise investments in human capital per

[1]This text featured on a 2-rupee coin issued by India in 1993, to promote family planning.

We received much appreciated comments from Elena Briones Alonso, Philip H. Ross, Pieter Serneels, Nik Stoop and the participants at seminars, conferences and workshops in Leuven (LICOS-KU Leuven seminar), Antwerp (Doctoral Day) and Oxford (2017 CSAE conference), as well as from the Editor and two anonymous referees of this journal.

Responsible Editor: Alessandro Cigno

✉ Marijke Verpoorten
marijke.verpoorten@uantwerp.be

Extended author information available on the last page of the article

child, leading in its turn to a stronger economy (Bongaarts 2009). Intuitively, the assumed causality between small family sizes and high schooling attainments makes sense, dividing scarce resources among fewer children, and leaves each child with more resources.

This intuition found support in well-known social science research. Blake (1989), studying US families, famously concluded that children from one- and two-child families are better educated and more successful than children in larger families because their parents have more time and money to invest in them. This "resource dilution model" was backed up by economic theory, in particular in a pioneering paper written by Becker (1960), in which the quantity and quality of children are modelled as substitutes from the parents' point of view.

However, there also exist theories that support a *positive* causal effect of family size on children's schooling. These theories break with Blake's and Becker's assumptions that children only imply a cost to parents, and that more children imply higher costs. As such, the quantity-quality trade-off needs no longer hold when allowing for part-time child work (Mueller 1984a; Marteleto and de Souza 2013), or when older children work to provide for the younger ones (the so-called 'chain arrangement'), or when there are economies of scale in raising children, with children sharing clothes, text books, transport to school or knowledge and skills (Guo and Van Wey 1999; Rosenzweig and Zhang 2009; Steelman et al. 2002; Qian 2009). Economies of scale can also be present in household chores, such that the time each child spends on chores reduces with the number of siblings, thus freeing up time for school.

Despite the diversity of theoretical predictions, it is hard to move away from the idea of a negative causal relation. An important reason for its stickiness lies in the strong negative correlation between family size and children's schooling, and the difficulty to empirically distinguish this correlation from the actual causal effect of family size on schooling. To make this distinction, one needs to purge the correlation of confounding factors. Most importantly, parents' characteristics determine preferences for both the number of children and their years of schooling. For instance, mothers who enjoyed more years of schooling generally prefer smaller families and, at the same time, are likely to give more importance to their children's schooling. Other confounding factors include wealth, social norms regarding fertility and child labour, labour market opportunities for adults and children and the availability and quality of old-age security schemes and education policies (Rosenzweig and Wolpin 1980; Angrist et al. 2010; Black et al. 2010). To the extent that these confounding factors are not perfectly observed and controlled for, the estimated relation between family size and children's schooling is plagued by endogeneity issues.

In this paper, we aim to remedy this endogeneity problem in a sample of children from 208,729 households across 34 sub-Saharan African countries. In 3844 of these households, twins were born, causing a quasi-exogenous increase in household size. Provided controls for certain mother characteristics (Smits and Monden 2011; Bhalotra and Clarke 2016)[2] and for the health condition of twins (Rosenzweig and Zhang 2009),[3] twin birth can be used as a plausibly exogenous instrumental variable to isolate

[2] According to Smits and Monden (2011), twinning is associated with maternal age, maternal height, smoking, oral contraceptive use, race and ethnicity.

[3] Rosenzweig and Zhang (2009) point out that twins have lower-average birth weight than singletons, and perhaps worse health or cognitive achievement later on, and this may threaten the exclusion restriction if parents therefore allocate resources away from twins, towards older singleton-birth children.

the causal effect of family size on the educational outcomes of children born *prior* to the twin birth (Angrist et al. 2010; Conley et al. 2012). The same does not apply for children born *after* the twins, because their birth can be the result of parental choice. Concretely, in our instrumental variable (IV) approach, we look at the outcomes of first-, first- and second-, and first-, second- and third-born children respectively in families of two or more, three or more and four or more children, using the birth of twins at the second birth, third birth and fourth order as the instrumental variable.

Our empirical investigation adds to the body of literature that has tried to empirically unearth the quantity-quality trade-off by relying on a number of techniques, such as tracing children's intellectual abilities in a longitudinal analysis of families (Guo and Van Wey 1999) and exploiting the gradual roll-out and subsequent relaxation of family planning programs (Liu 2014; Qian 2009), a randomized controlled trial in family planning (Sinha 2005; Joshi and Schultz 2007) and instrumental variable approaches based on reported miscarriages (Maralani 2008), siblings' sex composition (Angrist et al. 2010; Conley and Glauber 2006; Black et al. 2010; Fitzsimons and Malde 2014; Lee 2008) and twin birth (Rosenzweig and Wolpin 1980; Rosenzweig and Zhang 2009; Angrist et al. 2010; Black et al. 2005, 2010; Marteleto and de Souza 2012, 2013; Mogstad and Wiswall 2016; Bhalotra and Clarke 2016). These exercises in causal identification have not uniformly yielded negative estimates of the effect of sibship on schooling. Instead, the effect turns out to vary over time, across regions and subpopulations, across birth order and across the exact outcome of interest studied (e.g. private schooling, school enrolment, educational attainment or IQ).[4]

None of these exercises in causal identification has however looked specifically at sub-Saharan African countries.[5] Our paper fills this gap. There are several reasons why sub-Saharan Africa (SSA) provides an interesting setting for such analysis. First, in SSA, the majority of households face tight budget constraints, schooling is barely compulsory and children's participation in the labour market and in time-consuming household chores is socially still largely accepted (Bass 2004). Combined, these features make it very likely that a household's decision to invest in children's formal education involves important trade-offs. Second, in most African cultures, family members are bound to act for the benefit of the collective, be it the nuclear family or the extended family, the clan or ethnic group (Lloyd and Blanc 1996). Regarding the decision to invest in schooling, this implies that the benefits of schooling are expected to be shared—giving, for instance, way to the chain arrangement in which earlier-born children are sent to school and use their wage earnings to invest in their younger

[4] In the first study using twins as an instrument, Rosenzweig and Wolpin (1980) find that schooling levels of Indian children decrease with exogenous increases in fertility. Consecutive studies using twins find mixed results. Cáceres-Delpiano (2006), relying on US Census data, finds a negative effect of family size on the probability of enrolment in private schooling, Black et al. (2010) find that family size negatively affects the IQ of younger cohorts in Norway and Li et al. (2008), relying on Chinese twins for identification, find a negative effect of sibship on schooling attainments. However, family size is found to have no effect on children's educational attainment in Norway (Black et al. 2005) and in Israel (Angrist et al. 2010). Furthermore, Marteleto and de Souza (2012), studying the effect of family size on adolescents' schooling in Brazil, find a positive effect in periods and regions in the earlier stages of socioeconomic development, but these effects disappear for recent periods "when the opportunities for child farm work have declined, education has expanded, and fertility has declined to below-replacement levels" (p. 1473).

[5] Bhalotra and Clarke (2016) tested the quantity-quality trade-off in a large sample, with data from 72 countries among which many sub-Saharan African countries, but they did not specifically focus on SSA.

siblings later on, rendering a quantity-quality trade-off superfluous (Baland et al. 2016; Mueller 1984b). Third, in many of the least developed regions in SSA, the quality of schooling may be low (e.g. Chaudhury et al. 2006), or labour market opportunities may be lacking (e.g. Garcia and Fares 2008), both of which depress the returns to education. Hence, additional schooling may not be a way to invest in children's *quality*. Finally, SSA still is the region with the highest fertility and lowest educational enrolments and attainments,[6] increasing the relevance of research on these issues.

Our use of Demographic and Health Survey (DHS) data comes with both pros and cons. Among the pros, we count the availability of demographic and health data of mothers, allowing us to explicitly control for factors such as ethnicity and mothers' height and health that are likely to affect twinning. In doing so, we further purge this instrument of potential sources of endogeneity. Second, the detailed information on children's health allows us to verify that parents do not allocate resources away from twins—who may suffer from poorer health at birth—towards older singleton-birth children, thus further providing confidence in our instrument. Third, the DHS data allows us to distinguish between three distinct proximate causes that underlie differences in educational attainment: school enrolment, school starting age and dropout. Among the cons, we face the constraint that, beyond the fourth birth order, there are insufficient observations on twin births in the DHS to provide enough power in the first stage. Our analysis therefore only focuses on the effect of family size on outcomes of children of the first, second and third birth order, and its findings cannot readily be generalized to siblings at higher birth orders (Booth and Kee 2009; Qian 2009; Mogstad and Wiswall 2016). Another point of attention is that the DHS focuses on mothers of childbearing age (15–49 years), such that the observed number of children may be below the eventual number of children and the reported level of schooling may be below the eventual schooling attainment. Consequently, the relation that we observe between sibship and schooling captures a snapshot in time of a process in motion, not its final stage. Furthermore, the DHS provides no systematic information on health for children above 5 years old, nor information on test scores or other proxies for the quality of schooling. We therefore limit our analysis to the quantity of schooling, measured by the number of years of schooling. Finally, as will be explained further on, we need to address the complexity of family structure in SSA, which includes polygamy and a non-negligible number of children living outside the household, often with extended family.

In the next section, we explain the empirical strategy. Then, we describe our data and present the results. In our results section, we find overall no evidence of a quantity-quality trade-off. In particular, we cannot reject the null hypothesis of no relation between family size and schooling in the subsamples of families with two or more, or four or more children, while we find a significantly positive effect of family size on schooling in the sample of families with three or more children.

[6] Data on 46 sub-Saharan African countries reveal an average total fertility rate (TFR) of 4.97 in 2014 (World Bank 2016) and a gross school enrolment (GSE) rate of 98.96 in 2013 (UIS Stat. UNESCO, 2016). The 2014 TFR is markedly lower than that in 1985 (6.62) but still much higher than the 2014 TFR in South Asia (2.56) and in Latin America (2.11). Similarly, the 2013 GSE has greatly improved from its level in 1985 (79.24) but still compares unfavourably with the 2013 GSE in Southern Asia (109.06) and in Latin America and the Caribbean (109.03).

2 Empirical strategy

In our empirical analysis, we use an age-standardized z-score[7] for the years of schooling as our main outcome variable. We consider three subsamples: firstborn from families with at least two children (2+ sample), first and second born from families with at least three children (3+ sample) and first, second and third born from families with at least four children (4+ sample). We first examine the relation between schooling and family size using ordinary least squares (OLS). Then, we respectively use twinning at the second, third and fourth birth to instrument the number of children in the '2+ sample', '3+ sample' and '4+ sample'. Focusing on the schooling outcomes of $n-1$ children born prior to twins at the nth birth avoids selection problems that arise because "families who choose to have another child after a twin birth may differ from families who choose to have another child after a singleton birth" (Black et al. 2005). To ensure the validity of our instrument, we duly control for a battery of mother-level characteristics that may affect twinning and, at the same time, correlate with children's schooling.

Concretely, the OLS specification takes the following form:

$$\text{Education}_{hmfi} = \beta_0 + \beta_1 \text{ Number of children}_h + \beta_2 X_{hmfi} + \beta_3 X_{hm} + \beta_4 X_{hf} + \beta_5 X_h + \varepsilon_{hmfi} \quad (1)$$

where h indicates household, m mother, f father and i the individual child. Education$_{hmfi}$ is equal to the child's z-score. The Number of children$_h$ is a count variable that is equal to the total number of sons and daughters of the household head, residing in the household. X_{hmfi} is a vector of child-level characteristics that includes an indicator variable for sex, indicator variables for each age in the range 6 to 18, the child's birth year and its month of birth. X_{hm} is the set of mother-level characteristics including her years of schooling, age, age squared, height, religion, ethnicity,[8] the total number of her children who have died and whether a child of her has died before its first birthday[9]. X_{hf} is the set of father-level characteristics comprising his age and years of schooling. X_h includes household's residence area (urban/rural) and wealth quintile.[10] To account for a within-household correlation of the residuals, we cluster all error terms (ε_{hmfi}) at the

[7] Since the variation in completed grades across country and by child age is so large, it is unlikely to be completely captured by country and age fixed effects. We therefore use age-standardized z-score for which the reference group comprises children in the same country and birth cohort. In the robustness section, we show that our results remain stable when using completed grades as outcome variable.

[8] Mother's ethnicity is country specific and generated as follows: (country code × 1000) + ethnic group code. The insertion of the complete set of ethnicity fixed effects (as dummy variables) makes country fixed effects superfluous. In some DHS rounds, ethnicity is omitted (e.g. Rwanda, Burundi). These rounds are omitted from our baseline results but included in a robustness check that uses region of residence of the household as a proxy for ethnicity.

[9] Infant deaths proxy among others the reproductive health of the mother.

[10] For each round, the DHS separates all interviewed households into five wealth quintiles based on their wealth index. The wealth index is a composite measure of a household's cumulative living standard. It is calculated using principal component analysis on easy-to-collect data on a household's ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities (Standard Recode Manual for DHS 6, 2013).

DHS cluster level, which equate villages in rural areas and city blocks in urban areas.[11]

The second stage of the IV specification is captured by the following equation:

$$\text{Education}_{hmfi} = \delta_0 + \delta_1 \text{ Number of } \widehat{\text{children}}_h + \delta_2 X_{hmfi} + \delta_3 X_{hm} + \delta_4 X_{hf} + \delta_5 X_h + \omega_{hmfi} \quad (2)$$

in which family size is instrumented in the first-stage equation:

$$\text{Number of children}_h = \alpha_0 + \alpha_1 \text{ Twin}_h + \alpha_2 X_{hmfi} + \alpha_3 X_{hm} + \alpha_4 X_{hf} + \alpha_5 X_h + \vartheta_h \quad (3)$$

In the n+ sample, the indicator variable Twin_h is equal to 1 if the nth birth is a multiple birth and 0 otherwise.[12] The X vectors include control variables as previously defined. ε_{hmfi}, ϑ_h and ω_{hmfi} are the error terms.

In our results section, we scrutinize the exclusion restriction of our instrument, relying on insights of Altonji et al. (2005), Conley et al. (2012) and Bhalotra and Clarke (2016). In a series of nine robustness checks (cf. infra), we modify this empirical framework in several ways, using alternative samples, changing the decision unit (from household head to wife/wives) and modifying the definition of our key variables.

In all cases, we report heteroscedasticity-robust statistics and the usual post-estimation tests (under-identification test, weak identification test and overidentification test).

3 Data

In our empirical analysis, we rely on all DHS rounds implemented in sub-Saharan African countries in the period 1990–2014, for which we could construct the main variables.[13] In our baseline approach, we consider 59 survey rounds for which information on the ethnicity of the mother is available (Appendix Table 11 gives an overview of these survey rounds by country and year) and restrict the sample to children whose siblings of schooling age (6–18) all reside within the household.[14] This gives a dataset of 456,068 siblings of schooling age (6–18), to which we will refer as 'sample I'.

[11] The results are robust to clustering at the ethnicity or region level. The number of countries is insufficient to get a robust covariance matrix when clustering at the country level.

[12] For instance, in the 3+ sample, the indicator variable Twin_h is equal to 1 if the third birth is a multiple birth and 0 otherwise.

[13] In the period 1990–2014, 86 DHS rounds were administered in 34 sub-Saharan African countries. For the following surveys, we could not construct and merge the key variables because of issues with the unique identifiers of surveyed individuals: 1991 and 1998 in Cameroon, 1996–1997 in Chad, 1998–1999 in Cote d'Ivoire, 2001 in Mali and 1992 in Niger.

[14] A household is defined as one person or a group of persons who usually live and eat together. This is not the same as a family. A family includes people who are related, but a household includes any people who live together, whether or not they are related (*DHS Interviewer's Manual*, October 2012).

We focus on the educational attainments of 64,339 firstborn (2+ sample), 99,875 first- and second-born children (3+ sample) and 101,848 first-, second- and third-born children (4+ sample) in the age group 6 to 18, who live with their parents. The lower limit of 6 is the age at which many children in SSA start primary schooling. The upper limit of 18 is the age of secondary schooling completion, provided a swift grade progression. We do not extend the upper limit beyond 18 because post-secondary education in SSA still faces important supply side constraints, and because a considerable proportion of children above 18 live outside the household such that their schooling attainment goes unrecorded in the DHS rounds.

To capture the educational attainment of these children, we look at their completed years of schooling at the time of the survey and then construct age-standardized education z-score with children of the same country and birth cohort as reference group. Our explanatory variable of interest is the number of children in a household. In our baseline specification, we define this variable as the total number of the household head's sons and daughters residing in the household.

Our instrumental variables are the birth of twins at the second, third and fourth birth order. In the DHS birth records, there is a specific variable indicating whether a child is part of a twinship or not. To determine birth order, we consider all children of the household head including those who do not have their mother in the household.

The summary statistics of the principal variables in our analysis can be consulted in Table 1. The summary statistics of the other variables can be consulted in Table 12 of the Appendix.

4 Baseline results

4.1 Baseline OLS and IV estimation

The estimation of the OLS model (Eq. (1)) yields a negative and significant relation between family size and children's schooling. As shown in Table 2, the estimated coefficient is rather small in magnitude: one additional child is associated with a reduction of about 0.023 units of the z-score which corresponds to 0.057 years of schooling for a child of 10 years of age.[15] As this result does not isolate a causal link between our variables of interest, we turn to our IV estimations.

The estimates of the first stage (Eq. (3)) are shown in the second panel of Table 3. Unsurprisingly, they indicate that twinning increases average family size. The effect of twinning ranges from an additional family size of 0.356 (in the 4+ sample) to 0.509 (in the 3+ sample). The coefficients are precisely estimated, significant at 1% and similar to the range of coefficients found in previous research.[16] In all cases, the twin instrument has reassuring first-stage post-estimation statistics, with the Cragg-Donald Wald *F* statistic well above 100.

[15] The coefficients are multiplied by the standard deviation of completed years of education in sample I (see Table 1) and interpreted with reference to the average age of children in sample I.

[16] For instance, Bhalotra and Clarke (2016), Marteleto and de Souza (2012), Angrist et al. (2010) and Black et al. (2005)

Table 1 Sample means and proportions of main variables

Variables	Sample I	2+ sample	3+ sample	4+ sample
Number of children	5.30 (2.54)	3.86 (1.53)	4.39 (1.46)	5.14 (1.43)
Child completed years of schooling	2.05 (2.47)	2.47 (2.67)	2.29 (2.52)	2.31 (2.55)
Child's education z-score	0.00 (1.00)	0.156 (1.04)	0.104 (1.02)	0.086 (1.02)
Child's age	10.09 (3.04)	10.38 (3.11)	10.19 (2.94)	10.44 (2.96)
Child's sex				
Male	*51.76*	*51.03*	*51.36*	*51.68*
Female	*48.24*	*48.97*	*48.64*	*48.32*
Month of birth	6.05 (3.33)	6.05 (3.36)	6.05 (3.34)	6.06 (3.33)
Mother's age	35.39 (6.31)	30.17 (4.91)	30.91 (4.74)	32.05 (4.66)
Mother's education (single years)	2.65 (3.84)	3.61 (4.42)	3.20 (4.18)	2.80 (3.90)
Mother's religion				
Mother is Muslim	*43.49*	38.53	40.33	42.22
Mother is not Muslim	*56.51*	61.47	59.67	(57.78)
Father's age	44.73 (9.43)	38.18 (7.76)	39.23 (7.79)	40.79 (7.88)
Father's education (single years)	4.16 (4.85)	5.22 (5.20)	4.83 (5.08)	4.44 (4.93)
Residence				
Urban	*26.39*	*31.98*	*29.32*	*26.93*
Rural	*73.61*	*68.02*	*70.68*	*73.07*
Wealth quintile				
Poorest	*25.00*	*22.10*	*23.62*	*25.36*
Poor	*21.21*	*19.58*	*20.24*	*21.05*
Middle	*19.77*	*18.76*	*19.38*	*19.56*
Rich	*18.03*	*18.60*	*18.39*	*18.09*
Richest	*16.00*	*20.97*	*18.37*	*15.94*
Twin birth				
Twin at 2nd birth	*1.10*	*1.37*	*1.24*	*1.13*
Twin at 3rd birth	*1.53*	*1.60*	*1.61*	*1.71*
Twin at 4th birth	*1.64*	*1.88*	*1.82*	*1.93*
Twin at 5th birth	*1.89*	*2.28*	*2.16*	*2.12*
N (number of observations)	**456,068**	**64,339**	**99,875**	**101,848**

Source: Authors, based on data from 59 DHS rounds. The twin indicators are expressed in percentages of the total number of households, while the other percentages relate to the total number of children. Sample I includes all children of schooling age (6–18). The "*n*+ sample" is composed of lower birth order children from families with *n* or more children, whose siblings of schooling age (6–18 years) all reside within the household

Sample proportions are in italics and total number of observations in boldface

In contrast to the OLS estimates that are uniformly negative in all three subsamples, the second-stage IV estimates (first panel of Table 3) indicate either no impact (in the 2+ sample and 4+ sample) or a positive and significant effect (at 5%) of family size on education z-scores (in the 3+ sample). In the 3+ sample, a 1-unit increase in predicted family size increases the z-score by 0.097 units on average which is the equivalent of 0.240 years of schooling for a child of 10 years

Table 2 OLS estimates

	Dependent variable: education z-score		
	2+ sample	3+ sample	4+ sample
Number of children	− 0.023*** (0.003)	− 0.024*** (0.003)	− 0.023*** (0.003)
Clusters	20,339	20,117	18,307
Observations	64,339	99,875	101,848
R^2	0.398	0.375	0.363

The 2+ sample includes firstborn from families with two or more children, the 3+ sample includes first and second born from families with three or more children and the 4+ samples comprises first, second and third born from families with four or more children. Control variables include child-level characteristics (sex, age, birth year and month of birth), mother-level characteristics (mother's education, mother's age, mother's age squared, mother's height, mother's ethnicity, her religion, the number of her children who have died and an indicator variables taking one if any of the mother's children died before their first birthday), father-level characteristics (his age and education) and household-level characteristics (residence area and wealth quintile). Standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$

of age. We demonstrate the robustness of these IV results in Section 5 and tentatively explore plausible explanations in Section 6.

4.2 On instrument validity

A concern when implementing the IV estimation is the violation of the exclusion restriction. The exclusion restriction may be threatened because of the presence of confounding factors that affect both twinning and children's schooling. In particular, besides a mother's age, ethnicity and height, also less easily measurable characteristics such as her general health condition may affect the probability of twinning (Smits and Monden 2011). In theory, a mother's health behaviour, in particular smoking and multiple-birth-enhancing fertility treatments, could pose another threat to our instrument's validity, but in practice, this threat is neutralized in SSA due to its prohibitively high (social and monetary) cost (Inhorn 2003).

Table 4 explores the determinants of twinning in our sample. The first column shows the results of a regression of twinning on the mother-level characteristics included in our baseline model: her years of schooling, age, age squared, height, ethnicity, religion, the total number of her children who have died and a dummy variable capturing whether these children died before their first birthday. Out of these eight characteristics, only mother's religion does not significantly affect the probability of twinning at third birth. However, the overall explanatory power of the regression is very low (R^2 of 0.009). The second column adds four additional mother-level regressors: mothers' body mass index (BMI), access to prenatal health care, access to a doctor and access to a nurse.[17] All four additional regressors turn up significant while mother's education becomes non-significant.

[17] To measure access to prenatal care, we use the percentage of births (occurring within 5 years prior to the survey) with prenatal care in the DHS cluster. Access to a doctor or nurse is proxied by the DHS cluster-level percentage of births with prenatal care given by a doctor or nurse.

Table 3 IV estimates of the effect of family size on children schooling

	2+ sample	3+ sample	4+ sample
2nd stage			
Dependent variable: education z-score			
Number of children	0.019 (0.055)	0.097** (0.047)	− 0.060 (0.064)
Clusters	20,339	20,117	18,307
Observations	64,339	99,875	101,848
R^2	0.339	0.297	0.304
1st stage			
Dependent variable: number of children			
Twin at 2nd birth	0.503*** (0.042)		
Twin at 3rd birth		0.509*** (0.040)	
Twin at 4th birth			0.356*** (0.038)
F statistic (excluded instrument)	142.21	160.71	85.41
Under identification test p value	0.000	0.000	0.000
Weak identification test			
Cragg-Donald Wald F (Stock-Yogo critical values)	143.41 (16.38)	259.91 (16.38)	136.89 (16.38)
H_0: equation is weakly identified			

The 2+ sample includes firstborn from families with two or more children, the 3+ sample includes first and second born from families with three or more children and the 4+ sample comprises first, second and third born from families with four or more children. Control variables include child-level characteristics (sex, age, birth year and month of birth), mother-level characteristics (mother's education, mother's age, mother's age squared, mother's height, mother's ethnicity, her religion, the number of her children who have died and an indicator variables taking one if any of the mother's children died before their first birthday), father-level characteristics (his age and education) and household-level characteristics (residence area and wealth quintile). Standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$

Based on subsamples of individuals for which all additional four control variables are available, Table 5 shows how the estimated coefficient of interest changes when adding the additional four control variables to our baseline controls. Using the baseline specification, the estimates are 0.003 in the 2+ sample (panel A, column I), 0.137** in the 3+ sample (panel B, column I) and − 0.043 in the 4+ sample (panel C, column I). When adding the four additional mother-level controls, our coefficient of interest only slightly decreases in all three panels, to − 0.005, 0.126** and − 0.061, respectively. Thus, without the additional controls, our estimates are (slightly) biased upward, which is expected if the included mother characteristics positively correlate with both twinning and children's schooling. In the spirit of Altonji et al. (2005), we however argue that, if observed additional controls change the value of our estimated coefficients only to such a small extent, it is unlikely that there exist unobserved controls that would turn our results upside down.[18]

[18] We do not formally verify this intuition, since the formalized extension by Oster (2017) only applies to OLS estimates.

Table 4 Determinants of twinning

	Dependent variable: probability of twinning	
	I With H + E	II With H + E + AC
Mother's age	0.223*** (0.062)	0.208*** (0.062)
Mother's age squared	− 0.002** (0.001)	− 0.002* (0.001)
Mother completed years of education	0.037** (0.015)	0.012 (0.015)
Mother's height	0.005*** (0.001)	0.005*** (0.001)
Number of dead children	0.478*** (0.056)	0.497*** (0.056)
Infant death	0.453*** (0.127)	0.452*** (0.127)
Mother is Muslim	− 0.136 (0.145)	− 0.153 (0.145)
Mother BMI	–	0.000* (0.000)
Access to prenatal care	–	0.450* (0.234)
Access to a doctor	–	0.613* (0.326)
Access to a nurse	–	0.454** (0.199)
Clusters	121,597	121,597
Observations	250,837	250,837
R^2	0.009	0.009

This table shows linear probability estimates of twinning. The sample includes all children in sample I. H + E stands for mother's health characteristics and her ethnicity for which we control for in our baseline specification: her years of education, age, age squared, height, ethnicity, religion, number of her children who have died and whether children died before their first birthday. AC stands for four additional controls: mother' BMI, access to prenatal health care, access to a doctor and access to a nurse. We additionally control for mother's ethnicity and the year in which the birth occurred. Robust standard errors are clustered at the mother level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

To further safeguard our results, we follow Conley et al. (2012) and Bhalotra and Clarke (2016) in deriving bounds for our coefficient of interest using the 'plausexog' command in Stata. To do so, we first acquire insight into the direct effect of twinning on children's schooling (γ) by simply comparing education z-scores of children from twin mothers to those of children from non-twin mothers, controlling for the set variables in our baseline specification (see Table 13 in the Appendix). We then take the upper value of the estimated 95% confidence interval which is 0.005.[19] The standard deviation of γ (sd = 0.007) is obtained from a 100-replication bootstrap estimation of γ. Table 6 shows our bounds' estimates of the family size effect on children's schooling using Conley's union of confidence interval (UCI) and local to zero (LTZ) approaches. The results point to a rejection of a quantity-quality trade-off in all three samples and a confirmation of the positive effect of family size on children schooling in the 3+ sample.

Another potential violation of the exclusion restriction stems from parental behaviour towards twins due to twins' lower average birth weight and the potential

[19] Note that our prior value on γ is more conservative than the 0.004 used in Bhalotra and Clarke (2016). Our results hold even with γ set at 0.008.

Table 5 Stability of the estimated coefficient of interest and post-estimation statistics when expanding the set of mother-level control variables

	Dependent variable: education z-score	
	I With H + E	II With H + E + AC
Panel A: 2+ sample		
Number of children	0.003 (0.063)	− 0.005 (0.063)
Clusters	16,935	16,935
Observations	42,789	42,789
R^2	0.345	0.337
Cragg-Donald Wald F (Stock-Yogo critical values)	102.09 (16.38)	103.92 (16.38)
Panel B: 3+ sample		
Number of children	0.137** (0.056)	0.126** (0.054)
Clusters	16,940	16,940
Observations	66,654	66,654
R^2	0.273	0.287
Cragg-Donald Wald F (Stock-Yogo critical values)	178.99 (16.38)	182.76 (16.38)
Panel C: 4+ sample		
Number of children	− 0.043 (0.073)	− 0.061 (0.071)
Clusters	14,996	14,996
Observations	68,144	68,144
R^2	0.304	0.309
Cragg-Donald Wald F (Stock-Yogo critical values)	93.97 (16.38)	97.56 (16.38)

H + E stands for mother's health characteristics and her ethnicity for which we control in our baseline specification: her years of education, age, age squared, height, ethnicity, religion, number of her children who have died and whether children died before their first birthday. AC stands for five additional controls: mothers' BMI, access to prenatal health care, access to a doctor and access to a nurse. In column I, we control H + E in column II and H + E and AC in column III. Robust standard errors are clustered at the DHS cluster level and reported in parentheses. The 2+ sample, 3+ sample and 4+ sample are here restricted to subsamples for which the four additional control variables (AC) are available

**$p < 0.05$

consequences for their health or cognitive achievements later on (Rosenzweig and Zhang 2009). If the future earning potential of twins is thought to be lower, parents may divert resources from twins to singletons. Rosenzweig and Zhang (2009) suggest to include twins in the analytical sample and to include birth weight in the regression in order to control for this potential bias. This could indeed be a straightforward solution, where it is not that the DHS only includes birth weight for the under-5-year-old children. As a second best, we verify whether this 'diversion of resources' hypothesis is supported in the sample of under-5-year-old children.

Using information on 70,902 under-5-year-old children in our 59 DHS waves, we regress *birth* weight (in log terms) on a twin indicator and find that twins indeed have a significantly lower average birth weight than singletons (see column

Table 6 Bounds' estimates of family size effect on children schooling using Conley's UCI and LTZ approaches

	Union of confidence interval (UCI)		Local to zero (LTZ)		
	Lower bound	Upper bound	Lower bound	Coefficient	Upper bound
2+ sample	−0.108	0.127	−0.109	0.020	0.149
3+ sample	−0.002	0.176	0.017	0.119**	0.221
4+ sample	−0.196	0.048	−0.162	−0.033	0.095

The bound's estimates are derived using the 'plausexog' command in Stata and are based on the prior that being from a twin family has a direct effect ($\gamma = 0.005$) on educational outcomes (which, for UCI bounds, is more conservative compared to the 0.004 used in Bhalotra and Clarke 2016 for developing countries). The UCI bounds are derived based on $\text{gamma}_{\text{min}} = 0.000$ and $\text{gamma}_{\text{max}} = 0.010$ while the LTZ bounds are derived based on $\gamma = 0.005$ with a sd of 0.007 (the sd results from 100-replication bootstrap estimations, and we perform the test of normal distribution of γ; see details in Table 14 in the Appendix). Since the LTZ approach does not allow for factor variables, we exclude mother's ethnicity from the equation and all other variables enter as continuous variables. When applied to the baseline specification, this does not qualitatively change our results

$^{**}p < 0.05$

I of Table 14 in the Appendix).[20] When looking at the weight (column II) and body mass index percentile (column III) of under-5-year-old children *at the time of the survey*, we still find that twins have significantly lower weight and BMI than singletons.[21] However, the estimated coefficient is smaller, indicating that the gap has become smaller over time. In fact, when looking at the BMI, twins and singletons belong to the same decile on average.[22] The closing of the gap suggests that parents do not divert resources away from twins. Furthermore, controlling for the entire set of age dummies, wealth quintiles, region of residence, child sex, parental education and ethnicity fixed effect, we find that twins enjoy as much education as singletons (column IV of Table 14 in the Appendix).

4.3 Heterogeneity

As mentioned in the introduction, the studies that have set out to identify the causal relationship between family size and schooling have produced mixed results, suggesting that the relation is context dependent. Hence, we explore the heterogeneity of our result, by running separate regressions across subsamples with respect to gender, poverty status and region.

[20] We control for several characteristics likely to affect birth weight such as a mother's age, education, body mass, age at first birth and ethnicity; the child's sex, preceding birth interval and year of birth; and the household's residence and wealth quintile. All error terms are clustered at the DHS cluster level.

[21] We regress BMI and BMI percentile on the twin indicator, controlling for the factors mentioned in the previous footnote, as well as for birth weight, child age, the number of months of breastfeeding, the number of under-5-year-old children in the households and whether or not the mother lives with her husband.

[22] The difference in BMI is measured in terms of percentiles. A child is considered underweighted if its BMI is below the 10th percentile of the World Health Organization's reference BMI distribution (see Cole et al. 2007). According to this definition, 13.75% of twins are underweighted compared to only 10% of singletons. But, as shown in Table 3, on average, twins and singletons belong to the same decile.

To explore regional variation in the estimated relation, we contrast West and Central Africa with East and Southern Africa. This division is inspired by the regional clustering of TFR, which is, on average, relatively high in West and Central Africa (5.09) and lower in East and Southern Africa (4.45). To compare across poor and non-poor, we define households in the first and the second asset ownership quintiles as poor and those in the fourth and the fifth quintiles as non-poor. We discard the third quintile to achieve a sharper contrast between poor and non-poor. The results are lined up in Table 7. In our discussion below, we highlight the various significant coefficient estimates.

When only looking at the schooling of boys, we find a sizeable and significantly positive coefficient estimate in the 3+ sample (0.122**). For girls, the estimated coefficient is found to be negative and (slightly) significant (−0.183*) in the 4+ sample. In the subsample of poor households, the estimated coefficient is insignificant across the board. For the subsample of non-poor, we observe a significantly positive estimate in the 3+ sample (0.139***). In the regional subsamples, the estimated family size coefficient appears to be only slightly significant (and positive) in the 3+ sample in West and Central Africa (0.108*). It remains non-significant elsewhere.

When comparing the effect of family size on schooling across countries with persistently high fertility[23] and countries with either a low fertility or a downward fertility trend[24], we find a non-significant effect in low-fertility countries in all the 2+, 3+ and 4+ samples, but a positive and slightly significant (of 0.128*) effect in the 3+ sample of high-fertility countries.

Finally, we explore whether the relation between family size and schooling has changed over time. To do so, we focus on the year 2000, which marked the adoption of the Education For All initiative by 189 countries.[25] We compare the effect of family size on schooling across children born prior to 2000 and children born from 2000 onwards. In none of the samples, we find a differential effect of family size over time, but our results indicate that the association of parents' education, gender and residence area with children's educational outcomes has weakened considerably over time, suggesting a democratization of schooling (see Table 8).

In sum, in our various subsample analyses, the zero result is confirmed for the 2+ sample and the 4+ sample (with the notable exception of the subsample of girls in the 4+ sample[26]), while the positive result for the 3+ sample is shown to be mainly driven by children from non-poor families, living in high-fertility countries. We will come back to this latter result in Section 6. Now, we first discuss a series of robustness checks.

[23] Burkina Faso, Burundi, Congo DR, Guinea, Malawi, Mali, Niger, Nigeria and Chad (see Lesthaeghe 2014, p. 2)

[24] Ghana, Lesotho, Liberia, Madagascar, Namibia, Rwanda, Senegal, South Africa, Swaziland and Uganda (see Lesthaeghe 2014, p. 2)

[25] The initiative aimed at a global commitment to provide quality basic education for all children, youth and adults and was first launched in 1990 (http://www.worldbank.org/en/topic/education/brief/education-for-all). Admittedly, children born in 2000 reached schooling age only in 2006. On the other hand, the initiative's adoption itself is unlikely to have had an immediate impact on the ground. In any case, any cut-off year would be somewhat arbitrary. Choosing slightly different cut-off years (1998 and 1999 which give almost balanced samples across cohorts) yields similar results.

[26] Family size might have a different impact on girls' schooling (no impact or negative impact) because of time spent in household chores by older daughters (Mueller 1984; Tiefenthaler,1997) which might negatively affect their educational achievements.

Table 7 Subsample analyses with respect to gender, asset wealth and region

	Dependent variable: education z-score							
	Only boys	Only girls	Only children from poor families	Only children from non-poor families	Only children living in West and Central Africa	Only children living in East and Southern Africa	High fertility	Low Fertility
Panel A: 2+ sample								
Number of children	0.042 (0.077)	−0.023 (0.079)	−0.094 (0.072)	0.020 (0.095)	0.050 (0.076)	−0.059 (0.066)	0.026 (0.103)	0.244 (0.225)
Clusters	15,963	15,825	11,080	11,287	13,549	6790	7901	2924
Observations	32,831	31,508	26,814	25,458	44,276	20,063	30,767	6960
R^2	0.325	0.354	0.150	0.385	0.326	0.387	0.360	0.255
Cragg-Donald Wald F	72.25	69.98	57.60	60.97	71.19	109.35	39.38	15.07
Panel B: 3+ sample								
Number of children	0.122** (0.060)	0.064 (0.065)	0.060 (0.085)	0.139** (0.061)	0.108* (0.060)	0.075 (0.072)	0.128* (0.067)	0.058 (0.140)
Clusters	17,404	17,240	11,286	11,020	13,554	6563	7934	10,597
Observations	51,297	48,578	43,802	36,717	69,185	30,690	48,672	2847
R^2	0.282	0.315	0.150	0.333	0.286	0.347	0.306	0.326
Cragg-Donald Wald F	135.96	120.22	59.51	175.07	136.81	175.96	110.79	28.75
Panel C: 4+ sample								
Number of children	0.031 (0.072)	−0.183* (0.102)	−0.076 (0.108)	−0.132 (0.087)	−0.066 (0.084)	−0.046 (0.088)	−0.041 (0.101)	0.107 (0.283)
Clusters	15,994	15,900	10,314	9372	12,518	5789	7546	2039
Observations	52,637	49,211	47,269	34,662	71,406	30,442	50,298	3886
R^2	0.298	0.273	0.146	0.337	0.306	0.331	0.338	0.290
Cragg-Donald Wald F	95.68	120.22	33.06	83.93	67.14	124.04	45.44	28.75

The control variables are as specified in the note of Table 3. Robust standard errors are clustered at the DHS cluster level and reported in parentheses. Low- and high-fertility countries are as defined in Lesthaeghe (2014). The Stock-Yogo critical value is 16.38 in all columns

**$p < 0.05$; *$p < 0.1$

Table 8 Effect of family size across older and younger cohorts

	Dependent variable: education z-score		
	2+ sample	3+ sample	4+ sample
Number of children	− 0.067 (0.074)	0.079 (0.070)	− 0.138 (0.091)
Number of children × young cohort	0.222 (0.170)	0.030 (0.120)	0.152 (0.149)
Female	− 0.018*** (0.008)	− 0.029*** (0.007)	− 0.040*** (0.007)
Female × young cohort	0.042*** (0.014)	0.045*** (0.011)	0.028*** (0.011)
Mother's completed years of education	0.036*** (0.004)	0.040*** (0.003)	0.032*** (0.004)
Mother's completed years of education × young cohort	− 0.003 (0.006)	− 0.007* (0.004)	− 0.003 (0.006)
Father's completed years of education	0.032*** (0.001)	0.034*** (0.001)	0.033*** (0.002)
Father's completed years of education × young cohort	− 0.004* (0.002)	− 0.005* (0.002)	− 0.002 (0.003)
Rural	− 0.164*** (0.021)	− 0.206*** (0.017)	− 0.175*** (0.021)
Rural × young cohort	0.081*** (0.025)	0.091*** (0.019)	0.076*** (0.020)
Wealth quintile	0.116*** (0.007)	0.122*** (0.005)	0.124*** (0.005)
Wealth quintile × young cohort	0.048** (0.020)	0.023** (0.012)	0.026*** (0.010)
Clusters	20,339	20,117	18,307
Observations	64.339	99,875	101,848
R^2	0.328	0.301	0.292
Cragg-Donald Wald F (Stock-Yogo critical values)	31.40 (7.03)	58.73 (7.03)	33.02 (7.03)

This table shows the effect of the number of children and other relevant variables on children's schooling for children belonging to a young (born after 2000) and an older cohort (born before 2000). Control variables include child-level characteristics (age, birth year and month of birth), mother-level characteristics (mother's age, mother's age squared, mother's height, mother's ethnicity, her religion, the number of her children who have died and an indicator variables taking one if any of the mother's children died before their first birthday) and father's age. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

5 Robustness checks

We check the robustness of our results in ten different ways. Table 9 gives a line-up of the estimated coefficients on family size for each check in the 2+, 3+ and 4+ samples. The full results are reported in Tables 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 of the Appendix, for the total sample, as well as for the poor and non-poor subsamples.

In the first robustness check, we follow Angrist et al. (2010) in allowing for heterogeneity across subsamples in the predictive power of twin birth in the first stage. We do so by including interaction terms in our first-stage regressions between twin birth and a set of indicator variables (rural, mother is Muslim, rural West and Central Africa), thus sequentially adding the following regressors: $\text{Twin}_h \times \text{Rural}_h$, $\text{Twin}_h \times \text{Mother\_Islam}_m$ and $\text{Twin}_h \times$ Rural West and Central Africa_h. The rationale for including these regressors is that fertility is higher in rural areas (compared to urban areas), in Muslim families and in West and Central Africa (compared to East and Southern Africa), and the twin instrument tends to perform less well in larger families (Angrist et al. 2010).[27] When including all three interaction terms, the zero result remains in the 2+ and 4+ samples, while we still find a positive and significant in the 3+ sample.

In the second robustness check, we use an unrestricted specification with the 'partially missing instruments' method as described in Mogstad and Wiswall (2016), pp. 174–175). The partially missing instruments are constructed based on a polynomial function of mother's age, mother's education, father's age and father's education, controlling for child characteristics, household characteristics, mother's ethnicity and her health conditions. This change in specification does not alter our results: family size remains insignificant in the 2+ and 4+ samples and significantly positive in the 3+ sample, in particular when captured by the indicator variable 'more than 3 children' that apprehends the marginal effect of moving from three to four children.

Third, instead of restricting the sample to those children that are part of households where all children reside within the household, we expand the sample to include also children that live in households where one or more school-aged siblings reside outside the household. This approach yields a 2+ sample of 68,259 children, a 3+ sample of 112,285 children and a 4+ sample of 119,544 children. We find that the coefficient on family size is still positive and slightly significant in the 3+ sample and positive and non-significant in the 2+ sample but turns negative and slightly significant in the 4+ sample.

Fourth, recognizing the complexity of SSA households, we change the decision unit. Among the 525,646 children in our sample I, we count 165,418 living in polygamous households. While our baseline approach considers the household head as the unit of decision making, in this robustness check, we assume decisions to be taken at the level of each mother. In this decentralized approach, the number of children is defined for each of the household head's wives as her total number of sons and daughters living in the household. Birth order is also defined at the level of the mothers. Doing so, we find that family size loses its significance in the 3+ sample and remains so in the 2+ and 4+ samples.

In checks 5, 6 and 7, we use alternative definitions of our key variables. Instead of defining the number of children as the sons and daughters of the household head, we define them as the total number of births given by the household head's wives (unless the household head is female). To reduce measurement error in our schooling variable (for instance, parents reporting years in kindergarten as schooling), education z-scores are based on censored years of education.[28] And, estimates using completed grades (years of schooling) rather than educational z-scores are provided in robustness check 7. Our findings remain similar in all three cases: no significant effect in the 2+ and 4+ samples and a positive and significant effect in the 3+ sample.

In robustness checks 8 and 9, we use region of residence of the household and country-by-urban/rural fixed effects instead of mother's ethnicity fixed effects to take into account DHS rounds in which ethnicity is not included (e.g. Rwanda,

[27] Also following Angrist et al. (2010), we try out a second instrument, i.e. sibship sex composition. This alternative instrument relies on the idea that parents prefer to have both boys and girls rather than only children of the same sex. Hence, in the latter case, they may be more likely to have an additional child. However, in our case, only twinning performs well in the first stage. For instance, where the F-statistic surpasses 100 for twinning, it barely reaches 20 for the sex composition instrument (results not reported, but available on request). We therefore focus on twinning in our analysis.

[28] We censor the completed years of education to the child's age minus 6, assuming that schooling can start only from 6 years onwards since children are admitted to primary school at 6 in most of the countries (from http://data.worldbank.org/indicator). For instance, if a 12-year-old child in our sample is reported to have completed 9 years of schooling, we set years of schooling to 6.

Burundi).[29] In the former case, the estimated coefficient loses significance in the 3+ sample. Apart from that, the results remain qualitatively similar: a positive and significant coefficient in the 3+ sample and non-significant ones in the 2+ sample and the 4+ sample.

Finally, we use an alternative definition to discriminate between poor and non-poor families, defining the poverty line as the average value of the wealth index in each DHS round. This last approach confirms the positive effect observed only in non-poor families of the 3+ sample and the absence of effect in both poor and non-poor families across the 2+ and 4+ samples.

Overall, our results remain fairly robust in all three subsamples. The zero result in the 2+ sample remains so across all ten robustness checks while it becomes significantly negative in the 4+ sample in only one case (inclusion of households in which some school-aged children reside outside the household). In the 3+ sample, the positive coefficient is no longer significant in only two cases (decentralized approach and region of residence of the household instead of mother's ethnicity fixed effect). When running the robustness checks on the non-poor sample, the estimated coefficient on family size in the 3+ sample remains positive and significant across the board.[30]

Taken together, these results bolster the case against a quantity-quality trade-off in SSA, when quality is measured as educational attainment. At the same time, however, we note the heterogeneity of coefficient estimates, not only across gender, asset wealth and region but also across the 2+, 3+ and 4+ samples. Whether or not this heterogeneity is a statistical artefact needs to be determined by future work.

6 Mechanisms

To further guide future work, we fully exploit the DHS data, in order to provide some cues for the possible mechanisms underlying our results and their heterogeneity.

First, we use the DHS data to distinguish between three proximate causes that underlie differences in educational attainment: school enrolment, school starting age and dropout. We explore these proximate causes in all three analytical samples, by estimating Eqs. (2) and (3) with school enrolment, school starting age and dropout as dependent variables. Table 10 lines up the coefficients of interest. The full results are given in Tables 25, 26, 27, 28, 29, 30, 31 of the Appendix.

When using the total sample and the subsample of poor households, we find a zero effect of family size on enrolment, dropout and school starting age in the 2+, 3+ and 4+ samples, with one exception (in the total 3+ sample, the dropout of second born is slightly reduced). For the subsample of non-poor households, we find various significant coefficients. In the 2+ sample, firstborn's school starting age is reduced (− 0.499*) with an exogenous increase in family size. In the 3+ sample, they are second born that seem to start school earlier on. A closer examination of this effect reveals that it is largest and significant when the second born is relatively close in age with the firstborn

[29] This yields a more balanced sample between West and Central Africa (57.80%) and East and Southern Africa (42.20%).

[30] In the case of the 4+ sample, we obtain a negative and slightly significant (at 10%) coefficient in the subsample of non-poor households when using censored education z-scores (panel C of Table 20) and when using the household's region of residence instead of mother's ethnicity (panel C of Table 22).

Table 9 Summary of robustness checks

Robustness check	Table	Description	Coefficient on number of children		
			2+ sample	3+ sample	4+ sample
1	Table 15	Heterogeneity of twin at 3rd birth			
		Interaction with Rural	0.019 (0.056)	0.098** (0.047)	−0.061 (0.064)
		Interaction with Mother is Muslim	0.008 (0.054)	0.092** (0.046)	−0.047 (0.063)
		Interaction with rural West and Central Africa	0.018 (0.054)	0.094** (0.047)	−0.064 (0.062)
2	Table 16	Partially missing instruments with non-linear specification			
		Number of children > 2	−0.050 (0.264)		
		Number of children > 3	0.156 (0.122)	0.275** (0.127)	
		Number of children > 4	−0.083 (0.140)	−0.104 (0.115)	−0.110 (0.120)
		Number of children > 5	0.021 (0.166)	0.023 (.129)	0.043 (0.116)
3	Table 17	Expand sample including households in which some school-aged children reside outside the household	0.029 (0.056)	0.091* (0.048)	−0.114* (0.070)
4	Table 18	Decentralized approach (mother as decision unit)	0.001 (0.066)	0.059 (0.042)	0.003 (0.044)
5	Table 19	Alternative definition of the number of children (as the total number of births given by the household head's wives)	0.018 (0.051)	0.092* (0.045)	−0.055 (0.059)
6	Table 20	Education z-score calculated with censored completed years of education (to the child's age minus 6 at most)	0.050 (0.045)	0.113** (0.041)	−0.054 (0.057)
7	Table 21	Completed grade (years of schooling) as outcome variable	0.075 (0.095)	0.144* (0.079)	−0.031 (0.111)
8	Table 22	Region of residence of the household instead of mother's ethnicity FE	0.030 (0.048)	0.063 (0.041)	−0.080 (0.054)
9	Table 23	Country-by-urban/rural instead of mother's ethnicity FE	0.040 (0.048)	0.097** (0.044)	−0.054 (0.061)
10	Table 24	Alternative definition to discriminate between poor and non-poor families			
		Poor	−0.030 (0.071)	0.034 (0.073)	−0.072 (0.088)
		Non-poor	0.042 (0.087)	0.138** (0.059)	−0.099 (0.093)

Detailed tables of robustness checks are provided in Tables 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 in the Appendix

$^{**}p < 0.05$; $^{*}p < 0.1$

Table 10 Exploring the proximate cause underlying the positive effect of family size on schooling

	I Enrolment	II Starting age	III Starting age when a difference in age between 1st and 2nd born was ≤ 3 years	IV Dropout
Panel A: 2+ sample				
Effect on 1st born (all)	−0.009 (0.023)	−0.172 (0.149)	−0.214 (0.245)	−0.024 (0.026)
Effect on 1st born (poor)	0.000 (0.039)	0.134 (0.169)	0.192 (0.182)	−0.058 (0.057)
Effect on 1st born (non-poor)	−0.028 (0.023)	−0.499* (0.260)	−0.965* (0.528)	0.008 (0.039)
Panel B: 3+ sample				
Effect on 1st born (all)	0.001 (0.022)	−0.196 (0.139)	−0.048 (0.181)	−0.014 (0.030)
Effect on 1st born (poor)	−0.019 (0.047)	−0.039 (0.404)	0.550 (0.534)	0.065 (0.092)
Effect on 1st born (non-poor)	0.030 (0.023)	−0.128 (0.154)	−0.179 (0.210)	−0.017 (0.030)
Effect on 2nd born (all)	0.037 (0.024)	−0.018 (0.180)	−0.222 (0.208)	0.063* (0.035)
Effect on 2nd born (poor)	−0.010 (0.050)	0.524 (0.405)	0.691 (0.490)	0.062 (0.081)
Effect on 2nd born (non-poor)	0.023 (0.027)	−0.285 (0.180)	−0.623** (0.275)	0.060 (0.040)
Panel C: 4+ sample				
Effect on 2nd born (all)	−0.010 (0.050)	0.524 (0.405)	0.691 (0.490)	0.062 (0.081)
Effect on 2nd born (poor)	0.034 (0.063)	−0.462 (0.291)	−0.248 (0.326)	−0.008 (0.085)
Effect on 2nd born (non-poor)	−0.072* (0.042)	0.342 (0.441)	0.469 (0.435)	−0.048 (0.047)
Effect on 3rd born (all)	−0.004 (0.035)	−0.204 (0.234)	−0.109 (0.177)	−0.047 (0.040)
Effect on 3rd born (poor)	−0.027 (0.068)	0.240 (0.276)	−0.147 (0.213)	0.016 (0.087)
Effect on 3rd born (non-poor)	0.020 (0.046)	−0.557 (0.429)	−0.159 (0.380)	−0.076* (0.040)

The control variables are as specified in the note of Table 3. Robust standard errors are clustered at the DHS cluster level and reported in parentheses. In columns II to V, we consider only children enrolled in school at the time of the survey. In columns IV and V, the school starting age is generated based on censored years of education (see robustness check 5). Following Angrist and Pischke (2009), pp. 102–107), we use manual 2SLS as a linear probability model in columns I and VI, although "enrolment" and "dropout" are binary variables. In panel C, the age difference in columns III and V is between the 2nd and the 3rd born. Full results are reported in Tables 25, 26, 27, 28, 29, 30, 31 in the Appendix

$^{**}p < 0.05$; $^{*}p < 0.1$

(3 years apart or less[31]) (see column III of Table 10). In the 4+ sample, we find that family size reduces the probability of enrolment of the second born (− 0.072*). For the third born, results show a reduction in the probability of dropout (− 0.076*).[32]

Overall, this tentative exploration of the proximate causes suggests that, in response to an exogenous increase in family size at birth order n, relatively small and wealthy households tend to send the $n-1$th child earlier to school, a finding that is not replicated in poor households. This could indicate that, when financially possible, some households opt to speed the schooling of earlier-born children upon twin birth. Whether this finding can be broadly replicated, and whether it is explained by an attempt on the part of parents to maximize economies of scale[33] or simply to relieve the caregiver so he/she can focus on younger siblings, is a question left for future research. The non-linearity of economies of scale (see e.g. Holmes and Tiefenthaler 1997; Tiefenthaler 1997) together with the negative effect of reduced care time on children cognitive abilities (Lehmann et al. 2018) may account for the heterogeneity across the 2+, 3+ and 4+ samples.

Next to economies of scale, the introduction of the article mentioned three other mechanisms that could explain the absence of a quantity-quality trade-off or even a positive effect of family size on schooling: child labour, the chain arrangement and support from the extended family. While we do not have the data to explore the likelihood of the latter two mechanism, we tentatively discuss (and dismiss) the role of child labour.

Child labour, both at home and in the labour market,[34] still is common in many sub-Saharan African countries, but the group of children that are working is increasingly made up of children who combine part-time employment and schooling (Guarcello et al. 2015). The combination of work and schooling may allow for child labour to contribute to schooling rather than crowd it out, by providing resources for schooling fees, their own as well as those of their siblings. Should child labour explain the positive effect of family size in the 3+ sample, we would however expect the effect to be larger in poor households and lower in non-poor ones, given that the latter rely less on resources provided by children. Instead, we find the reverse. To test more formally for the child labour mechanism, we exploit the available child labour information in a subset of the DHS rounds.[35] If child labour contributes to schooling and explains the positive effect (or zero effect) of family size, the estimated coefficient on family size would be reduced after controlling for child labour in our model. Results in Table 32 of the Appendix show that, rather than attenuating the positive effect of family size, the inclusion of child labour (both own labour and the labour of his/her siblings) slightly reinforces the positive effect in the 3+ sample while it leaves the estimated coefficients in the 2+ sample and the 4+ sample almost unaltered.

[31] This result remains significant but is lower when considering an age difference of ≤ 4 years, but it loses significance when the age difference is 2 years or less, both in the 3+ and 2+ samples.

[32] Enrolment, school starting age and dropout of firstborn in the 4+ sample are not affected by family size (results not reported).

[33] When sharing books, tutoring or transporting to school, the per-child cost of schooling declines when two or more children can be sent to school simultaneously (Qian 2009).

[34] In our analysis, child labour is broadly defined as the sum of time spent in household chores, the number of working hours for a family member (including himself/herself) or for someone outside the household. In an alternative definition, we exclude household chores; our findings remain robust to this change.

[35] In 21 surveys of 15 countries, information on children's time allocation is provided (the rounds and countries are listed in Table 11 in the Appendix). The average total of hours worked by a child, including time spent on household chores, is 9.78 per week and is significantly larger in poor families in a sample mean t test (11.73 h) compared to 7.59 in richer ones.

7 Conclusion

The aim of this study was to test the quantity-quality trade-off in SSA, focusing on children's schooling. To do so, we investigated how an increase in family size affects schooling using twin birth as an instrumental variable to deal with endogeneity issues. Overall, we find no significant effect of family size on children's schooling, thus casting doubt on the generally assumed negative relation between family size and schooling.

In the subsample of first and second born from relatively rich households with three or more children, we find a positive effect of family size on schooling and this effect survives various robustness checks. To exclude that this result is a statistical artifact, its replication in other samples is required.

In a tentative exploration of the underlying mechanisms, it emerges that upon a fertility shock at the nth birth, relatively small and rich families tend to send their $n-1$th born to school relatively early on. Doing so may optimize economies of scale in schooling and/or relieve the caregiver, although both the replicability of the finding and its explanation need further study.

Exploring the heterogeneity of our results, we find that the effect of family size does not vary substantially across time. This in sharp contrast to the role of other factors such as parental education, gender and residence area (urban/rural) that significantly decreased over time, in line with the 'democratization' of education. Only in the 3+ sample we find a regional difference: family size positively impacts children's schooling in countries with persistently high fertility while we find no effect in countries with low or declining fertility.

Our research suffers from a number of limitations. First, the positive impact of family size on schooling of first and second born children in the 3+ sample cannot readily be generalized to higher-parity children, which is clear from our results in the 4+ sample. Second, the DHS provides only a snapshot in time of children whose mothers are of childbearing age (15–49 years). The number of children observed as well as their schooling attainment reflect therefore only an intermediate situation, not the final one. This leaves open the possibility that, in the longer run, the positive effect of family size on schooling in non-poor households with three or more children (driven by early enrolment of the second born) fades away. Third, the available data are not well suited to distinguish between competing underlying mechanisms to the differential effect of family size across samples. The short-term horizon does not allow for explicitly testing of the chain arrangement. And, lacking more detailed data on household consumption and transfers, we cannot thoroughly test for the economies of scale and extended family mechanisms. Finally, the number of years of schooling is only one way in which parents can invest in their children. Important omitted dimensions include the quality of schooling and health care.

These gaps should be filled by future research, relying on other types of data such as pooled census data in which families and their split-offs are traced over time and micro-economic surveys that provide detailed information on household members' consumption of schooling inputs and their time allocation, as well as surveys that include more information on health and the quality of schooling of a children in various age cohorts. A more open line of questioning, in qualitative research, could also reveal the reasoning underlying parent's decision making.

Acknowledgements We received much appreciated comments from Elena Briones Alonso, Philip H. Ross, Pieter Serneels, NikStoop and the participants at seminars, conferences and workshops in Leuven (LICOS-KU Leuven seminar),Antwerp (Doctoral Day) and Oxford (2017 CSAE conference), as well as from the Editor and two anonymousreferees of this journal.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Table 11 Cross-country summary statistics based on sample I

Country	DHS round	Exploitable data on child labour
Benin	1996	No
	2001	No
	2006	No
	2012	No
Burkina	1993	No
	2003	No
	2010	Yes
Cameroon	2004	No
	2011	Yes
Central African Rep.	1995	No
Chad	2004	Yes
Congo	2012	Yes
Congo Dem.	2007	Yes
	2014	Yes
Cote d'Ivoire	2012	Yes
Ethiopia	2000	No
	2005	No
	2011	Yes
Gabon	2000	No
	2012	Yes
Gambia	2013	No
Ghana	1993	No
	1998	No
	2003	No
	2008	No
	2014	No
Guinea	1999	No
	2005	No
	2012	Yes
Kenya	1993	No
	1998	No

Table 11 (continued)

Country	DHS round	Exploitable data on child labour
	2003	No
	2014	No
Liberia	2013	No
Malawi	2000	Yes
	2004	Yes
	2010	No
Mali	1996	No
	2006	Yes
	2013	Yes
Mozambique	1997	No
	2011	No
Namibia	1992	No
	2000	No
Niger	1998	No
	2006	Yes
Nigeria	2008	No
	2013	No
Senegal	1993	No
	2005	Yes
	2011	Yes
	2013	Yes
	2014	No
Sierra Leone	2008	Yes
	2013	Yes
Togo	1998	No
	2014	Yes
Uganda	2011	No
Zimbabwe	1994	No
All together	59	

Source: Authors, based on the DHS data

Table 12 Summary statistics of variables

Variables	Sample I	2+Sample	3+Sample	4+Sample
Mother height (mm)	1593.42	1591.67	1591.93	1592.59
	(56.98)	(59.30)	(59.11)	(58.81)
Number of dead children of mother	0.84	0.26	0.33	0.38
	(1.21)	(0.61)	(0.66)	(0.72)
At least one of the mother's children died before 1st birthday	32.44	13.65	17.13	19.54
Mother BMI	2394.24	2403.17	2385.13	2370.13
	(1019.26)	(1036.25)	(1006.89)	(979.36)
Access to prenatal care	0.82	0.82	0.81	0.80
	(0.28)	(0.28)	(0.28)	(0.29)
Access to doctor for prenatal care	0.10	0.13	0.12	0.11
	(0.18)	(0.20)	(0.19)	(0.18)
Access to nurse for prenatal care	0.59	0.60	0.59	0.58
	(0.34)	(0.34)	(0.34)	(0.34)
Child school enrolment	66.76	73.17	72.01	71.17
Child dropout	5.39	5.09	5.07	5.58
Child labour				
Less than 8 h/week	63.21	65.25	63.27	62.29
8 to 21 h/week	21.50	20.54	21.67	21.85
22 to 35 h/week	6.85	6.57	6.94	7.06
35 h/week and+	8.44	7.63	8.11	8.80
N (number of observations)	456,068	64,339	99,875	101,848

Source: Authors, based on data from 59 DHS rounds. The twin indicators are expressed in percentages of the total number of households, while the other percentages relate to the total number of children. Sample I includes all children of schooling age (6–18). "N+Sample" is composed of lower birth order children from families with N or more children, whose siblings of schooling age (6–18) all reside within the household

Table 13 Estimate of the direct effect of twinning on children's schooling (γ)

	Prior on γ	95% CI	100 replications Bootstrap estimated Standard-deviation of γ
Twin in the family	−0.013	−0.0296 to 0.0042	0.007
	(0.009)		
Observations	392,033		
R^2	0.337		

Jarque-Bera normality test for γ: 1.629 Chi(2) 0.4428

Skewness/Kurtosis tests for Normality of γ 0.3325

Note. The sample includes all 6–18 years children in the DHS data from 59 DHS rounds (for which ethnicity variable is available) out of 86 DHS rounds. Control variables include child-level characteristics (sex, age, birth-year, and month of birth), mother-level characteristics (mother's education, mother's age, mother-age squared, mother height, mother ethnicity, her religion, the number of her children who have died, and an indicator variables taking one if any of the mother's children died before their first birthday), father-level characteristics (his age and education), and household-level characteristics (residence area and wealth quintile). Standard errors are clustered at the DHS cluster level and reported in parentheses

Table 14 Testing twin instrument validity using data on children under 5

	(I) Log birth weight	(II) Log current weight	(III) Child's current body mass index decile	(IV) Education z-score
Twin	−0.170*** (0.005)	−0.070*** (0.007)	−0.303*** (0.097)	−0.010 (0.011)
Log birth weight		0.130 (0.005)	1.28*** (0.075)	
Clusters	14,383	12,528	12,494	24,465
Observations	70,902	40,614	40,251	395,202
R^2	0.100	0.725	0.114	0,297

Note: In Col. I, we control for mother characteristics (age, education, body mass index, age at first birth, ethnicity), child characteristics (sex, preceding birth interval, birth year), and household characteristics (residence and wealth quintile). In Cols. (II) and (III), we additionally control for number of under 5 year olds living in the household, months of breastfeeding, and whether the mother lives with her husband. Robust standard errors are clustered at the DHS cluster level and reported in parentheses. Cols. I to III are based on data on under 5 year olds from 59 (for which ethnicity variable is available) out of our 86 DHS rounds. In Col. IV, we use data on schooling-age children from the same 59 DHS to compare education of twins to that of singletons controlling for the entire set of age dummies, wealth quintiles, region of residence, child sex, parental education, and ethnicity fixed effects

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 15 Robustness check IV estimates of the effect of family size on children schooling accounting for heterogeneity of twin effect on family size

	Dependent variable: education *z*-score		
	(I)	(II)	(III)
	Panel A : 2+Sample		
Number of children	0.019	0.008	0.018
	(0.056)	(0.054)	(0.054)
Clusters	20,339	20,339	20,339
Observations	64,339	99,875	99,875
R^2	0.339	0.340	0.339
Cragg-Donald Wald F	72.16	74.29	74.31
(Stock-Yogo critical values)	(19.93)	(19.93)	(19.93)
	Panel B : 3+Sample		
Number of children	0.098**	0.092**	0.094**
	(0.047)	(0.046)	(0.047)
Clusters	20,117	20,117	20,117
Observations	99,875	99,875	99,875
R^2	0.297	0.299	0.299
Cragg-Donald Wald F	131.75	132.73	133.01
(Stock-Yogo critical values)	(19.93)	(19.93)	(19.93)
	Panel C: 4+Sample		
Number of children	− 0.061	− 0.047	− 0.064
	(0.064)	(0.063)	(0.062)
Clusters	18,307	18,307	18,307
Observations	101,848	101,848	101,848
R^2	0.304	0.306	0.304
Cragg-Donald Wald F	68.47	70.85	70.92
(Stock-Yogo critical values)	(19.93)	(19.93)	(19.93)

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Note. In Col. (I), we interact twin at the second, third, and fourth order respectively in 2+, 3+, and 4+Samples with residence in rural area; in Col. (II), we interact with mother's religion and in Col. (III) we interact with residence in rural West and Central Africa. Control variables include child-level characteristics (sex, age, birth-year, and month of birth), mother-level characteristics (mother's education, mother's age, mother-age squared, mother height, mother ethnicity, her religion, the number of her children who have died, and an indicator variables taking one if any of the mother's children died before their first birthday), father-level characteristics (his age and education), and household-level characteristics (residence area and wealth quintile). We add residence in rural West and Central Africa in Col. III. Standard errors are clustered at the DHS cluster level and reported in parentheses

Table 16 Robustness check IV estimates of the effect of family size on children schooling using partially missing instruments and an unrestricted specification

Dependent variable: education z-score			
	All	Poor	Non-poor
	Panel A : 2+Sample		
Number of children > 2	− 0.050	− 0.326	− 0.277
	(0.264)	(0.384)	(0.407)
Number of children > 3	0.156	0.048	0.354**
	(0.122)	(0.180)	(0.176)
Number of children > 4	− 0.083	− 0.158	− 0.153
	(0.140)	(0.183)	(0.222)
Number of children > 5	0.021	0.079	0.179
	(0.166)	(0.229)	(0.290)
Clusters	20,339	11,287	11,080
Observations	64,339	26,814	25,458
R^2	0.396	0.241	0.427
	Panel B : 3+Sample		
Number of children > 3	0.275**	0.189	0.473**
	(0.127)	(0.180)	(0.194)
Number of children > 4	− 0.104	− 0.144	− 0.239
	(0.115)	(0.152)	(0.185)
Number of children > 5	0.023	0.019	0.124
	(0.129)	(0.178)	(0.230)
Clusters	20,117	11,286	11,020
Observations	99,875	43,802	36,717
R^2	0.374	0.225	0.407
	Panel C : 4+Sample		
Number of children > 4	− 0.110	− 0.092	− 0.288
	(0.120)	(0.161)	(0.194)
Number of children > 5	0.043	− 0.002	0.120
	(0.116)	(0.161)	(0.204)
Clusters	18,307	10,314	9,372
Observations	101,848	47,269	34,662
R^2	0.362	0.226	0.390

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Note. The partially missing instruments are constructed based on the initial twin at the second, third, fourth, and fifth birth dummies and a polynomial function of mother age, mother education, father age and father education, controlling for child characteristics, households characteristics, and mother ethnicity and her health conditions. Number of children>x is the marginal effect from moving from x to x+1 children. Control variables are as specified in the note of Table 15. Standard errors are clustered at the DHS cluster level and reported in parentheses

Table 17 Robustness check using samples including children with school-aged siblings living outside the household

	Dependent variable: education z-score		
	All	Poor	Non-poor
	Panel A: 2+Sample		
Number of children	0.029 (0.056)	−0.076 (0.074)	0.035 (0.095)
Clusters	20,671	11,452	11,550
Observations	68,259	28,848	26,563
R^2	0.332	0.156	0.377
Cragg-Donald Wald F	135.95	54.48	58.77
(Stock-Yogo critical values)	(16.38)	(16.38)	(16.38)
	Panel B : 3+Sample		
Number of children	0.091* (0.048)	0.092 (0.090)	0.111* (0.062)
Clusters	20,905	12,207	11,737
Observations	112,285	50,312	40,115
R^2	0.291	0.132	0.335
Cragg-Donald Wald F	249.61	60.27	159.71
(Stock-Yogo critical values)	(16.38)	(16.38)	(16.38)
	PANEL C: 4+Sample		
Number of children	−0.114* (0.070)	−0.113 (0.103)	−0.152 (0.097)
Clusters	19,491	11,514	10,261
Observations	119,544	56,975	39,087
R^2	0.278	0.123	0.317
Cragg-Donald Wald F	111.85	36.61	65.13
(Stock-Yogo critical values)	(16.38)	(16.38)	(16.38)

Note. Here, the 2+, 3+, and 4+Samples include both children without and with school-aged children living outside the household. Controls variables are as specified in Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 18 Regression results for decentralized level of decision (mother level analysis – with number of children defined as the mother's total number of sons and daughters living in the household) using decentralized samples

	Dependent variable: education *z*-score		
	All	Poor	Non-poor
	Panel A: 2+Sample decentralized		
Number of children	0.001 (0.066)	−0.073 (0.084)	0.042 (0.115)
Clusters	19,820	11,012	10,887
Observations	63,994	28,226	23,275
R^2	0.330	0.153	0.379
Cragg-Donald Wald F (Stock-Yogo critical values)	175.97 (16.38)	83.09 (16.38)	64.29 (16.38)
	Panel B: 3+Sample decentralized		
Number of children	0.059 (0.042)	0.037 (0.066)	0.128** (0.060)
Clusters	20,780	11,861	11,843
Observations	116,571	51,301	42,580
R^2	0.314	0.154	0.351
Cragg-Donald Wald F (Stock-Yogo critical values)	561.96 (16.38)	189.33 (16.38)	271.82 (16.38)
	Panel C: 4+Sample decentralized		
Number of children	0.003 (0.044)	−0.009 (0.059)	−0.055 (0.070)
Clusters	18,087	10,291	9,219
Observations	111,179	51,827	37,573
R^2	0.298	0.152	0.342
Cragg-Donald Wald F (Stock-Yogo critical values)	618.43 (16.38)	268.98 (16.38)	274.51 (16.38)

Note. This table shows the effect of the number of children on children's schooling using mother as unit of decision (decentralized approach). Control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 19 Robustness check with number of children defined as the total number of births given by the household head's wives

	Dependent variable: education *z*-score		
	All	Poor	Non-poor
	Panel A: 2+Sample		
Number of children	0.018 (0.051)	−0.086 (0.065)	0.019 (0.088)
Clusters	20,339	11,080	11,287
Observations	64,339	26,814	25,458
R^2	0.339	0.154	0.385
Cragg-Donald Wald F (Stock-Yogo critical values)	159.85 (16.38)	65.99 (16.38)	70.89 (16.38)
	Panel B: 3+Sample		
Number of children	0.092** (0.045)	0.054 (0.076)	0.136** (0.060)
Clusters	20,117	11,286	11,020
Observations	99,875	43,802	36,717
R^2	0.298	0.151	0.334
Cragg-Donald Wald F (Stock-Yogo critical values)	270.95 (16.38)	69.42 (16.38)	179.53 (16.38)
	Panel C: 4+Sample		
Number of children	−0.055 (0.059)	−0.072 (0.103)	−0.116 (0.076)
Clusters	18,087	10,291	9,219
Observations	101,848	47,269	34,662
R^2	0.305	0.146	0.342
Cragg-Donald Wald F (Stock-Yogo critical values)	141.20 (16.38)	31.24 (16.38)	101.39 (16.38)

Note. This table shows the effect of the number of children on children's schooling. The number of children is defined as the total number of births given by the household head's wives. Control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 20 Robustness check using censored education z-score

	Dependent variable: censored education z-score		
	All	Poor	Non-poor
	Panel A: 2+Sample		
Number of children	0.050 (0.045)	− 0.054 (0.061)	0.045 (0.072)
Clusters	20,316	11,059	11,268
Observations	64,106	26,709	25,373
R^2	0.340	0.176	0.432
Cragg-Donald Wald F (Stock-Yogo critical values)	218.50 (16.38)	68.46 (16.38)	95.15 (16.38)
	Panel B: 3+Sample		
Number of children	0.113*** (0.041)	0.115 (0.076)	0.130*** (0.049)
Clusters	20,101	11,267	11,009
Observations	99,543	43,639	36,609
R^2	0.297	0.132	0.383
Cragg-Donald Wald F (Stock-Yogo critical values)	339.57 (16.38)	88.92 (16.38)	209.97 (16.38)
	Panel C: 4+Sample		
Number of children	− 0.054 (0.057)	− 0.083 (0.100)	− 0.133* (0.073)
Clusters	18,296	10,305	9,364
Observations	101,547	47,129	34,543
R^2	0.312	0.156	0.379
Cragg-Donald Wald F (Stock-Yogo critical values)	167.94 (16.38)	42.35 (16.38)	104.46 (16.38)

Note. This table shows the effect of the number of children on education. Education z-scores are constructed based on censored completed years of schooling which is at most equal to the child's age minus 6. Control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 21 Robustness check with years of education as outcome variable

	Dependent variable: completed years of education		
	All	Poor	Non-poor
	Panel A: 2+Sample		
Number of children	0.075 (0.095)	−0.042 (0.139)	0.058(0.145)
Clusters	20,339	11,080	11,287
Observations	64,339	26,814	25,458
R^2	0.508	0.337	0.632
Cragg-Donald Wald F (Stock-Yogo critical values)	219.73 (16.38)	88.71 (16.38)	94.98 (16.38)
	Panel B: 3+Sample		
Number of children	0.144* (0.079)	0.027 (0.141)	0.254** (0.101)
Clusters	20,117	11,286	11,020
Observations	99,875	43,802	36,717
R^2	0.484	0.331	0.586
Cragg-Donald Wald F (Stock-Yogo critical values)	340.80 (16.38)	89.76 (16.38)	209.95 (16.38)
	Panel C: 4+Sample		
Number of children	−0.031 (0.111)	−0.113 (0.187)	−0.103 (0.154)
Clusters	18,307	10,314	9,372
Observations	101,848	47,269	34,662
R^2	0.479	0.321	0.588
Cragg-Donald Wald F (Stock-Yogo critical values)	169.55 (16.38)	42.42 (16.38)	104.46 (16.38)

Note. This table shows the effect of the number of children on education. We use completed years of education as outcome variable instead of education z-scores. Control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 22 Robustness check with a control for the household's region of residence instead of mother's ethnicity

	Dependent variable: education z-score		
	All	Poor	Non-poor
	Panel A: 2+Sample		
Number of children	0.030 (0.048)	− 0.057 (0.061)	− 0.021 (0.080)
Clusters	27,401	14,669	15,555
Observations	85,046	34,598	34,841
R^2	0.334	0.181	0.396
Cragg-Donald Wald F (Stock-Yogo critical values)	192.90 (16.38)	84.05 (16.38)	87.59 (16.38)
	Panel B: 3+Sample		
Number of children	0.063 (0.041)	0.025 (0.079)	0.107** (0.052)
Clusters	26,798	14,876	14,959
Observations	130,066	55,967	49,221
R^2	0.305	0.173	0.349
Cragg-Donald Wald F (Stock-Yogo critical values)	339.85 (16.38)	73.59 (16.38)	232.67 (16.38)
	Panel C: 4+Sample		
Number of children	− 0.080 (0.054)	− 0.125 (0.094)	− 0.116* (0.069)
Clusters	24,053	13,392	12,488
Observations	130,514	59,488	45,595
R^2	0.295	0.140	0.339
Cragg-Donald Wald F (Stock-Yogo critical values)	211.08 (16.38)	53.04 (16.38)	131.58 (16.38)

Note. Region of residence is used instead of mother's ethnicity. This expands the sample to all DHS rounds in which the ethnicity variable is missing. Control variables are as specified in the note of Table A4 with region of residence instead of mother's ethnicity. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 23 Robustness check with a control for country-urban by rural FE instead of mother's ethnicity

	Dependent variable: education *z*-score		
	All	Poor	Non-poor
	Panel A: 2+Sample		
Number of children	0.040 (0.048)	− 0.039 (0.055)	− 0.045 (0.085)
Clusters	34	34	34
Observations	85,046	34,598	34,841
R^2	0.299	0.189	0.359
Cragg-Donald Wald F (Stock-Yogo critical values)	169.77 (16.38)	74.77 (16.38)	79.14 (16.38)
	Panel B: 3+Sample		
Number of children	0.097** (0.044)	0.074 (0.068)	0.139** (0.053)
Clusters	34	34	34
Observations	130,066	55,967	49,221
R^2	0.261	0.163	0.349
Cragg-Donald Wald F (Stock-Yogo critical values)	278.66 (16.38)	59.02 (16.38)	181.86 (16.38)
	Panel C: 4+Sample		
Number of children	− 0.054 (0.061)	− 0.130 (0.095)	− 0.088 (0.084)
Clusters	34	34	34
Observations	130,514	59,488	45,595
R^2	0.272	0.144	0.312
Cragg-Donald Wald F (Stock-Yogo critical values)	195.11 (16.38)	51.34 (16.38)	125.93 (16.38)

Note. Country-urban by rural FE is used instead of mother's ethnicity. This expands the sample to all DHS rounds in which the ethnicity variable is missing. Control variables are as specified in the note of Table 15 with country-urban by rural FE instead of mother's ethnicity. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 24 Robustness check with an alternative definition of poverty

	Dependent variable: education *z*-score	
	Poor	Non-poor
	Panel A: 2+Sample	
Number of children	− 0.030 (0.071)	0.042 (0.093)
Clusters	13,378	12,060
Observations	36,548	27,791
R^2	0.198	0.357
Cragg-Donald Wald F (Stock-Yogo critical values)	68.78 (16.38)	74.27 (16.38)
	Panel B: 3+sample	
Number of children	0.034 (0.073)	0.138** (0.059)
Clusters	13,544	11,758
Observations	59,493	40,382
R^2	0.182	0.308
Cragg-Donald Wald F (Stock-Yogo critical values)	90.16 (16.38)	187.11 (16.38)
	Panel C: 4+Sample	
Number of children	− 0.072 (0.088)	− 0.099 (0.093)
Clusters	12,454	10,035
Observations	63,399	38,449
R^2	0.172	0.324
Cragg-Donald Wald F (Stock-Yogo critical values)	58.90 (16.38)	78.87 (16.38)

Note. The poverty line is defined as the average wealth index in each DHS round. Control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 25 Exploring the proximate cause underlying the zero-effect of family size on schooling using data on first-born in 2+Sample

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the first and second born ≤ 3 years	(IV) Censored starting age	(V) Censored starting age when difference in age between the first and second born ≤ 3 years	(VI) Dropout
	Panel A : all sample					
Number of children	−0.009 (0.023)	−0.172 (0.149)	−0.214 (0.245)	−0.157 (0.129)	−0.162 (0.199)	−0.024 (0.026)
Clusters	20,339	6556	3969	6556	3969	6693
Observations	64,339	13,004	5772	13,004	5772	13,256
R^2	0.346	0.371	0.297	0.349	0.292	0.100
Cragg-Donald Wald F	–	47.83	23.66	47.83	23.66	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B : poor					
Number of children	0.000 (0.039)	0.134 (0.169)	0.192 (0.182)	0.055 (0.163)	0.118 (0.176)	−0.058 (0.057)
Clusters	11,080	2665	1565	2665	1565	2739
Observations	26,814	4053	1974	4053	1974	4145
R^2	0.351	0.363	0.217	0.371	0.239	0.159
Cragg-Donald Wald F	–	19.86	15.06	19.86	15.06	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel C : non-poor					
Number of children	−0.028 (0.023)	−0.499* (0.260)	−0.965* (0.528)	−0.329 (0.204)	−0.695* (0.144)	0.008 (0.039)
Clusters	11,287	3817	2083	3817	2083	3890
Observations	25,458	6580	2707	6580	2707	6686
R^2	0.275	0.159	−0.299	0.159	−0.299	0.111
Cragg-Donald Wald F	–	23.30	9.12	23.30	9.12	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 6). Following Angrist & Pischke (2009: 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although "enrolment" and "drop out" are binary variables

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 26 Exploring the proximate cause underlying the positive effect of family size on schooling, across first and second born in 3+Sample

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the first and second born ≤ 3 years	(IV) Censored starting age	(V) Censored starting age when difference in age between the first and second born ≤ 3 years	(VI) Dropout
	Panel A : first born					
Number of children	0.001 (0.022)	−0.196 (0.139)	−0.048 (0.181)	−0.119 (0.133)	−0.028 (0.165)	−0.014 (0.030)
Clusters	18,912	5872	3894	5872	3894	6019
Observations	53,346	10,546	5626	10,546	5626	10,869
R^2	0.342	0.360	0.322	0.337	0.310	0.104
Cragg-Donald Wald F	–	60.03	33.40	60.03	33.40	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B: second born					
Number of children	0.037 (0.024)	−0.018 (0.180)	−0.222 (0.208)	0.032 (0.141)	−0.126 (0.172)	0.063* (0.035)
Clusters	18,018	5289	3664	5289	3664	5305
Observations	46,529	8805	5187	8805	5187	8818
R^2	0.345	0.388	0.353	0.380	0.358	0.115
Cragg-Donald Wald F	–	46.14	25.15	46.14	25.15	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 5). Following Angrist and Pischke (2009): 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although "enrolment" and "drop out" are binary variables

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 27 Exploring the proximate cause underlying the positive effect of family size on schooling in 3+Sample, across first and second born of poor households using 3+Sample

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the first and second born ≤ 3 years	(IV) Censored starting age	(V) Censored starting age when difference in age between the first and second born ≤ 3 years	(VI) Dropout
	Panel A : first born in poor households					
Number of children	−0.019 (0.047)	−0.039 (0.404)	0.550 (0.534)	−0.115 (0.401)	0.525 (0.523)	0.065 (0.092)
Clusters	10,305	2426	1535	2426	1535	2498
Observations	23,343	3515	1930	3515	1930	3622
R^2	0.349	0.361	0.046	0.355	0.063	0.168
Cragg-Donald Wald F	–	5.54	4.34	5.54	4.34	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B: second born in poor households					
Number of children	−0.010 (0.050)	0.524 (0.405)	0.691 (0.490)	0.468 (0.359)	0.682 (0.433)	0.062 (0.081)
Clusters	9647	2116	1944	2116	1944	2097
Observations	20,459	2925	1386	2925	1386	2858
R^2	0.340	0.184	0.020	0.219	−0.002	0.218
Cragg-Donald Wald F	–	8.30	6.21	8.30	6.21	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 5). Following Angrist and Pischke (2009): 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although "enrolment" and "drop out" are binary variables

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 28 Exploring the proximate cause underlying the positive effect of family size on schooling in 3+Sample, across first and second born of non-poor households using 3+Sample

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the first and second born ≤ 3 years	(IV) Censored starting age	(V) Censored starting age when difference in age between the first and second born ≤ 3 years	(VI) Dropout
	Panel A : first born in non-poor households					
Number of children	0.030 (0.023)	− 0.128 (0.154)	− 0.179 (0.210)	− 0.167 (0.144)	− 0.183 (0191)	− 0.017 (0.030)
Clusters	9,934	3259	2036	3259	2036	3343
Observations	19,664	5010	2633	5010	2633	5115
R^2	0.269	0.319	0.262	0.283	0.237	0.116
Cragg-Donald Wald F	–	44.94	23.96	44.94	23.96	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B: second born in non-poor households					
Number of children	0.023 (0.027)	− 0.285 (0.180)	− 0.623** (0.275)	− 0.146 (0.141)	− 0.413** (0.205)	0.060 (0.040)
Clusters	9282	2877	1944	2877	1944	
Observations	17,053	4212	2501	4212	2501	4287
R^2	0.278	0.259	0.039	0.281	0.108	0.128
Cragg-Donald Wald F	–	34.44	15.30	34.44	15.30	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 3. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 5). Following Angrist and Pischke (2009): 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although "enrolment" and "drop out" are binary variables

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 29 Exploring the proximate cause underlying the positive effect of family size on schooling in 4+Sample, across the second born and third born

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the second and third born ≤ 3 years	(IV) Censored starting age	(V) Censored starting age when difference in age between the second and third born ≤ 3years	(VI) Dropout
	Panel A : second born					
Number of children	−0.010 (0.050)	0.524 (0.405)	0.691 (0.490)	0.468 (0.359)	0.682 (0.433)	0.062 (0.081)
Clusters	15,593	4308	2659	4308	2659	4353
Observations	35,061	6514	3403	6514	3403	6602
R^2	0.336	0.372	0.329	0.366	0.294	0.118
Cragg-Donald Wald F	–	27.04	21.52	27.04	21.52	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B: third born					
Number of children	−0.004 (0.035)	−0.204 (0.234)	−0.109 (0.177)	−0.064 (0.206)	0.063 (0.132)	−0.047 (0.040)
Clusters	15,338	4118	2273	4118	2273	6055
Observations	32,845	5988	2851	5988	2851	4149
R^2	0.384	0.360	0.376	0.387	0.400	0.104
Cragg-Donald Wald F	–	19.67	17.81	19.67	17.81	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 3. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 6). Following Angrist and Pischke (2009): 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although “enrolment” and “drop out” are binary variables

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 30 Exploring the proximate cause underlying the positive effect of family size on schooling in 4+Sample, across the second born and third born of poor households

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the second and third born ≤ 3 years	(IV) Censored starting age	(V) Censored starting age when difference in age between the second and third born ≤ 3 years	(VI) Dropout
	Panel A : second born in poor households					
Number of children	0.034 (0.063)	−0.462 (0.291)	−0.248 (0.326)	−0.303 (0.242)	−0.163 (0.278)	−0.008 (0.085)
Clusters	8453	1793	1061	1793	1061	1798
Observations	16,345	2353	1250	2353	1250	2336
R^2	0.339	0.291	0.284	0.358	0.314	0.177
Cragg-Donald Wald F	–	11.92	8.77	11.92	8.77	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B: third born in poor households					
Number of children	−0.027 (0.068)	0.240 (0.276)	−0.147 (0.213)	0.311 (0.280)	−0.036 (0.225)	0.016 (0.087)
Clusters	8170	1656	992	1656	992	1662
Observations	14,998	2068	865	2068	865	2047
R^2	0.332	0.362	0.445	0.330	0.477	0.170
Cragg-Donald Wald F	–	8.41	5.10	8.41	5.10	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 3. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 6). Following Angrist and Pischke (2009): 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although "enrolment" and "drop out" are binary variables

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 31 Exploring the proximate cause underlying the positive effect of family size on schooling in 4+Sample, across the second born and third born of non-poor households

	(I) Enrolment	(II) Starting age	(III) Starting age when difference in age between the second and third born ≤3years	(IV) Censored starting age	(V) Censored starting age when difference in age between the second and third born ≤ 3 years	(VI) Drop out
	Panel A : second born in non-poor households					
Number of children	−0.072* (0.042)	0.342 (0.441)	0.469 (0.435)	0.375 (0.370)	0.641 (0.414)	−0.048 (0.047)
Clusters	7329	2184	1265	2184	1265	2228
Observations	11,842	2916	1497	2916	1497	2998
R^2	0.273	0.198	0.157	0.198	−0.057	0.132
Cragg-Donald Wald F	–	6.75	6.89	6.75	6.89	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–
	Panel B: third born in non-poor households					
Number of children	0.020 (0.046)	−0.557 (0.429)	−0.159 (0.380)	−0.320 (0.141)	0.425 (0.314)	−0.076* (0.040)
Clusters	7205	2149	1125	2149	1125	2180
Observations	11,451	2852	1325	2852	1325	2922
R^2	0.301	0.050	0.307	0.189	0.307	0.139
Cragg-Donald Wald F	–	8.64	5.66	8.64	5.66	–
(Stock-Yogo critical values)	–	(16.38)	(16.38)	(16.38)	(16.38)	–

Note. The control variables are as specified in the note of Table 3. Robust standard errors are clustered at the DHS-cluster level and reported in parentheses. In Cols (2) to (5), we consider only children enrolled in school at the time of the survey. In Cols (4) and (5), the school starting age is generated based on censored years of education (see robustness check 6). Following Angrist and Pischke (2009): 102–107), we use manual 2SLS as a linear probability model in Cols (1) and (6), although "enrolment" and "drop out" are binary variables

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table 32 Formal tests of child labour mechanism

	Dependent variable: education *z*-score		
	(I)	(II)	(III)
	Panel B: 2+Sample		
Number of children	−0.030 (0.091)	−0.037 (0.090)	−0.035 (0.086)
Child labour (less than 8 h/week)		–	
8 to 21 h/week		0.034* (0.015)	
22 to 35 h/week		−0.066*** (0.026)	
35 h/week and+		−0.144*** (0.029)	
Siblings labour (total hours)			−0.002 (0.002)
Clusters	6901	6901	6901
Observations	22,485	22,485	22,485
R^2	0.299	0.301	0.300
Cragg-Donald Wald F (Stock-Yogo critical values)	48.16 (16.38)	48.72 (16.38)	57.04 (16.38)
	Panel B: 3+Sample		
Number of children	0.211** (0.099)	0.222** (0.100)	0.237** (0.108)
Child labour (less than 8 h/week)		–	
8 to 21 h/week		−0.002 (0.013)	
22 to 35 h/week		−0.066*** (0.023)	
35 h/week and+		−0.169*** (0.025)	
Siblings labour (total hours)			−0.005*** (0.002)
Clusters	6851	6851	6851
Observations	35,153	35,153	35,153
R^2	0.192	0.184	0.183
Cragg-Donald Wald F (Stock-Yogo critical values)	68.05 (16.38)	67.21 (16.38)	62.40 (16.38)

Table 32 (continued)

	Dependent variable: education z-score		
	(I)	(II)	(III)
	Panel C: 4+Sample		
Number of children	− 0.138 (0.088)	− 0.131 (0.088)	− 0.139 (0.094)
Child labour (less than 8 h/week)		–	
8 to 21 h/week		− 0.002 (0.013)	
22 to 35 h/week		− 0.037* (0.021)	
35 h/week and+		− 0.128*** (0.022)	
Siblings labour (total hours)			0.000 (0.001)
Clusters	6288	6288	6288
Observations	35,865	35,865	35,865
R^2	0.251	0.255	0.250
Cragg-Donald Wald F (Stock-Yogo critical values)	52.66 (16.38)	52.33 (16.38)	48.98 (16.38)

Note: Col. (I) does not include child labour; Col. (II) includes own child labour and Col.(III) includes siblings' child labour. Control variables are as specified in the note of Table 15. Robust standard errors are clustered at the DHS cluster level and reported in parentheses. 2+, 3+, and 4+Samples are restricted to countries for which data on child labour is available. Child labour is defined as the sum of time spent in household chores, number of working hours for a family member (including him/her self), and number of working hours for someone outside the household

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Altonji JG, Elder TE, Taber CR (2005) Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. J Polit Econ 113:151–184

Angrist JD, Pischke J-S (2009) Mostly harmless econometrics: an empiricist's companion, 1 edition. ed. Princeton University Press, Princeton

Angrist J, Lavy V, Schlosser A (2010) Multiple experiments for the causal link between the quantity and quality of children. J Labor Econ 28(4):773–824

Baland JM, Bonjean I, Guirkinger C, Ziparo R (2016) The economic consequences of mutual help in extended families. J Dev Econ 123:38–56

Bass LE (2004) Child labor in sub-Saharan Africa. Lynne Rienner Publishers

Becker GS (1960) An economic analysis of fertility. In: National Bureau of Economic Research (ed.). Demographic and economic change in developed countries. Princeton: Princeton University Press: 209–231

Bhalotra S R, Clarke D (2016) The twin instrument. IZA Discussion Paper No. 10405

Black SE, Devereux PJ, Salvanes KG (2005) The more the merrier? The effect of family composition on children's education. Q J Econ 120:669–700

Black SE, Devereux PJ, Salvanes KG (2010) Small family, smart family? Family size and the IQ scores of young men. J Hum Resour 45(1):33–58

Blake J (1989) Family size and achievement. University of California Press, Berkeley

Bongaarts J (2009) Human population growth and the demographic transition. Philos Trans R Soc Lond B 364(1532):2985–2990

Booth AL, Kee HJ (2009) Birth order matters: the effect of family size and birth order on educational attainment. J Popul Econ 22:367–397

Cáceres-Delpiano J (2006) The impacts of family size on investment in child quality. J Hum Resour 41(4):738–754

Chaudhury N, Hammer J, Kremer M, Muralidharan K, Rogers FH (2006) Missing in action: teacher and health worker absence in developing countries. J Econ Perspect 20(1):91–116

Cole TJ, Flegal KM, Nicholls D, Jackson AA (2007) Body mass index cut offs to define thinness in children and adolescents: international survey. Bmj 335(7612):194

Conley D, 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. J Hum Resour 41(4):722–737

Conley TG, Hansen CB, Rossi PE (2012) Plausibly exogenous. Rev Econ Stat 94(1):260–272

Fitzsimons E, Malde B (2014) Empirically probing the quantity–quality model. J Popul Econ 27:33–68

Garcia M, Fares J (2008) Youth in Africa's labor market (English). Directions in development; human development. Washington, DC: World Bank.

Guarcello L, Lyon S, Valdivia C (2015) Evolution of the relationship between child labour and education since 2000: evidence from 19 developing countries. Working Paper March 2015. Understanding Children's Work (UCW) Programme

Guo G, Van Wey L (1999) Sibship size and intellectual development: is the relationship causal? Am Sociol Rev 64:169–187

Holmes J, Tiefenthaler J (1997) Cheaper by the dozen? The marginal time costs of children in the Philippines. Popul Res Policy Rev 16(6):561–578

Inhorn MC (2003) Global infertility and the globalization of new reproductive technologies: illustrations from Egypt. Soc Sci Med 56(9):1837–1851

Joshi S, Schultz TP (2007) Family planning as an investment in development: evaluation of a program's consequences in Matlab, Bangladesh. Yale University Economic Growth Center Discussion Paper, (951)

Lee J (2008) Sibling size and investment in children's education: an Asian instrument. J Popul Econ 21:855–875

Lehmann JYK, Nuevo-Chiquero A, Vidal-Fernandez M (2018) The early origins of birth order differences in children's outcomes and parental behavior. J Hum Resour 53(1):123–156

Lesthaeghe R (2014) The fertility transition in sub-Saharan Africa into the 21st century. University of Michigan, Ann Arbor

Li H, Zhang J, Zhu Y (2008) The quantity-quality trade-off of children in a developing country: identification using Chinese twins. Demography 45(1):223–243

Liu H (2014) The quality–quantity trade-off: evidence from the relaxation of China's one-child policy. J Popul Econ 27:565–602
Lloyd CB, Blanc AK (1996) Children's schooling in sub-Saharan Africa: the role of fathers, mothers, and others. Popul Dev Rev 22(2):265–298
Maralani V (2008) The changing relationship between family size and educational attainment over the course of socioeconomic development: evidence from Indonesia. Demography 45(3):693–717
Marteleto LJ, de Souza LR (2012) The changing impact of family size on adolescents' schooling: assessing the exogenous variation in fertility using twins in Brazil. Demography 49(4):1453–1477
Marteleto LJ, de Souza LR (2013) The implications of family size for adolescents' education and work in Brazil: Gender and birth order differences. Social Forces 92(1):275–302
Mogstad M, Wiswall M (2016) Testing the quantity–quality model of fertility: estimation using unrestricted family size models. Quant Econ 7:157–192
Mueller E (1984a) The value and allocation of time in rural Botswana. J Dev Econ 15(1):329–360
Mueller E (1984b) Income, aspirations, and fertility in rural areas of less developed countries. In: Schutjer WA, Shannon Stokes C (eds) Rural development and human fertility. Macmillan, New York, pp 121–150
Oster E (2017) Unobservable selection and coefficient stability: theory and validation. J Bus Econ Stat. Forthcoming
Qian N(2009) Quantity-quality and the one child policy: the only-child disadvantage in school enrollment in rural China (No. w14973). National Bureau of Economic Research
Rosenzweig MR, Wolpin KI (1980) Testing the quantity-quality fertility model: The use of twins as a natural experiment. Econometrica: journal of the Econometric Society 227–240
Rosenzweig MR, Zhang J (2009) Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. Rev Econ Stud 76(3):1149–1174
Sinha N (2005) Fertility, child work, and schooling consequences of family planning programs: evidence from an experiment in rural Bangladesh. Econ Dev Cult Chang 54(1):97–128
Smits J, Monden C (2011) Twinning across the developing world. PLoS One 6(9):e25239
Steelman LC, Powell B, Werum R, Carter S (2002) Reconsidering the effects of sibling configuration: recent advances and challenges. Annu Rev Sociol 28:243–269
Tiefenthaler J (1997) Fertility and family time allocation in the Philippines. Popul Dev Rev 23:377–397
UIS Stat. UNESCO. 2016. http://data.uis.unesco.org
World Bank (2016) World Development Indicators. http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators. Accessed 6 Feb 2016

Affiliations

Sahawal Alidou [1,2] **· Marijke Verpoorten** [1,2]

Sahawal Alidou
sahawal.alidou@uantwerp.be

[1] IOB–Institute of Development Policy and Management, University of Antwerp, Antwerp, Belgium

[2] LICOS–Centre for Institutions and Economic Performance, KU Leuven, Leuven, Belgium