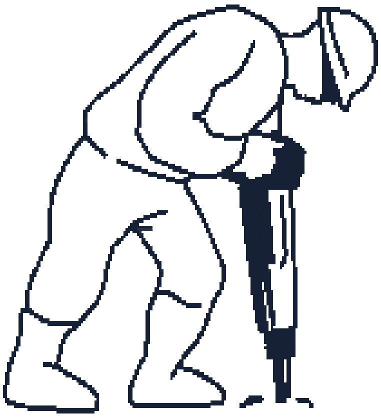
Southern Africa Labour and Development Research Unit



Gender and Birth Order Effects on Intra-household Schooling Choices and Education Attainments in Kenya

by
Fredrick M. Wamalwa and Justine Burns

Working Paper Series
Number 203, Version 0

About the Author(s) and Acknowledgments

Fredrick Wamalwa: School of Economics, Faculty of Commerce, University of Cape Town, Private Bag, Rondebosch 7701, Cape Town, South Africa. E-mail: fredwamalwa@yahoo.co.uk - corresponding author
Justine Burns: School of Economics, University of Cape Town, South Africa

We are grateful to Dr. Soham Sahoo of the University of Goettingen (Germany) for useful comments. All remaining errors are our own.

Recommended citation

Wamalwa., F.M, Burns, J. (2017). Gender and Birth Order Effects on Intra-household Schooling Choices and Education Attainments in Kenya. Tcpqgm l 0, Cape Town: SALDRU, UCT. (SALDRU Working Paper Number 203).

ISBN: 978-1-928281-64-1

Working Papers can be downloaded in Adobe Acrobat format from www.saldru.uct.ac.za.
Printed copies of Working Papers are available for R25.00 each plus vat and postage charges.

Orders may be directed to:
The Senior Administrative Officer, SALDRU, University of Cape Town, Private Bag, Rondebosch, 7701,
Tel: (021) 650 1808, Fax: (021) 650 5697, Email: tania.hendricks@uct.ac.za

Gender and Birth Order Effects on Intra-household Schooling Choices and Education Attainments in Kenya*

Fredrick M. Wamalwa† & Justine Burns‡

April 10, 2017

In this paper, we investigate the effect of two important family characteristics- *gender* and *birth order*- on *intra-household investments in*, and *educational outcomes of*, children in Kenya. We measure *intra-household education investments in children* by household's decision to enrol children in private schools and *educational outcomes* by two variables, *completed years of education* and *relative grade attainment.* We use a large household survey data that allows us to apply the family fixed effects models that address the potential endogeneity of children's gender and family size as well as factors that are unobservable at the household level. Although we do not find an intra-household gender preference in terms of investments in children's education, there is a female advantage in terms of the two measured education outcomes. Such female advantage is in contrast with literature generally reported from developing countries. It is, however, in line with global trends which show that more girls are getting educated and the gender gap in education has narrowed considerably. Regarding birth order effects, we find significant negative birth order effects on private enrolment, completed years of education and relative grade attainment. The negative birth order effects are not in line with the evidence from many other developing countries but are in line with results from developed countries. Our results are robust to different sample restrictions. We find that household wealth plays a significant role in propagating the birth order but not the

*We are grateful to Dr. Soham Sahoo of the University of Goettingen (Germany) for useful comments. All remaining errors are our own.

†*Corresponding author*: School of Economics, Faculty of Commerce, University of Cape Town, Private Bag, Rondebosch 7701, Cape Town, South Africa. E-mail: fredwamalwa@yahoo.co.uk.

‡School of Economics, University of Cape Town, South Africa

gender effects we observe.

Key words- birth order, gender, household fixed effects, fully interacted models, Kenya.

1 Introduction

When faced with resource constraints, parents can send some of their children to work and let others concentrate on studying at school (Basu and Van 1998, Tenikue and Verheyden 2010). Similarly, if resources become scarce, parents can send some of their children to fee paying private schools and let others attend public schools (which are generally free of charge in most developing countries following the free primary education policies). As we shall show in table 1, 15 percent of the households in our data send some children to private and others to public schools. What specific family characteristics influence such parental decisions? Although most schooling decisions are made within the family, research on family characteristics that influence children's education outcomes is still inconclusive and has not been fully exploited (Tenikue and Verheyden 2010, Black et al. 2005).

A number of family characteristics have been found to influence children's education outcomes. These include sibling features such as birth order (De Haan et al. 2014, Price 2008, Booth and Kee 2009, Conley and Glauber 2006, Tenikue and Verheyden 2010, 2007, Kristensen and Bjerkedal 2007, Harkonen 2014, Ejrnaes and Portner 2004), birth spacing (De Haan 2010, Black et al. 2005), family size (Black et al. 2005, De Haan 2010) and gender of the child (Sahoo 2016, Maitra et al. 2016). In particular, the role that birth order plays in influencing child outcomes has dominated the economic literature. This stems from interesting findings in the psychology and sociology literature (Zajonc 1976, Zajonc and Markus 1976, Zajonc 1983) which shows that child outcomes are higher for first born children. This finding is attributed to different factors including the intellectual environment within the household (Zajonc 1976, Zajonc and Markus 1976) and family resources (Blake 1981, Downey 2001) both of which are hypothesized to diminish with each additional child in the family.

Unlike literature in psychology and sociology, economic literature, both theoretical and empirical, does not demonstrate a universal first-born advantage. While literature in developed countries points to a first-born advantage in terms of education outcomes (Black et al. 2005, De Haan 2010), evidence from developing countries generally finds a first-born disadvantage, that is, birth order effects are positively related to education outcomes (De Haan et al. 2014, Basu and Van 1998, Tenikue and Verheyden 2007, 2010, Ejrnaes and Portner 2004). Birth order effects on child outcomes are therefore still the subject of further research. Another family characteristic that has received a fair amount of theoretical attention is the gender of the child. Generally, most studies in developing countries document the existence of a gender bias in favor of boys in terms of education and other social outcomes (Maitra et al. 2014, 2016, Sahoo 2016).

This paper contributes to this literature by investigating the effect of two important family characteristics, *gender* and *birth order*, on *intra-household investments in*, and *educational outcomes of*, children in Kenya. Following Caceres-Delpiano (2006), we measure *intra-household education investments in children* by household decisions to enrol children in private schools.[1] We define *educational outcomes* by two variables: *completed years of education* and *relative grade progression.* In Kenya, we are only aware of the study by Tenikue and Verheyden (2010) that empirically examined the effect of birth order on educational attainments in twelve countries in sub-Sahara Africa, Kenya included. This study does not however explicitly make a distinction between variables that measure child investments (such as, enrolment of children in private schools) and those that measure child outcomes (completed years of education and relative grade attainment) in examining birth order effects.

An obvious obstacle here is the endogeneity of gender, birth order and even family size, Following Tenikue and Verheyden (2010), Moshoeshoe (2016), Maitra et al. (2016), Black et al. (2005), Sahoo (2016) and De Haan et al. (2014), we apply the family fixed effects model which allows us to remove all sources of such common household level unobserved heterogeneity.

Here is the summary of our results. Although we do not find an intra-household gender preference in terms of investments in children's education, as measured by household enrolment of children in private schools, there is a female advantage in terms of completed years of education and relative grade progression. Relative to their male siblings, female siblings complete 0.138 more years of education. Female siblings also progress through school faster, accumulating 0.025 more years of education relative to their male siblings. Regarding birth order, our results show significant negative effects of birth order on private school enrolment, completed years of education and relative grade progression. Our results are robust to different robustness checks including correction of selectivity bias due to non-enrolment of children and further attempts to measure birth order effects more accurately. Lastly, we find that

[1] Recall that private schools in Kenya are not free. Even in low cost private schools, parents pay tuition fees at an average of less than 10 US Dollars per month (Piper et al. 2014, Piper and Mugenda 2010, Piper et al. 2015).

household wealth plays a significant role in propagating the birth order but not the gender effects we observe.

The rest of the paper is organized follows. In section 2, we review the theoretical and empirical evidence of gender and birth order effects. Section 3 provides a description of the data while section 4 provides the emprical model. In section 5, we discuss the empirical results and provide a conclusion in section 6.

2 Who is the Most Favoured Child? Theoretical and Empirical Evidence

One of the earliest theories originating from the psychology literature is the confluence model. According to this model, a child's intellectual ability is determined by the sum of the intellectual level in his or her family. As more children are born in the household, the average intellectual environment declines(Zajonc 1976, Zajonc and Markus 1976). As a result, first-born children are advantaged since they are born in a household with a higher average intellectual environment. Last-borns enter the household when the average intellectual level is at its lowest. The model therefore predicts a negative correlation between birth order and outcomes such as educational attainments.

The confluence model further predicts that first-borns also develop skills by tutoring younger siblings, thus earning skills that help them excel in academic related outcomes. Such tutoring effect explains the model's predictions that *only children* are likely to do worse in outcomes (such as academic achievements) than *first-borns in large families* because the former do not have anyone to tutor (Zajonc and Markus 1976). The model further posits that being born in a household with long, rather than short, spacing between siblings reduces the first-born advantage (Zajonc 1983).

Away from psychology and closer to economic literature is the resource dilution hypothesis. According to this theory, all forms of parental resources, such as time and finances, are generally limited (Jaeger 2009, Downey 2001). In particular, as the size of the family enlarges, *per capita family* economic and material resources which cannot be shared by siblings decline (Downey 2001). Like the confluence model, this model predicts that an increase in family size leads to poorer child outcomes. Closely related to the resource dilution hypothesis is the quantity-quality theory by Becker (1960) whose key feature is the interaction of child quality and quantity in the household budget constraint. According to this model, when there are capital market imperfections and parents have many children (like in developing countries), they can, for a given income, invest less in each child than if they have fewer children. The model predicts a negative relationship between sibling size and child outcomes.

There exists a number of studies that have looked at the empirical evidence of the above theoretical predictions. We start with empirical evidence from developed countries. Much of the evidence in developed countries is generally in line with the theoretical predictions that birth order negatively affects schooling outcomes (Price 2008, Black et al. 2005, Booth and Kee 2009, De Haan 2010, Harkonen 2014, Kristensen and Bjerkedal 2007). A study by Black et al. (2005) makes use of survey data from Norway to estimate the effects of birth order on children's education attainments. They estimate a household fixed-effects model and find that birth order is negatively and significantly related to child education. They do not find evidence that the results are driven by parental resources (measured by mother's level of education) and household residential location (rural/urban). They also show that birth order effects do not significantly differ by gender.

Booth and Kee (2009) use the 13th wave of the British Household Panel Survey to estimate the effects of family size and birth order on children's educational attainment. They control for parental household income, parental age at birth and a number of household level attributes. They find that children from larger households have lower levels of education. In line with Black et al. (2005), they find negative birth order effects on children's educational attainment. Using data from the American Time Use Survey in the US, Price (2008) estimates the relationship between birth order and parental time use. He finds that first born children receive significantly more quality time from their parents relative to later-born children.

De Haan (2010) estimates the effect of birth order on education attainments of children in USA and Norway using the Wisconsin Longitudinal Study data and the Brabant Survey data, respectively. To address the potential endogeneity resulting from the possible correlation between birth order and number of sibling, they follow Black et al. (2005) by estimating birth order by household size. Like Black et al. (2005), their results reveal a negative and significant effect of birth order on child education. The authors further find that the age gap between children does not affect the effect of birth order. In addition, the negative effect does not differ between children from higher or lower educated parents. Other studies based in developed countries that find similar negative birth order effects on educational attainment include Harkonen (2014) for Germany and Conley and Glauber (2006)and Kristensen and Bjerkedal (2007) for USA.

Unlike developed countries, much of the evidence in developing countries finds a first-born disadvantage in terms of education outcomes, mainly due to high poverty and low education levels in developing countries (De Haan et al. 2014). High poverty levels in developing countries force households to invest less in the human capital of first-born children (Ejrnaes and Portner 2004, Basu and Van 1998, Tenikue and Verheyden 2010, 2007), for example. The income from older children contributes to family resources, allowing young children to go to school thus explaining the positive birth order effects on child schooling.

Empirical evidence seems to support the above resource constraint hypothesis. Tenikue and Verheyden (2010) estimate a household fixed effects model to examine the impact of birth order on educational attainment using the demographic and health survey data for 12 sub-Saharan Africa countries including Kenya. Educational attainment is measured by the number of completed years of education (for the children in their sample as at the time of the survey). In line with the resource constraint hypothesis, they find positive birth order effects (for example, a first-born disadvantage) in poor households while the opposite holds for richer households. An earlier study by Tenikue and Verheyden (2007) also applied a family fixed effects model to examine the impact of birth order on educational attainment using the Cameroon Household Survey of 2001. This study finds a first-born disadvantage among poor households which however disappears in wealthier households indicating that wealth has the potential to reverse the first-born advantage.

Empirical evidence further points to other factors for the positive birth order effects in developing countries apart from resource constraints. De Haan et al. (2014) estimate the causal effect of birth order on the probability of school enrolment and child labor using nationally representative data from Ecuador. Their estimates, based on family fixed effects, show that earlier born (early order) children are less likely to enrol in school but more likely to participate in child labor relative to later born (later order) children. Analyzing the mechanisms driving their results, De Haan et al. (2014) find that earlier born children receive less quality time from their mothers and more so, are breast feed for a shorter time. The evidence presented in their paper also shows that birth order effects are particularly large among poor households.

Another sibling characteristic that has received theoretical attention is the gender of the child. Generally, most studies in developing countries document the existence of gender bias in favor of boys. For instance, in Ghana, a study by Garg and Morduch (1998) found that children are better off in terms of measured health indicators if they have sisters and no brothers. Another study, by Gupta (1987) based in rural, India shows that although there is a preference for boys in terms of food allocation, clothing, and education and medical expenses, parents also discriminate selectively against some of their daughters. In particular, in households where there is more than one daughter, younger daughters are worse off relative to their older sisters.

In another study from India, Sahoo (2016) explores the role of gender on intra-household schooling choices between private and government schools. He estimates a household fixed effects model on a three-period longitudinal data set based on rural households from Uttar Pradesh and finds that girls are less likely to be enrolled in private schools by 6 percentage points. Similarly, Maitra et al. (2016) examines the role of gender in private school choice using two nationally representative data sets from household surveys conducted in India in 2005 and 2012. They find that female siblings are significantly disadvantaged in both survey years. They find a significant female disadvantage in both surveys (4 percentage points in 2005 and 6 percentage point in 2012) which varies across sub-samples and years.

This paper uses the case study of Kenya to contribute to this literature by empirically testing the prediction of the theoretical models we have reviewed with regard to gender and birth order. Apart from Tenikue and Verheyden (2007), we are not aware of any other study that has explored the gender and birth order effects on intra-household schooling choices and education attainment in Kenya.

3 Data and Descriptive Statistics

3.1 The Uwezo Survey Data

We use the third round of Uwezo survey for Kenya collected in 2012. The Uwezo[2] initiative has been implementing large-scale household surveys that assess literacy and numeracy competencies of school age children since 2009. A detailed description of the sampling strategy is provided in Jones et al. (2014). The third round of Uwezo survey was based on a two-stage random sampling design. First, 30 primary sampling units [3] from each district were selected with the probability of selection proportional to population size. Second, about 20 households in each enumeration area were selected via systematic random sampling.[4] Uwezo

[2]Uwezo which means 'capability' in Kiswahili, is a non-governmental organization that aims to improve competencies in literacy and numeracy among children aged 6-16 years in Kenya. More details about Uwezo can be found at: http://www.uwezo.net/.

[3]Primary Sampling Units generally represent enumeration areas and/or villages

[4]The sample design was provided by the Kenya National Bureau of Statistics (KNBS).

targets children aged 6-16 years who are regular residents of the household. Households without such children were therefore excluded.

In each enumeration area, data collection involved three steps. First, data was collected from one randomly selected local public primary school within the enumeration area.[5] Information was gathered on school enrollment, teachers, classroom facilities as well as school facilities among others. Close to 4,465 public schools were covered. Second, a questionnaire was administered to the administrator of the sampled enumeration areas (villages). Among others, it gathered information on availability of: (i) social amenities (such as chief's office, shopping center and police post), (ii) infrastructure (such as tarmacked roads, all-weather roads, protected water points and electricity), and (iii) the number of educational and health facilities in the village.[6]

Finally, households were visited. A questionnaire was then administered to the head of the household (or representative). For children aged 6-16 years, information was gathered about their age, gender, disability, school grade, whether they were enrolled in school and for those enrolled, the type of school they were enrolled in (private or public) and the time taken to arrive at school. The household questionnaire also collected information on parental age and education as well as indicators of household socioeconomic status. Lastly, each child of school age (6-16 years), whether or not attending school, was assessed in language (English and Kiswahili) and mathematics.

3.2 Descriptive Statistics

The Uwezo survey does not allow us to distinguish between biological and non-biological children of the household. We thus treat all children as though they are biological children of the family. We are of the opinion that this assumption will not have a huge implication on our estimates because information was collected from children *who regularly live* in the household. These children, whether biological or not, by virtue of being regular inhabitants of the household, directly influence intra-household decisions including those related to schooling.[7]

[5]In cases where there was no public primary school in the sampled enumeration area, the nearest public school attended by the majority of children in the sampled enumeration area was selected. When more than one school was available in the enumeration area, the school that attended by a majority of the students was selected.

[6]Information gathered included the number of primary and secondary schools (public and private) as well as the number of village polytechnics. Data was collected on the number of health facilities run by government and non-governmental organizations.

[7]In India, a very recent study by Maitra et al. (2016) examined the effect of gender and birth order on intra-household private school choice. The authors compare estimates based on a sample of only biological

Like Caceres-Delpiano (2006), our sample is limited to children in the household aged 6 to 16 years who are the target group in the Uwezo household assessments. First, we exclude younger children (less than 6 years) because compulsory primary schooling in Kenya begins at age 6. The upper age limit is right-censored since Uwezo targets children of maximum age of 16. Focusing on children aged between 6 to 16 years has an advantage. For instance, the incidence of children below 16 years moving out of the household (to look for a job or otherwise) is low, an aspect that minimizes potential measurement errors in our birth order measures.

Since are exploiting within-family variations in the fixed effects models, it is a requirement that we restrict the sample to include only families with a minimum of two children (Black et al. 2005, Sahoo 2016, Tenikue and Verheyden 2010, Maitra et al. 2016, De Haan et al. 2014).[8] We further limit the sample size to households with a maximum of five children. This restriction ensures we have included families that are not likely to have more children. Limiting the sample to households with five children does not have a huge implication on our estimates since fewer than 1.5 percent of households in our sample have more than five children aged 6-16 years. Following De Haan et al. (2014), we drop households with multiple births (twins) because it is not clear how birth order can be assigned in such cases.[9] We also exclude households and/or children with missing data on key variables such as age.

After imposing all these restrictions, we are left with 96,920 children and 31,568 households. We construct indicators of birth order for the children based on their reported year of birth. In this case, birth order equals 1, 2, 3, 4 and 5 for the first, second, third, fourth and fifth born children aged 6 to 16 years, respectively. According to De Haan (2010), the most accurate measure of birth order is one based on households where all children of the mother are still alive and live at home at the moment of the survey. Although our data does not allow us to know if there are children staying away from the household at the time of the survey and if some of the children were not alive, we provide some evidence in the robustness test section to allay fears about the accuracy of our measure of birth order.

We are interested in three dependent variables. The first is a dummy variable that equals *1 if the child attends a private school* and *0 if the child attends a public school*. Following

children and estimates based on a sample extended to all children who were in the household at the time of the survey. The results between the two samples are similar.

[8]By imposing this restriction, we lose about 39.2 percent of the 72,216 households in the sample.

[9]Uwezo did not directly ask if the household had twins or not. If two or more children in the household have the same age, we assume that such children are twins. We lose about 5.3 percent of the 72,216 households in the sample.

Tenikue and Verheyden (2010), we define our second dependent variable as the child's highest *number of completed years of education* as at the time of the survey. In a developing country like Kenya, researchers such as Mani et al. (2013) have advocated for an education outcome measure that accounts for such late entry and early exit. This measure is called the *relative grade progression* and it is our third dependent variable. It is defined as *completed years of schooling* divided by *potential years of schooling*.[10] The potential years of schooling are *the total number of schooling years a child accumulates if he/she starts schooling on time*[11] *and adds one more in each subsequent year* (Mani et al. 2013).[12]

Since private schooling is not free in Kenya, we consider private school enrolment as a measure of parental investment in children's education allowing us to accurately test economic theories such as the resource dilution hypothesis (Caceres-Delpiano 2006). We associate the next two dependent variables (completed years of education and relative grade progression) as measures of child well-being (education outcomes), that result from inputs such as parental resource allocation (Caceres-Delpiano 2006). Table 1 shows selected descriptive statistics for the whole sample, for households that send their children to public schools and for households that send their children to private schools.

Panel A (column 1) shows the results for our three dependent variables for the children. As can be seen from the table, about 12 percent of children have never been enrolled while about 1 percent have dropped out of school.[13] The majority of children, 87 percent are enrolled and of these 14 percent are enrolled in private schools. The average completed years of education of those who are enrolled is 3.43. If children in our sample enter schools on time (at age 6) and accumulate every year without repeating, we expect the relative grade progression to be equal to 1, showing an efficient education system. As can be seen from table 1, our average relative grade is 0.64 (less than 1 by 37 percentage points), meaning that children in our sample accumulate 0.64 grades per year of schooling.

Column 2 and 3 show that children in public schools have slightly more years of education (probably because they are older). However, children in public schools have slightly lower

[10]Relative grade progression is calculated as $[\frac{Completed\, years\, of\, education}{(Age-6)}]$ where 6 in the denominator denotes the official age of starting school in Kenya.

[11]That is age 6 in the case of Kenya.

[12]Relative grade progression is defined accounting for the fact that official primary school starting age in Kenya is 6. In this case, all children of age 6 and below are assumed to have accumulated zero years of education. As a result, this dependent variable has a small sample size.

[13]Later, we address concerns that such non-enrolment might be a source of selection bias through the Heckman-type selection model since a decision to enroll in a private or a public school is only observed after a decision either to enrol in school or not to enrol has been taken.

relative grade progression level meaning that they are more likely to start school late and/or repeat classes relative to their counterparts in private schools.

Table 1: Mean Statistics for Children and Households

	(1) Whole Sample	(2) Public School	(3) Private School	(4) Public-Private diff
Panel A: Dependent Variables				
Never enrolled	0.12	–	–	–
Dropped out	0.01	–	–	–
Enrolled	0.87	–	–	–
Enrolled in private school	0.14	–	–	–
Completed years of education	3.43	3.50	2.95	0.56***
Relative grade progression	0.64	0.64	0.65	-0.01***
Panel B: Child Characteristics				
Age of the child (in years)	10.87	11.24	10.29	0.95***
Child is female	0.48	0.49	0.49	0.00
Child attends tuition classes	0.39	0.44	0.61	-0.18***
Child has some form of disability	0.03	0.03	0.02	0.01***
Child is first-born	0.38	0.41	0.36	0.05***
Child is second-born	0.36	0.36	0.40	-0.04***
Child is third-born	0.18	0.17	0.18	-0.01***
Child is fourth-born	0.07	0.05	0.06	0.00
Child is fifth-born	0.02	0.01	0.01	0.00
Panel C: Household Characteristics				
Household is in a rural area	0.78	0.79	0.60	0.18***
Age of mother (in years)	36.77	37.09	35.58	1.51***
Years of education of the mother	5.53	5.74	6.88	-1.13***
Number of children aged 6 to 16 years in the family	2.76	2.77	2.61	0.16***
Within household variation in school choice	0.15	–	–	–
Household wealth index (normalized to 0-1 range)	0.40	0.41	0.52	-0.11***
Household belongs to the upper wealth index	0.26	0.25	0.51	-0.27***
Number of children	**96,920**	**74,245**	**9,407**	
Number of households	**31,568**	**22, 423**	**3,330**	

Source: Own calculations based on Uwezo 2012. Notes: (1) Our sample is restricted to households with 2-5 children aged 6-16; (2) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school; (2) Completed years of education is number of years of education the child had completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) The household wealth index is constructed using the principal component analysis (PCA) based on household ownership of durable and livestock assets, type of material used to construct the wall of the dwelling unit, type of lighting regularly used by the household, number of meals taken per day and household sanitation status; (6) We normalize the family wealth index to the range of 0 to 1 and classified household whose index lies 0.5 and above as rich while the rest were classified as poor and; (7) By mother being a teen at birth of the oldest child (in years) means the mother gave birth her first born child when she was 18 years or less.

Looking at child demographic characteristics in panel B (column 1), we find that children are on average 10.87 years old and just below half of them, 48 percent, are female. Less than half of the children, 39 percent, reported that they attend paid tuition classes. A relatively small proportion of children, about 3 percent, were reported to have some form of physical disability. Column 2 and 3 show that relative to their counterparts in public schools, private school students are more likely to be young, female and with no disability. The share of first and second born in the sample is 38 and 36 percent, respectively. The share of third and fourth born children is, respectively, 18 percent and 7 percent. The fifth born are least represented at 2 percent.

In terms of household characteristics, shown in panel C (column 1), we find that the majority of households, 78 percent, are in rural areas. The mean age of the mother is 36.77 years. On average, mothers of children in our sample have 5.53 years of education. There are, on average, 3 children, aged 6-16 years. Lastly, 15 percent of the households exhibit such within-household variation in school type choice where some children attend public schools while others attend private schools even when they are from the same household.

The Uwezo survey does not ask about household income or expenditure, two conventional measures of household living standards. However, the survey includes questions related to household ownership of durable assets (television, radio, computer, mobile phone, car, bicycle, motorbike and cart) and livestock assets (number of cattle, sheep, horses/donkeys/ camels, chicken etc.), type of material used to construct the wall of the dwelling unit, type of lighting regularly used by the household, number of meals taken per day and household sanitation status (whether the household has a source of water and a latrine at home). We use these household socioeconomic characteristics to construct an index of household wealth.[14]

The household family wealth index is constructed using the (ordinary) principal components analysis (PCA).[15] Details of the PCA approach are described and defended in Vyas and Kumaranayake (2006) and Filmer and Pritchett (2001) among others who show that the PCA wealth index performs just as well as traditional measures of household living standards such as household-size-adjusted consumption expenditures, in predicting household outcomes such

[14]The variables used for the construction of the index are (1) a set of eight dummy variables which is equal to one if a household owns each of the following durable assets: television, radio, computer, mobile phone, car, bicycle, motorbike and cart; (2) a set of four dummy variables which is equal to one if a household owns each of the following livestock assets: cattle, donkey, camel and sheep/goat; (3) a set of four dummy variables which is equal to one if the household dwelling unit is made of the following materials: mud, iron, timber, and bricks/stone; (4) a set of two dummy variables which is equal to one if the household's regular source of lighting is: electricity and paraffin; (5) a set of three dummy variables which is equal to one if the household's number of meals consumed per day are one meal, two meals and three meals; (6) a set of two dummy variables which is equal to one if the household has a source of water and a toilet at home; (7) Number of years of education of the mother and father (entered as a continuous variable).

[15]Principal components analysis (PCA) is a technique that summarizes information contained in a large number of variables in a smaller number by creating a set of mutually uncorrelated components of the data (Filmer 2005, Filmer et al. 2008). This is defined in such a way that the first principal component accounts for as much of the variability in the variables as possible. This first principal component can be interpreted as a *wealth index* on the assumption that the underlying variable with the largest explanatory power is a household's long-run wealth (Filmer 2005).

as educational attainment.[16] Despite this, we acknowledge and are aware that this measure of household resources has flaws.[17]

Following De Haan et al. (2014), we normalize the household wealth index to the range of 0 and 1 and classify households with a wealth index of 0.5 and above as rich and those below this wealth index as poor. Panel C of table 1 shows that the average wealth index (0-1 range) is 0.40, meaning that the majority of households belong to the lower part of the wealth index and can therefore be classified as poor. Only 26 percent of households can be classified as rich.[18]

Panel C (column 1 and 2) further shows that households that send children to private schools are more likely to be: smaller in size, based in urban areas, wealthier and with more young and more educated mothers. Thus, it seems, that children who attend private schools already have disproportionately higher academic potential and access to complementary educational resources. They are more likely to attend private schools, accumulate more years of education and progress faster through schooling. We use family fixed effects models to address such potential source of endogeneity.

4 Econometric Issues and Estimation Strategies

We are interested in estimating the effect of *gender* (being female) and *birth order* on three variables related to children in the households, namely, *private school enrolment, completed years of education* and *relative grade progression*. Following Maitra et al. (2016), Tenikue and Verheyden (2010), Sahoo (2016), De Haan et al. (2014) and Black et al. (2005), we estimate the following household fixed effects model by OLS as outlined in 1:

$$y_{ij} = \beta_0 + \beta_1 GIRL_{ij} + \beta_2 second_{ij} + \beta_3 third_{ij} + \beta_4 fourth_{ij} + \beta_5 fifth_{ij} + \beta_6 X'_{ij} + \psi_j + \varepsilon_{ij} \quad (1)$$

[16] The PCA based wealth index has been used by many studies (Filmer and Pritchett 2001, Tenikue and Verheyden 2010) to investigate the determinants of education outcomes in the developing world and has been found to significantly determine children education outcomes.

[17] A detailed exposition of these flaws are discussed in Vyas and Kumaranayake (2006) and more recently Wittenberg and Leibbrandt (2015). The centered PCA index has been criticized for giving negative scores to assets that are mainly owned by rural households such as ownership of livestock in the case of Kenya leading to artificial increase in the rural-urban divide (Wittenberg and Leibbrandt 2015).

[18] A household is classified as rich if its wealth index is 0.5 and above..

where y_{ij} represents the three outcomes of interest: (a) type of school (equals to one if private) that child i from household j is currently enrolled; (b) child i's completed years of education; and lastly (c) child i's relative grade progression. Each of the measures of y_{ij} depends on: (a) the child's gender, where $GIRL_{ij}$ is a dummy variable for gender being a girl; (b) child's order of birth, where $second_{ij}$, $third_{ij}$, $fourth_{ij}$ and $fifth_{ij}$ denote the second-born, third-born, fourth-born and fifth-born child, respectively (first-born being the base category); and (c) a vector of other child characteristics X_{ij} (age, a dummy variable for whether the child goes for paid up tuition and a dummy variable for whether a child has some form of disability).

As mentioned, an estimation bias may arise from the potential endogeneity of the child's gender and family size. We are able to deal with such endogeneity and remove all sources of unobserved heterogeneity common to all children in the family through the family fixed effects, denoted ψ_j. Finally, ε_{ij} is the error term.

We interpret the estimated coefficients on the dummies $second_{ij}$, $third_{ij}$, $fourth_{ij}$ and $fifth_{ij}$ as the effect of being the second, third, fourth and fifth child relative to the first child (the first child being the reference category). Our birth order variable is likely to be correlated (positively) with the age of the child since children in our sample were observed when they were still in school. Others like De Haan et al. (2014), Moshoeshoe (2016) and Tenikue and Verheyden (2010) have solved this challenge by including dummies for child's age in the regressions. We apply this same strategy. We also include two additional child level variables: a variable indicating whether a child goes for paid tuition and whether the child has some form of disability. Many studies focusing on developing countries (such as De Haan et al. (2014), Moshoeshoe (2016), Tenikue and Verheyden (2010)) do not account for these variables which may be correlated with our measured outcomes variables.

Each household in our data is considered a cluster since there are individual children within the household who are genetically linked. We cluster standard errors within the household fixed effects in order to derive homoscedastic idiosyncratic error given such clustering (Nichols and Schaffer 2007, Maitra et al. 2016).

5 Results and Discussion

5.1 Gender and Birth Order Effects

The dependent variables *completed years of education* and *relative grade progression* are continuous variables and therefore the estimation strategy used is the standard OLS with fixed effects. For private enrolment, the dependent variable is binary, indicating whether the child is enrolled in private (equals to 1) or not (equals to 0). Ideally, we should use the conventional probit or logit that bound the maximum likelihood estimated probabilities on the unit interval (Horrace and Oaxaca 2006). However, results estimated by probit or logit can be inconsistent when fixed effects are included (Baltagi 2008, Wooldridge 2003). We therefore use the Linear Probability Model (LPM) which is generally acceptable when using fixed effects in the context of a dummy dependent variable (Horrace and Oaxaca 2006).

The LPM, however, comes with two potential challenges. First, its predicted probabilities are not bounded on the unit interval (for example, they can lie outside the 0 to 1 range) and second is the challenge of heteroscedasticity. To deal with heteroscedasticity, we use the robust standard errors, which are clustered at the household level. Regarding the first challenge, Horrace and Oaxaca (2006) show that this bias increases with an increase in the relative share of predicted probabilities falling outside the unit interval. They argue that if no predicted probabilities, or just a few predicted probabilities, lie outside the unit interval, then the LPM is expected to be unbiased and consistent. In our case, the proportion of LPM predicted probabilities that lie outside the unit interval is, in fact, zero.

In table 2, we present the OLS and household fixed effects regression results for the effects of gender and birth order on the three dependent variables.[19] Throughout this paper, we refer to the results in table 2 as our *headline results/regression.* Since parents of first-born children are more likely to be younger than those of fourth or fifth born children, we include in the OLS regression the mother's age to control for such cohort effects for the parents (Black et al. 2005). In addition, we include mother's education, and a full set of dummies for the number of children in the family, which control for the correlation between birth order and family size.

[19] In the interest of brevity we only report estimates of our variables of interest (gender and birth order) throughout this paper.

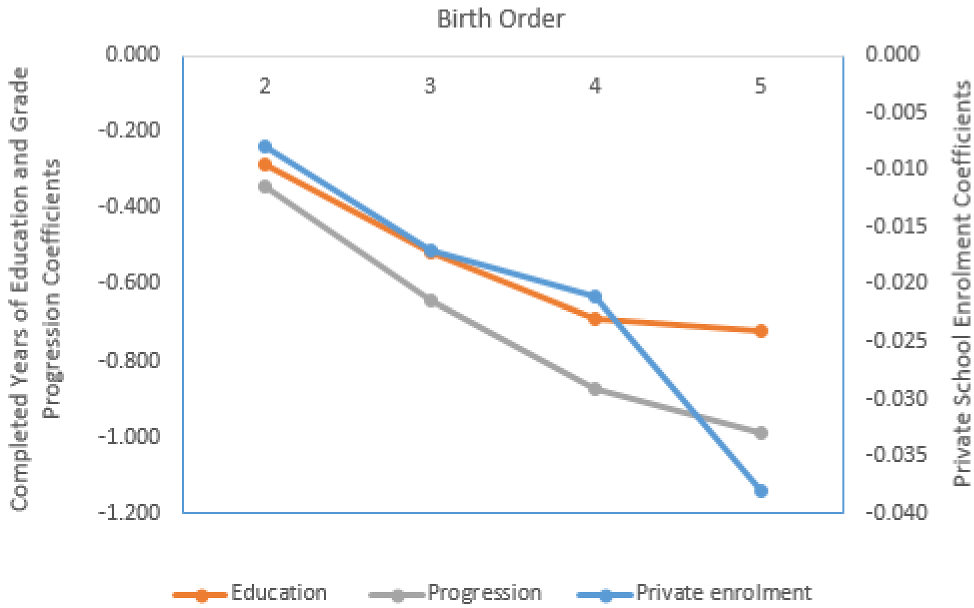
Table 2: Effects of Gender and Birth Order on Private School Enrolment, Completed Years of Education and Relative Grade Progression

	(1)	(2)	(3)	(4)	(5)	(6)
	Private School enrolment		Completed Years of Education		Relative Grade Progression	
	OLS	Fixed Effects	OLS	Fixed Effects	OLS	Fixed Effects
Female	-0.002	0.001	0.146***	0.138***	0.028***	0.025***
	(0.002)	(0.002)	(0.008)	(0.009)	(0.002)	(0.002)
Second born	-0.009***	-0.008**	-0.087***	-0.205***	-0.019***	-0.061***
	(0.003)	(0.003)	(0.011)	(0.015)	(0.002)	(0.004)
Third born	-0.020***	-0.013**	-0.165***	-0.403***	-0.048***	-0.124***
	(0.005)	(0.005)	(0.018)	(0.026)	(0.004)	(0.006)
Fourth child	-0.028***	-0.017**	-0.188***	-0.541***	-0.079***	-0.183***
	(0.007)	(0.008)	(0.026)	(0.038)	(0.007)	(0.010)
Fifth born	-0.050***	-0.032***	-0.120***	-0.594***	-0.138***	-0.271***
	(0.012)	(0.012)	(0.039)	(0.055)	(0.018)	(0.019)
Family Size						
Three children	-0.016***		-0.028**		-0.006**	
	(0.004)		(0.013)		(0.003)	
Four children	-0.022***		-0.020		0.004	
	(0.004)		(0.019)		(0.003)	
Five children	-0.018***		-0.076**		0.008	
	(0.007)		(0.030)		(0.005)	
Constant	0.322***	0.328***	-5.510***	-3.979***	-0.412***	-0.020
	(0.028)	(0.017)	(0.093)	(0.085)	(0.030)	(0.026)
Observations	83,652	83,652	84,622	84,622	80,389	80,389
R-squared	0.035	0.039	0.747	0.823	0.128	0.119
No. of households	37,282		37,432		36,272	

Notes: (1) The OLS and fixed effects regressions include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not. In addition, the OLS regression includes mother's level of education and mother's age (to control for parental cohort effects), dummies for the family size; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education the child had completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

Looking at the results in table 2, the OLS estimates in column 1 shows that being a female child does not have an effect on the probability of being enrolled in a private school. Results in columns 3 and 5 show that relative to their male siblings, female siblings are more likely to complete more years of education and progress faster through school. With regards to birth order, the results reveal a clear pattern of a negative and significant effects of birth order on all the dependent variables which increases in absolute terms with birth order. In figure 1, we show the effects of birth on all the measured outcomes based on the household fixed effects model (based on columns 2, 4 and 6 of table 2).

Figure 1: Effects of Gender and Birth Order on Private School Enrolment, Completed Years of Education and Relative Grade Progression based on Household fixed effects model



We refrain from giving OLS estimates a causal interpretation because of the possible bias that may arise from unobserved household characteristics. We therefore turn to the household fixed effects estimates that address these identification problems (columns 2, 4 and 6). The results generally show a similar pattern as those in OLS. There is no gender effect on our measure of intra-household investment in education (private enrolment). However, relative to their male siblings, female siblings complete 0.138 more years of education. In addition, they are more likely to progress through school faster, accumulating 0.025 more years of education per year of schooling relative to their male counterparts.

The female advantage we observe in terms of *completed years of education* and *relative grade progression* is in contrast with literature generally reported from developing countries (Garg and Morduch 1998, Gupta 1987, Maitra et al. 2016, Sahoo 2016). It is, however, in line with global trends which show that more girls are getting educated and the gender gap in education has narrowed considerably (King and Winthrop 2015). In Kenya, for example, the Uwezo surveys have consistently shown that over time, more girls than boys are enroling and progressing faster through school (Uwezo 2012, 2014).

Turning to birth order effects in the fixed effects model, we find a significant negative association between birth order and the three dependent variables. Relative to the first-born, all other later-born children are less likely to be enrolled in private schools and are likely to complete less years of education. For example, the second and fifth-born children are less likely to be enrolled in a private school by 0.8 percentage points and 3.2 percentage points, respectively, relative to their first-born sibling. Similarly, the second and fifth-born children are likely to complete 0.205 and 0.594 less years of education, respectively, relative to their first-born sibling. Furthermore, we find that first-born children progress through school much faster than their younger siblings. For instance, the second-born is likely to accumulate 0.061 less years of education per year of schooling relative to his/her first-born counterpart and this disadvantage increases to 0.271 less years of education in the case of the fifth-born.

Surprisingly, the first-born advantage we find in terms of private school attendance advantage, completed years of education and relative grade progression are generally in contrast with the findings in developing countries (as for instance Black et al. (2005), De Haan (2010), Tenikue and Verheyden (2010), Ejrnaes and Portner (2004)). The results are however in agreement with evidence in developed countries, for instance USA (Price 2008, Black et al. 2005, Booth and Kee 2009, De Haan 2010, Harkonen 2014, Kristensen and Bjerkedal 2007, Conley and Glauber 2006), Norway (Black et al. 2005, De Haan 2010), UK (Booth and Kee 2009) and Germany (Harkonen 2014). Our results on private school enrolment are consistent with Sahoo (2016) who documents a negative birth order effect on private school choice in India.

There are a number of reasons for the potential negative birth order effects we find. The first-borns, as posited by the confluence model, enjoy a higher intellectual environment which declines with entry of additional children (Zajonc and Markus 1976, Zajonc 1976). First-born children are also born into a family when limitations on the available parental resources such as finances and time are not thinly spread among many children (Blake 1981, Downey 2001, Becker 1960). Their cognitive ability and development is therefore more likely to be malleable at childhood, leading to better future outcomes (Cunha and Heckman 2007). In a developing country context like Kenya, parents can also favor the eldest child because these children are more likely to start earning income earlier, thus supplementing the family income. Although limited by data, we try to investigate if some of these factors are driving our results in section 5.6.

5.2 Robustness Checks

In this section, we carry out a number of robustness checks to determine the validity of our results presented in the headline regressions in table 2. Our robustness checks are based on the household fixed effects estimates.

School choice, whether to go to private school or not, is only observed if the child is enrolled in school. As we show in table 1, about 11 percent of children are not enrolled in school.[20] Non-enrolment can be a potential source of selectivity bias. Next, we check whether our estimates in table 1 compares with those based on a model with selection correction. As observed by Maitra et al. (2016), the standard Heckman-type selection model is not suitable for a household fixed-effects. In this regard, we closely follow Maitra et al. (2016) by estimating a selection equation defined in terms of the decision to enrol a child in school or not as shown in equation (2).

$$\begin{aligned} enrol^{*}_{ijk} = \beta_0 + \beta_1 GIRL_{ijk} + \beta_2 second_{ijk} + \beta_3 third_{ijk} + \beta_4 forth_{ijk} + \beta_5 fifth_{ijk} \\ + \beta_6 X^{'}_{ijk} + \beta\psi_{ijk} + \beta_8\phi_{ijk} + \mu_{ijk} \end{aligned} \tag{2}$$

where $enrol^{*}_{ijk}$ is the propensity to attend school (enrol) for student i in household j located in village k. Since this propensity is not observable, we only observe $enrol_{ij}$ when the child is enrolled e.g. $enrol_{ij} = 1$ if $enrol_{ij} > 0$ otherwise if the child is not enrolled then, $enrol_{ij} = 0$ if $enrol_{ij} \leq 0$. We let the propensity to enrol depend on: (a) the child's gender, where $GIRL_{ij}$ is a dummy for female child, (b) child's order of birth where as before, $second_{ij}$, $third_{ij}$, $fourth_{ij}$ and $fifth_{ij}$ denote the second-born, third-born, fourth-born and fifth-born child, respectively (first-born being the base category), (c) a vector of other child characteristics, X_{ij} (age, a dummy variable for whether the child goes for paid up tuition or not, and a dummy variable for whether a child has some form of disability or not), (d) family characteristics, ψ_{ijk}, (e) village characteristics, ϕ_{ijk}. μ_{ijk} is a random error.

Like Maitra et al. (2016), we estimate the selection equation equation (2) using a probit model after which we compute the inverse Mill's ratio, denoted as λ_{ij} for each sample child.[21]

[20] We do not know the underlining factors leading to their decision not to enrol.

[21] Recall that in equation (2) we are estimating a probit model and so we do not account for household fixed effects for reasons we have already explained.

At the second stage, we include the inverse Mill's ratio λ_{ij} in equation (1) and estimate a household fixed effects selectivity corrected model as shown in Equation (3):

$$y_{ij} = \beta_0 + \beta_1 GIRL_{ij} + \beta_2 second_{ij} + \beta_3 third_{ij} + \beta_4 forth_{ij} + \beta_5 fifth_{ij} + \beta_6 X'_{ij} + \psi_j + \lambda_{ij} + \varepsilon_{ij} \quad (3)$$

Like Maitra et al. (2016), we knowledge that this process of correction for selection may not be ideal, we simply want to test the robustness of our results. Table A.1 in the appendix shows the results of the selection equation for child enrolment based on equation (2).[22] In table 3, we present the selectivity corrected gender and birth order effects estimated from the household fixed effects model (equation (3)). The estimates do not differ from those presented in our headline regressions in table 2. We find no significant differences between female and male children in terms of private enrolment. However, there is a female advantage in terms of completed years of education and relative grade progression. Similarly, we find a significant negative association between birth order and the three dependent variables.

[22] As can be seen from the table, the child, household and community/village characteristics we control for exert a significant influence on enrolment probability.

Table 3: Effects of Gender and Birth Order using Household Fixed Effects Selectivity Corrected Model

	(1)	(2)	(3)
	Private School enrolment	Completed Years of Education	Relative Grade Progression
Female	0.001	0.156***	0.027***
	(0.002)	(0.011)	(0.003)
Second born	-0.011***	-0.317***	-0.060***
	(0.004)	(0.019)	(0.005)
Third born	-0.021***	-0.620***	-0.122***
	(0.006)	(0.034)	(0.008)
Fourth child	-0.023**	-0.907***	-0.185***
	(0.009)	(0.050)	(0.013)
Fifth born	-0.042***	-1.198***	-0.278***
	(0.014)	(0.071)	(0.024)
Mills Ratio (λ_{ij})	-0.100***	0.938***	-0.014
	(0.015)	(0.068)	(0.036)
Constant	0.533***	-6.899***	0.002
	(0.037)	(0.172)	(0.063)
Observations	55,943	56,510	48,418
R-squared	0.041	0.803	0.121
Number of households	22,221	22,289	21,588

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

Our sample is based on children aged 6 to16 years. The upper age is right censored since Uwezo targets children of the maximum age of 16. One of the assumptions we make is that all children are biologically born by the parents in the household.[23] It is possible that our *observed oldest child* might not be the *actual oldest living child* of the household. Our next robustness check addresses these concerns. According to a report by Kenya National Bureau of Statistics (KNBS) based on 2009 Kenya Demographic Health Survey (DHS), almost one-third of women (32%) in Kenya are married by age 18. The survey further reveals that the median age at first marriage is 20 for women in the age bracket 25–49, and that age

[23]We also assume that there is a low probability that there are older children living outside the household.

at marriage greatly increases with education where women with no education are likely to marry at age 17.5 while those with secondary school education are likely to marry at age 22.4 (KNBS 2014).

Based on this, we check the robustness of our results by estimating the gender and birth order effects on two samples. The first sample comprise households whose mothers had their *first-born* child when they were teenagers, for example at 18 years of age. In this sample, the mother's age ranges from 17 years to 34 years while the mean age is 28 years. The second sample constitutes households where mothers were 34 years or younger as at the time of the survey. We think that in these two samples, the first child was more likely to be less than 16 years in 2012 (the year when the survey was conducted), thus allowing us to estimate birth order effects in a more accurate manner. Results based on these two restricted sample are shown in table 4. In both cases, our estimates of gender and birth order do not deviate from those in the headline regression.

One might be concerned that family size is likely to be correlated with birth order and that this might confound our birth order estimates. Although the use of household fixed effects model helps to deal with such biases, we follow Black et al. (2005) and De Haan (2010) by estimating the effects of birth order effects by household size to account for possible endogeneity induced by household size. Results are shown in table 5. The results for gender are consistent with our headline findings: irrespective of the family size, there is no female advantage in terms private school enrolment but we find a female advantage in terms of completed years of education and relative grade progression. Similarly, there is no gender preference in terms of private school enrolment. We still find negative birth order effects on all the three dependent variables. However, the relationship is insignificant in the case of private school enrolment especially among large families.

5.3 Heterogeneous Effects of Gender and Birth Order

We are using a large, nationally representative survey with data collected from almost all regions of Kenya. Kenya reflects large inter-regional variation in human development. This is reflected in cultural and social practices as well as geography, and economics of the different regions. In this section, we test to see if the results in our headline regression reflect this heterogeneity. The regressions will be based on the household fixed effects models.

Table 4: Effects of Gender and Birth Order by Mother's Age

	(1)	(2)	(3)	(4)	(5)	(6)
	Households whose mothers had their first-borns at 18 years			**Households whose mothers were 34 years or less in 2012**		
Variables	Private School enrolment	Completed Years of Education	Relative Grade Progression	Private School enrolment	Completed Years of Education	Relative Grade Progression
Female	-0.002	0.136***	0.023***	-0.002	0.133***	0.025***
	(0.003)	(0.018)	(0.004)	(0.003)	(0.014)	(0.004)
Second born	-0.015***	-0.221***	-0.078***	-0.016***	-0.249***	-0.075***
	(0.006)	(0.030)	(0.007)	(0.005)	(0.023)	(0.006)
Third born	-0.024**	-0.455***	-0.153***	-0.023**	-0.471***	-0.157***
	(0.010)	(0.052)	(0.013)	(0.009)	(0.041)	(0.011)
Fourth child	-0.034**	-0.594***	-0.242***	-0.026*	-0.623***	-0.253***
	(0.015)	(0.074)	(0.020)	(0.014)	(0.061)	(0.017)
Fifth born	-0.081***	-0.605***	-0.322***	-0.065***	-0.645***	-0.332***
	(0.022)	(0.111)	(0.043)	(0.022)	(0.094)	(0.039)
Constant	0.348***	-3.714***	-0.089*	0.353***	-3.771***	-0.072
	(0.033)	(0.163)	(0.052)	(0.028)	(0.134)	(0.044)
Observations	22,313	22,600	21,335	33,745	34,152	31,880
R-squared	0.052	0.820	0.135	0.049	0.824	0.138
Number of households	9,815	9,862	9,805	15,441	15,512	15,375

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

Table 5: Effects of Gender and Birth Order by Household Size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Completed Years of Education				**Relative Grade Progression**				**Private School enrolment**			
	Two child family	Three child family	Four child family	Five child family	Two child family	Three child family	Four child family	Five child family	Two child family	Three child family	Four child family	Five child family
Female	0.138***	0.142***	0.143***	0.123***	0.030***	0.023***	0.026***	0.016**	-0.005	0.004	0.002	0.005
	(0.018)	(0.015)	(0.019)	(0.029)	(0.005)	(0.004)	(0.004)	(0.007)	(0.004)	(0.004)	(0.004)	(0.007)
Second-born	-0.166***	-0.205***	-0.253***	-0.268***	-0.074***	-0.048***	-0.066***	-0.054***	-0.012**	-0.014***	-0.006	-0.006
	(0.027)	(0.025)	(0.037)	(0.068)	(0.007)	(0.006)	(0.008)	(0.013)	(0.005)	(0.005)	(0.007)	(0.011)
Third-born		-0.516***	-0.653***	-0.761***		-0.139***	-0.138***	-0.125***		-0.001	-0.006	-0.015
		(0.044)	(0.059)	(0.106)		(0.011)	(0.014)	(0.023)		(0.007)	(0.009)	(0.015)
Fourth-born			-1.116***	-1.242***			-0.271***	-0.200***			0.021*	-0.007
			(0.084)	(0.142)			(0.020)	(0.032)			(0.012)	(0.020)
Fifth-born				-1.659***				-0.352***				0.002
				(0.179)				(0.042)				(0.026)
Constant	-4.121***	-3.536***	-2.069***	-1.549***	0.008	0.019	0.346***	0.427***	0.336***	0.255***	0.243***	0.153*
	(0.160)	(0.159)	(0.236)	(0.426)	(0.051)	(0.047)	(0.068)	(0.123)	(0.033)	(0.033)	(0.046)	(0.091)
Observations	31,184	28,935	17,505	6,998	29,426	27,589	16,695	6,679	30,842	28,610	17,292	6,908
R-squared	0.935	0.827	0.826	0.833	0.768	0.129	0.128	0.142	0.018	0.009	0.006	0.005
No. of households	15,391	11,490	5,346	1,768	15,391	11,448	5,335	1,767	15,391	1,838	5,326	1,761

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

5.3.1 Gender and Birth Order Effects by Family Size

We begin by exploring differences in the effects of gender and birth order by household size. Table 5 has already been explored as part of the robustness tests. We use it once more to show if there are heterogeneous effects in gender and birth order effects by family size. The results on gender are consistent with our headline results: in all the different family sizes, there is no significant differences between female and male children in terms of private enrolment. However, there is a female advantage in terms of completed years of education and relative grade progression.

It is more revealing to analyze birth order effects in table 5 by column (*within families)* and by row *(across families)*. Column 1 to 4 shows the effects of gender and birth order on *years of schooling completed.* Reading results by column, we see significant negative birth order effects irrespective of the family size. Reading results by row, we also find that birth order effects increase *across* the family sizes as we move from a family of two to five children. Put differently, the disadvantage of being the second, third and fourth-born increases with an increase in the family size. For example, the *second-born in a family of two children* completes 0.166 fewer years of education relative to the first-born sibling while the counterpart *second born in a family of five* completes 0.268 fewer years of education relative to the first-born sibling (row 2). These results generally support the theoretical predictions of the confluence (Zajonc and Markus 1976, Zajonc 1976), resource hypothesis (Downey 2001) and quantity-quality (Becker 1960) models which predict a negative relationship between child outcomes and family size.

Similarly, results in column 5 to 8 (read by column) shows significant negative birth order effects on *relative grade progression* irrespective of the family size in line with theoretical predictions. However, results (read by row) shows birth order effects on *relative grade progression* reduce *across* family sizes. The disadvantage (gap) of being the second, third and fourth-born, *relative to the first-born*, reduces with an increase in the family sizes (see column 5 to 8). For example, *relative to the first-born,* a *second-born child in a family of two children* accumulates 0.074 fewer years of education while the *same second-born in a five-child family* accumulates 0.054 fewer years of education. This result seems to be in line with the *tutoring effect* of the confluence model (Zajonc and Markus 1976, Zajonc 1976) which posits that large families can narrow latter-born disadvantage in child outcomes. We think that latter-born siblings in large families are likely to progress faster through school for two potential reasons: they are likely to *start school on time* by accompanying their older siblings who are already in school and are also likely to develop skills by tutoring their younger siblings or being tutored by older siblings.

Next, we test for the equality of the gender and birth order coefficients *across family sizes*. To do so, we follow De Haan et al. (2014) and Moshoeshoe (2016) by estimating a fully interacted model where every variable in the family fixed effects model, not just gender and birth order, is interacted with family size. In the next sub-section, we briefly explain the intuition behind fully interacted models. Results are reported in table 6. The coefficient of the interaction between gender and family size is insignificant in all the outcomes variables. This means that an increase in family size is not associated with a female advantage in terms of private school enrollment, completed years of education and relative grade progression.

The interaction term between birth order and private school enrolment variable is insignificant showing the absence of significant differences in the effects of birth order on private school enrolments across family sizes. The interaction terms between birth order and family size are negative and significant in the completed years of education regression confirming results in table 5 that the effects of birth order on completed years of education get larger as the size of the family increases. However, birth order interaction terms with relative grade progression are generally positive and significant again confirming results in table 5 that birth order effects on relative grade progression get smaller as the size of the family increases.

Table 6: Effects of Gender and Birth Order: Fully Interacted Models with Family Size

	(1)	(2)	(3)
	Private School enrolment	Completed Years of Education	Relative Grade Progression
Female	-0.004	0.147***	0.036***
	(0.006)	(0.031)	(0.008)
Second born	-0.018*	-0.092*	-0.071***
	(0.011)	(0.052)	(0.013)
Third born	-0.017	-0.303***	-0.222***
	(0.021)	(0.101)	(0.024)
Fourth child	-0.011	-0.511***	-0.461***
	(0.043)	(0.196)	(0.056)
Fifth born	0.043*	-1.614***	-0.360***
	(0.023)	(0.111)	(0.029)
Female* family Size	0.001	-0.002	-0.003
	(0.002)	(0.010)	(0.002)
Second-born* family Size	0.004	-0.038**	0.003
	(0.003)	(0.017)	(0.004)
Third-born* family Size	0.005	-0.083***	0.022***
	(0.006)	(0.031)	(0.007)
Forth-born* family Size	0.008	-0.137***	0.052***
	(0.011)	(0.051)	(0.014)
Fifth-born* family Size	–	–	–
	–	–	–
Constant	0.275***	-3.279***	0.113***
	(0.020)	(0.097)	(0.029)
Observations	83,652	84,622	80,389
R-squared	0.040	0.824	0.123
Number of households	36,329	36,501	36,272

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\,years\,of\,education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

5.3.2 Gender and Birth Order Effects by Location

Family size may matter more or less depending on whether the family is located in a rural or urban area. In a country like Kenya, a large household may be a source of labour on the farm, hence more beneficial in rural than urban areas. Also, urban areas in Kenya have more education opportunities, especially high concentration of private schools than rural areas (Piper and Mugenda 2010, Piper et al. 2014). For these reasons, we check for heterogeneity in gender and birth order effects by location (rural versus urban) and test if our results in the main regression are an artefact of household location. We estimate our household fixed effects model by location whose results are shown in table 7. The results on gender, in both rural and urban areas, are consistent with our headline regressions table 2: we do not find significant differences between female and male children in terms of private enrolment in both rural and urban locations but there is a female advantage in terms of completed years of education and relative grade progression.

Interestingly, we find significant negative birth order effects on private school enrolment in the rural sample but none in the urban sample (column 1 and 4). The prevalence of private schools in Kenya is quite high in urban areas relative to rural areas. For some families in urban informal settlements (especially in Nairobi), the choice is sometimes not between a government primary school and a non-formal private school, but between the non-formal private school and no school at all (Piper et al. 2014). With such near universal access to private schools in urban areas, there are no incentives for intra-household discrimination among siblings thus explaining the lack of significant birth order effects on private school enrolment in urban areas. On the other hand, private schools are not as prevalent in rural areas as they are in urban areas. The few private schools in rural areas are likely to charge higher fees thus inducing intra-household discrimination among children in favor of older siblings in terms of their access.

Table 7 further shows that birth order effects are different between rural and urban areas in terms of completed years of education (column 2 and 5) and relative grade progression (column 3 and 6). In particular, there is strong latter-sibling disadvantage in terms of completed years of education in rural households. For instance, a fifth-born child in a rural household is likely to complete 0.721 fewer years of education relative to their first-born sibling while a fifth-born child in an urban household is likely to complete 0.327 fewer years of education relative to their first-born sibling. Here, the disadvantage facing the fifth-born, relative to her first-born counterpart, in a rural household, is higher than that of a fifth-born in an urban household.

Table 7: Effects of Gender and Birth Order by Rural and Urban

Variables	(1) Rural Areas: Private School enrolment	(2) Rural Areas: Completed Years of Education	(3) Rural Areas: Relative Grade Progression	(4) Urban Areas: Private School enrolment	(5) Urban Areas: Completed Years of Education	(6) Urban Areas: Relative Grade Progression
Female	-0.001	0.164***	0.031***	0.004	0.074***	0.009**
	(0.002)	(0.011)	(0.003)	(0.005)	(0.020)	(0.005)
Second born	-0.008**	-0.216***	-0.060***	-0.007	-0.174***	-0.060***
	(0.003)	(0.018)	(0.004)	(0.008)	(0.033)	(0.008)
Third born	-0.015***	-0.441***	-0.126***	-0.002	-0.314***	-0.119***
	(0.006)	(0.031)	(0.008)	(0.014)	(0.057)	(0.014)
Fourth child	-0.019**	-0.620***	-0.185***	-0.015	-0.335***	-0.182***
	(0.009)	(0.045)	(0.011)	(0.021)	(0.083)	(0.021)
Fifth born	-0.036***	-0.721***	-0.280***	-0.016	-0.327***	-0.295***
	(0.013)	(0.063)	(0.022)	(0.032)	(0.126)	(0.043)
Constant	0.301***	-3.796***	0.025	0.459***	-4.401***	-0.073
	(0.019)	(0.100)	(0.031)	(0.046)	(0.177)	(0.057)
Observations	60,841	61,563	58,535	17,020	17,172	16,280
R-squared	0.041	0.820	0.115	0.035	0.839	0.132
No. of households	26,178	26,305	26,131	7,574	7,599	7,559

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\,years\,of\,education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

Similarly, to formally test for the equality of the effects of gender and birth order between rural and urban areas, we estimate a fully interacted model by interacting every variable with the rural location dummy.[24] Results are shown in table 8. Since this is a fully interacted family fixed effects model with *a rural dummy*, the coefficients on gender (female) and birth order dummies in table 8 are for urban areas. Notice that the coefficients in the table are identical to those for urban areas presented in table 7 (columns 4, 5 and 6). Second, the interaction terms in table 8 capture the difference between rural and urban location on each of the model parameters (Gordon 2015). A significant interaction term on the variable means that the effect of the variable on the measured dependent variable is significant between rural

[24]By a fully interacted model, the rural location dummy is interacted with our variables of interest (gender and birth order dummies) as well as all other variables (e.g. age, disability and tuition attendance status) in the family fixed effects model. In doing so, we allow all the parameters of the model including the intercept to vary by location (Gujarati 1970b,a, Gordon 2015). Two most recent studies that have investigated heterogeneity in birth order effects using fully interacted family fixed effects models are De Haan et al. (2014) and Moshoeshoe (2016).

and urban. The interaction terms therefore provide a formal test of equality of gender and birth order effects between rural and urban areas (see Gordon (2015), Gujarati (1970a) and Gujarati (1970b) for theoretical foundations of fully interacted models and De Haan et al. (2014) and Moshoeshoe (2016) for empirical application).

Table 8: Effects of Gender and Birth Order: Fully Interacted Models with Rural Dummy

	(1)	(2)	(3)
	Private School enrolment	Completed Years of Education	Relative Grade Progression
Female	0.004	0.074***	0.009
	(0.006)	(0.027)	(0.006)
Second born	-0.007	-0.174***	-0.060***
	(0.010)	(0.043)	(0.011)
Third born	-0.002	-0.314***	-0.119***
	(0.019)	(0.075)	(0.019)
Fourth child	-0.015	-0.335***	-0.182***
	(0.028)	(0.110)	(0.029)
Fifth born	-0.016	-0.327*	-0.295***
	(0.043)	(0.167)	(0.058)
Female*rural	-0.005	0.090***	0.021***
	(0.007)	(0.030)	(0.007)
Second-born*rural	-0.001	-0.042	0.000
	(0.011)	(0.049)	(0.012)
Third-born*rural	-0.013*	-0.127*	-0.007
	(0.007)	(0.086)	(0.021)
Forth-born*rural	-0.005**	-0.285**	-0.003
	(0.002)	(0.125)	(0.033)
Fifth-born*rural	-0.020**	-0.394**	0.015
	(0.010)	(0.187)	(0.065)
Constant	0.336***	-3.928***	0.004
	(0.018)	(0.087)	(0.027)
Observations	77,861	78,735	74,815
R-squared	0.791	0.909	0.676
No. of households	36,329	36,501	36,272

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

As can be seen in table 8, the interaction term between female dummy and rural dummy is significant only in the completed years of education regression ((column 2). This means that there is a strong female advantage in terms of completed years of education in rural than urban areas. However, there is no evidence of significant female advantage in terms of private school enrolment and relative grade progression between rural and urban areas. The interaction terms between birth order and rural dummy are negative and statistically significant in the private school enrolment and completed years of education regression. This means there are stronger negative birth order effects in terms of private school enrolment and completed years of education in rural than urban areas.

5.4 Possible Explanation for the Gender and Birth Order Effects

We now examine the possible explanation for the gender and birth order effects observed. In particular, our data allows us to test if the effects are propagated through the resource hypothesis or birth-spacing (or child-spacing) channel. All the estimates in this section are based on the family fixed effects model.

5.5 Gender, Birth Order Effects and Household Wealth

We begin by testing to see if the effects of gender and birth order are propagated through household wealth. In order to do this, we use the family wealth index constructed in 3.2. A number of studies have used a similar index to test the mechanisms through which birth order effects are propagated. A study by Tenikue and Verheyden (2010) uses a similar wealth index based on the Demographic and Health Survey data to examine birth order effects between siblings in rich and poor families in terms of schooling and child labor in 12 sub-Saharan Africa countries. They find that the education levels of earlier born children are lower than their later born siblings in poor households, whereas earlier-born children are more educated in richer ones.

There are a number of reasons why latter-born children do worse. The resource hypothesis (Downey 2001) and quantity-quality (Becker 1960) models argue that investment in children increases at higher levels of economic status. They argue that as the size of the family increases, per capita familial resources reduce thus reducing investment in children. If constraints on household resources are to some extent responsible for the negative effect of birth order as predicted by theoretical models, we might expect the negative effect of birth order to *be attenuated or even reversed* among rich households. Evidence in developing countries further shows that girls are likely to suffer more in resource constrained households (Garg and Morduch 1998, Gupta 1987). Therefore, we might expect a girl preference in rich households.

Table 9: Effects of Gender and Birth Order by Household Wealth

	(1)	(2)	(3)	(4)	(5)	(6)
	Households Classified as Poor			**Households Classified as Rich**		
	Private School enrolment	Completed Years of Education	Relative Grade Progression	Private School enrolment	Completed Years of Education	Relative Grade Progression
Female	-0.001	0.141***	0.026***	0.003	0.123***	0.022***
	(0.002)	(0.014)	(0.003)	(0.006)	(0.024)	(0.006)
Second-born	-0.001	-0.239***	-0.065***	-0.023**	-0.137***	-0.048***
	(0.004)	(0.023)	(0.006)	(0.010)	(0.038)	(0.010)
Third-born	0.000	-0.512***	-0.132***	-0.038**	-0.219***	-0.096***
	(0.007)	(0.040)	(0.010)	(0.018)	(0.067)	(0.017)
Fourth-born	-0.001	-0.719***	-0.193***	-0.034	-0.307***	-0.153***
	(0.011)	(0.058)	(0.015)	(0.027)	(0.098)	(0.027)
Fifth-born	-0.001	-0.859***	-0.282***	-0.088**	-0.208	-0.237***
	(0.017)	(0.083)	(0.030)	(0.039)	(0.143)	(0.057)
Constant	0.241***	-3.314***	-0.025	0.519***	-5.238***	-0.022
	(0.023)	(0.130)	(0.042)	(0.057)	(0.213)	(0.069)
Observations	62,055	62,843	59,798	21,597	21,779	20,591
R-squared	0.772	0.903	0.671	0.800	0.923	0.634

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

In table 9, we present the gender and birth order estimates for poor and rich households.[25] As can be seen from table 9, the results on the effect of gender on measured outcomes is similar to those in the headline regression- whether in poor or rich households, we find no presence of female advantage in terms of private school enrolments but there is a female advantage in terms of completed years of education and relative grade progression.

[25]Following De Haan et al. (2014), we normalized the family wealth index to the range of 0 and 1 and classified households with a wealth index of 0.5 and above as rich and vice versa. We have confirmed that the results do not significantly change even when we use the family wealth index in its original form without normalization.

Table 10: Effects of Gender and Birth Order: Fully Interacted Models with Rich Family Dummy

	(1) Private School enrolment	(2) Completed Years of Education	(3) Relative Grade Progression
Female	-0.001	0.141***	0.026***
	(0.002)	(0.014)	(0.003)
Second-born	-0.001	-0.239***	-0.065***
	(0.004)	(0.023)	(0.006)
Third-born	0.000	-0.512***	-0.132***
	(0.007)	(0.041)	(0.010)
Fourth-born	-0.001	-0.719***	-0.193***
	(0.011)	(0.059)	(0.015)
Fifth-born	-0.001	-0.859***	-0.282***
	(0.017)	(0.084)	(0.030)
Female*Rich household	0.003	-0.018	-0.004
	(0.007)	(0.027)	(0.007)
Second born*Rich household	-0.022**	0.101**	0.017
	(0.011)	(0.044)	(0.011)
Third born*Rich household	-0.039**	0.293***	0.036*
	(0.019)	(0.078)	(0.019)
Fourth child*Rich household	-0.033	0.412***	0.040*
	(0.028)	(0.113)	(0.020)
Fifth born*Rich household	-0.087**	0.651***	0.044
	(0.042)	(0.164)	(0.063)
Constant	0.313***	-3.809***	-0.024
	(0.017)	(0.084)	(0.026)
Observations	83,652	84,622	80,389
R-squared	0.045	0.827	0.119
No. of Households	36,329	36,501	36,272

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\, years\, of\, education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

However, we see that among poor households (column 1), there are no significant birth order effects on private school enrolment. However, in rich households (column 4), young siblings are significantly less likely to attend private schools. For instance, a fifth-born child in a poor family has equal chance of attending a private school (relative to her eldest sibling) (column 1) while her counterpart in a rich family is 8.8 percentage point less likely to attend a private school (relative to her eldest sibling) (column 4). The results suggest that family wealth *intensifies* rather than *attenuates* the negative effect of birth order on private enrolment. This is surprising result that is contrary to our expectation. This is because attending a private school in Kenya involves some costs and if our measure of household wealth is accurate, we expect *family wealth* to *lessen* (or even reverse) rather than *worsen* the latter-born disadvantage in terms of private school access. We return to explain this peculiar result shortly.

In line with our theoretical predictions, table 9 shows that higher family wealth *attenuates* the negative effect of birth order on completed years of education and relative grade progression. The effects of birth order on the number of completed years of schooling are two times larger among poor households relative to rich households (columns 2 and 5). Similarly, birth order effects on relative grade progression are relatively higher among poor households relative to rich households (columns 3 and 6).

In table 10, we show results from a fully interacted family fixed effects model with a dummy variable indicating that the family is rich. The interaction terms between female dummy and a dummy variable for the rich household are insignificant in all the regressions. This means that family wealth does not have a gender effect (female advantage) in terms of private school enrolment, completed years of schooling and relative grade progression. The interaction of birth order effects and a dummy variable for the rich household are largely significant in all the dependent variables. The interaction terms with private school enrolments are negative (column 1) confirming results in table 9 that family wealth indeed *intensifies* rather than *attenuates* the negative effect of birth order on *private enrolment.* The interaction terms with completed years of education and relative grade progression (column 2 and 3) are positive showing that family wealth *attenuates* the negative effect of birth order on *completed years of education* and *relative grade progression.*

One might be concerned that our results are an artifact of our classification of poor and rich households. For instance, we normalized the family wealth to a range of 0 to 1 and classified households whose index was 0.5 and above as rich. Our results might be an artifact of such a cut-off point. To allay such fears, following De Haan et al. (2014), we allow the wealth index, in its continuous form, to be fully interacted with the gender, birth order dummies and all control variables. Results are shown in table A.2 in the appendix. Generally, the results do not deviate from those in table 9. The interaction terms between female dummy and family wealth in table A.2 are largely insignificant.[26] On the other hand, the interaction term with birth order dummies are largely significant and further show that wealth *amplifies* the effects of birth order on private school enrolment and *attenuates* birth order effects on completed years of education and relative grade progression.

5.6 Family wealth and the latter-born disadvantage in private school enrolment

As promised, we return to the results in column 1 and 2 of table 9. Contrary to our expectation, there are no birth order effects among poor households and negative birth order effects among rich households. We maintain that attending a private school involves some costs and therefore *family wealth* should *lessen* rather than *worsen* the latter-born disadvantage in terms of private school access. To disentangle these results, we show the effects of birth order on private school enrolment by *region (rural and urban)* and *household wealth (poor and rich)* as shown in table 11. Recall that we attributed the lack of birth order effects in urban areas in table 7 (column 4) to the near universal access to private schools in urban areas. We further argued that the scarcity of private schools in rural areas could be driving up their prices thus inducing households to discriminate against some children in terms of private school access (see table 7 column 1).

Looking in column 1 and 2 of table 11, we see no birth order effects among both rich and poor houses in urban areas, consistent with results in table 7 (column 4). This is also consistent with our hypothesis that the near universal access to private schools in urban areas could have indeed eroded incentives for households to discriminate against some children in terms of their access. It therefore follows that the *negative birth order effects* we see in column 4 of table 9 among *rich households* is mainly driven by *rural rich households*. Column 3 of table 11 confirms this assertion: we see a significant negative association between birth order and private enrolment among children from rich households in rural areas. In column 4 of table 11, we do not see evidence of birth order effects among rural poor. This lack of birth order effects among rural poor combined with the lack of birth order effects among the urban

[26]Nonetheless, the interaction term between female dummy and family wealth is significant but only at 5 percent in the completed years of education regression showing that family wealth increases female advantage in terms of completed years of education.

poor in columns 2 of the same table means that the general lack of birth order effects on private school enrolments in table 9 (column 1) is being driven by both rural and urban poor households.

Table 11: Effects of Gender and Birth Order by Region and Household Wealth

	Private School enrolment			
	(1)	(2)	(3)	(4)
	Urban Areas		**Rural Areas**	
	Rich Household	Poor Household	Rich Household	Poor Household
Female	-0.001	0.008	0.002	-0.001
	(0.011)	(0.008)	(0.007)	(0.003)
Second born	-0.004	-0.007	-0.037***	0.000
	(0.018)	(0.012)	(0.012)	(0.005)
Third born	0.011	-0.004	-0.069***	0.002
	(0.034)	(0.022)	(0.021)	(0.008)
Fourth child	0.017	-0.020	-0.071**	0.001
	(0.050)	(0.031)	(0.032)	(0.012)
Fifth born	0.033	-0.030	-0.163***	0.002
	(0.075)	(0.050)	(0.045)	(0.018)
Constant	0.542***	0.350***	0.501***	0.235***
	(0.104)	(0.069)	(0.070)	(0.026)
Observations	8,022	8,998	12,436	48,405
R-squared	0.804	0.784	0.790	0.769

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

We can now speculate what is driving the peculiar results in column 1 and 4 of table 9. The near universal access of private schools in urban areas means that all children, from both rich and poor households, have a higher chance of accessing them. On the other hand, we think that the few private schools in rural areas seem to be accessible only by rich households. However, these households seem to be constrained due to the relatively high fees charged by these private schools and as a result, it is the first-borns who are favored. For the rural poor households, we think that these rural private schools/academies are almost non-accessible to them mainly due due to affordability. Put differently, for the rural poor, it is not a question of intra-household discriminating against some children in terms of access to private schools but rather a general lack of access to private schools in rural areas due to costs.

To further show that the rural poor have almost no access to private schools, we estimate the probability of private school enrolment (using a probit model) while accounting for among others, the interaction between a rural dummy and a dummy denoting whether a child is from a poor household. Results are shown in table A.3 in the appendix. As it can be seen from the table, results show that children from rural poor households are indeed less likely to be enrolled in private schools.

In summary, our results show that private schools are highly concentrated in urban areas and as a result, they are accessible to all children, regardless of economic status. In rural areas, private schools are scarce and hence charging high fees. The high cost of attending private schools in rural areas seems to have, locked out the poor households and induced an intra-household child-discrimination among the rural households working in favour of first-born children.

6 Conclusion

In this paper, we examine the effect of gender and birth order on intra-household education investment in children and the resultant educational outcomes in Kenya. We measure intra-household education investment in children by the household's decision to enrol children in private schools and educational outcomes by two variables: completed years of education and relative grade progression. We use the Uwezo survey and implement the household fixed effects models that control for potential endogeneity of gender, birth order, family size and the household level unobserved heterogeneity. We find no intra-household gender preference in terms of whether or not to enrol a child in private school. However, there is a consistent female advantage in terms of completed years of education and relative school progression. Regarding birth order effects, the results reveal a clear pattern of significant negative effects of birth order on private school enrolment, completed years of education and relative school progression.

The female advantage we find (in terms of *completed years of education* and *relative grade progression)* is generally not in consistent with literature reported from developing countries but in line with global trends which show that more girls are getting educated and the gender gap in education has narrowed considerably. Similarly, the first-born advantage we find seems to be in line with the findings in developed countries but not developing countries. Nevertheless, these results generally support the theoretical predictions of the confluence, resource hypothesis and quantity-quality models which predict a first-born advantage. Our results are robust to different robustness checks including correction of selectivity bias originating from non-enrolment of children and further attempts to measure birth order effects more accurately. We check the heterogeneity of the gender and birth order effects and find that these effects significantly differ by family size and by location of the family (rural or urban).

We end the discussion by testing for the drivers of our observed gender and birth order effects. We find evidence to support the hypothesis that our birth order effects (but not gender effects) are driven household wealth. Following De Haan (2010) and Moshoeshoe (2016), we also check if the gender and birth order effects are driven by child spacing. However, we do not report these results since the coefficients for the interaction of gender and birth order with average child spacing are effectively zero.

There are implications for policymakers from the finding that latter-born children, especially in poor and large households, lag behind in terms of education outcomes. This call for efforts to sensitive the population on the importance of family planning. There is also need to institute interventions such as cash transfer and other financial assistance to large and poor families. Given that its is latter-born children who are most disadvantaged, such support systems should be designed to improve the conditions of latter-born children.

7 Appendix

Table A.1: Probit Results for Children Enrolment

	Coefficient	Standard Error
Student Characteristics		
Birth order (ref: First child)		
Second child	-0.066**	(0.026)
Third child	-0.203***	(0.031)
Fourth child	-0.285***	(0.038)
Fifth child	-0.408***	(0.054)
Age	1.182***	(0.020)
Age squared	-0.050***	(0.001)
Child is female	0.008	(0.016)
Child has some disability	-0.491***	(0.041)
Child goes for paid tuition		
Household Characteristics		
Mother's age	0.029***	(0.006)
Mother's age squared	-0.000***	(0.000)
Mother's Education level(ref: None)		
Has primary education level	0.196***	(0.026)
Has secondary education level	0.246***	(0.037)
Has post secondary education level	0.264*	(0.135)
Father's Education level(ref: None)		
Has primary education level	0.246***	(0.027)
Has secondary education level	0.375***	(0.034)
Has post secondary education level	0.381***	(0.084)
Household size	0.005	(0.004)
Household has source of water at home	0.077***	(0.028)
Household has toilet/latrine at home	0.332***	(0.021)
Distance to school	-0.191***	(0.033)
Index of household Assets	0.055***	(0.008)
Meals taken per day(ref: Less than three meals)		
Three meals	0.102***	(0.018)
Wall material for dwelling place (ref: Mud)		
Polythene and iron	0.220***	(0.035)
Timber	0.240***	(0.037)
Bricks_stone	0.150***	(0.027)
Regular source of lighting (ref: Paraffin)		
Electricity/solar/gas	0.033	(0.034)
Other	-0.324***	(0.026)
Village Characteristics		
Village has chief's office	0.066***	(0.018)
Village has shopping center	0.030	(0.021)
Village has electricity	-0.010	(0.021)
Village has tarmac road	-0.014	(0.026)
Village has all-weather road	0.031	(0.019)
Village has an education committee	0.033*	(0.018)
Village has all protected water point	-0.002	(0.017)
Village is rural	-0.018	(0.024)
Constant	-6.311***	(0.158)
Observations	57,631	

Notes: (1) We use the ordinary principal component analysis (PCA) to construct the index of household assets. The index is based on household ownership of the following assets: durable assets (TV, radio, car, computer, mobile phone, bicycle, motorbike and cart) and livestock assets (cattle, donkey, camel, sheep/goat); (2) Standard errors are in parenthesis; (3) ***1% significance level, **5% significance level and *10% significance level.

Table A.2: Effects of Gender and Birth Order: Fully Interacted with Family Wealth, as a Continuous Variable

	(1)	(2)	(3)
	Private School enrolment	Completed Years of Education	Relative Grade Progression
Female	-0.007	0.076**	0.016*
	(0.007)	(0.033)	(0.008)
Second born	0.015	-0.388***	-0.095***
	(0.011)	(0.056)	(0.014)
Third born	0.032	-0.854***	-0.187***
	(0.020)	(0.099)	(0.024)
Fourth child	0.024	-1.265***	-0.270***
	(0.029)	(0.142)	(0.037)
Fifth born	0.080*	-1.685***	-0.361***
	(0.044)	(0.200)	(0.072)
Female*Family wealth	0.018	0.142*	0.024
	(0.018)	(0.074)	(0.018)
Second born*Family wealth	-0.057*	0.425***	0.084***
	(0.029)	(0.123)	(0.031)
Third born*Family wealth	-0.099*	0.989***	0.153***
	(0.050)	(0.219)	(0.053)
Forth child*Family wealth	-0.078	1.530***	0.212**
	(0.075)	(0.316)	(0.082)
Fifth born*Family wealth	-0.246**	2.329***	0.216
	(0.110)	(0.455)	(0.169)
Constant	0.308***	-3.741***	-0.022
	(0.022)	(0.110)	(0.036)
Observations	83,652	84,622	80,389
	0.792	0.911	0.677
	36,329	36,501	36,272

Notes: (1) All regressions are based on the household fixed effects model and include the following control variables: age of the child (in years), whether the child has some form of disability or not and whether the child goes for paid up tuition or not; (2) We only show estimates for gender and birth order dummies (our variables of interest); (3) Private school enrolment is defined as a dummy which equals 1 if the child was enrolled in a private school or 0 if a child was enrolled in a public school and this is estimated using LPM; (4) Completed years of education is number of years of education completed as at the time of the survey; (4) Relative grade progression equals $\frac{Completed\,years\,of\,education}{(Age-6)}$ where 6 is the school starting age in Kenya; (5) We clustered the standard errors (in parenthesis) at the household level; and (6) ***significant at 1%, **significant at 5%, *significant at 10%.

Table A.3: Probit Results for Children Private School Enrolment

	Coefficient	Standard Error
Rural	-0.099***	(0.029)
Poor	-0.042	(0.042)
Rural*poor	-0.075*	(0.038)
Student Characteristics		
Birth order (ref: First child)		
Second child	-0.087***	(0.021)
Third child	-0.177***	(0.028)
Fourth child	-0.218***	(0.040)
Fifth child	-0.351***	(0.078)
Age	-0.105***	(0.022)
Age squared	0.001	(0.001)
Child is female	0.017	(0.016)
Child has some disability	-0.087*	(0.052)
Child attends paid tuition	-0..433***	(0.018)
Household Characteristics		
Mother's age	0.006	(0.007)
Mother's age squared	-0.000**	(0.000)
Mother's Education level(ref: None)		
Has primary education level	-0.252***	(0.028)
Has secondary education level	-0.035	(0.033)
Has post secondary education level	0.136*	(0.074)
Father's Education level(ref: None)		
Has primary education level	-0.259***	(0.030)
Has secondary education level	-0.098***	(0.033)
Has post secondary education level	-0.028	(0.055)
Household size	0.005	(0.004)
Household has source of water at home	0.118***	(0.021)
Household has toilet/latrine at home	-0.032	(0.024)
Distance to school	-0.040	(0.037)
Index of household Assets	0.053***	(0.008)
Meals taken per day(ref: Less than three meals)		
Three meals	0.152***	(0.021)
Wall material for dwelling place (ref: Mud)		
Polythene and iron	0.105***	(0.030)
Timber	-0.043	(0.031)
Bricks_stone	0.040	(0.025)
Regular source of lighting (ref: Paraffin)		
Electricity/solar/gas	0.374***	(0.025)
Other	-0.026	(0.037)
Village Characteristics		
Village has chief's office	0.003	(0.018)
Village has shopping center	0.160***	(0.019)
Village has electricity	0.001	(0.019)
Village has tarmac road	0.088***	(0.021)
Village has all-weather road	-0.092***	(0.019)
Village has an education committee	-0.059***	(0.017)
Village has all protected water point	0.027*	(0.016)
Constant	0.063	(0.178)
Observations	50,558	

Notes: (1) We use the ordinary principal component analysis (PCA) to construct the index of household assets. The index is based on household ownership of the following assets: durable assets (TV, radio, car, computer, mobile phone, bicycle, motorbike and cart) and livestock assets (cattle, donkey, camel, sheep/goat); (2) Standard errors are in parenthesis; (3) ***1% significance level, **5% significance level and *10% significance level.

References

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Basu, K. and Van, P. H. (1998). The economics of child labor. *American economic review*, pages 412–427.

Becker, G. S. (1960). An economic analysis of fertility. In *Demographic and economic change in developed countries*, pages 209–240. Columbia University Press.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, pages 669–700.

Blake, J. (1981). Family size and the quality of children. *Demography*, 18(4):421–442.

Booth, A. L. and Kee, H. J. (2009). Birth order matters: the effect of family size and birth order on educational attainment. *Journal of Population Economics*, 22(2):367–397.

Caceres-Delpiano, J. (2006). The impacts of family size on investment in child quality. *Journal of Human Resources*, 41(4):738–754.

Conley, D. and Glauber, R. (2006). Parental educational investment and childrens academic risk estimates of the impact of sibship size and birth order from exogenous variation in fertility. *Journal of human resources*, 41(4):722–737.

Cunha, F. and Heckman, J. (2007). The technology of skill formation. Technical report, National Bureau of Economic Research.

De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review*, 29(4):576–588.

De Haan, M., Plug, E., and Rosero, J. (2014). Birth order and human capital development evidence from ecuador. *Journal of Human Resources*, 49(2):359–392.

Downey, D. B. (2001). Number of siblings and intellectual development: The resource dilution explanation. *American Psychologist*, 56(6-7):497.

Ejrnaes, M. and Portner, C. C. (2004). Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics*, 86(4):1008–1019.

Filmer, D. (2005). Gender and wealth disparities in schooling: Evidence from 44 countries. *International Journal of Educational Research*, 43(6):351–369.

Filmer, D. and Pritchett, L. H. (2001). Estimating wealth effects without expenditure data or tears. an application to educational enrollments in states of india. *Demography*, 38(1):115–132.

Filmer, D., Scott, K., et al. (2008). *Assessing asset indices*, volume 4605. World Bank Publications.

Garg, A. and Morduch, J. (1998). Sibling rivalry and the gender gap: Evidence from child health outcomes in ghana. *Journal of Population Economics*, 11(4):471–493.

Gordon, R. A. (2015). *Regression Analysis for the Social Sciences.* New York: Routlegde.

Gujarati, D. (1970a). Use of dummy variables in testing for equality between sets of coefficients in linear regressions: A generalization. *The American Statistician*, 24(5):18–22.

Gujarati, D. (1970b). Use of dummy variables in testing for equality between sets of coefficients in two linear regressions: a note. *The American Statistician*, 24(1):50–52.

Gupta, M. D. (1987). Selective discrimination against female children in rural punjab, india. *Population and development review*, pages 77–100.

Harkonen, J. (2014). Birth order effects on educational attainment and educational transitions in west germany. *European sociological review*, 30(2):166–179.

Horrace, W. C. and Oaxaca, R. L. (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters*, 90(3):321–327.

Jaeger, M. M. (2009). Sibship size and educational attainment. a joint test of the confluence model and the resource dilution hypothesis. *Research in social stratification and mobility*, 27(1):1–12.

Jones, S., Schipper, Y., Ruto, S., and Rajani, R. (2014). Can your child read and count? measuring learning outcomes in east africa. *Journal of African economies*, 23(5):643–672.

King, E. and Winthrop, R. (2015). Today's challenges for girl's education. Technical report, Brookings Institute Working Paper No. 90.

KNBS (2014). 2014 demographic and health survey key findings. Technical report, Kenya National Bureau of Statistics.

Kristensen, P. and Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science*, 316(5832):1717–1717.

Maitra, P., Pal, S., and Sharma, A. (2014). What explains the gender gap in private school enrolment? recent evidence from india. *Recent Evidence from India (October 31, 2014).*

Maitra, P., Pal, S., and Sharma, A. (2016). Absence of altruism? female disadvantage in private school enrollment in india. *World Development*, 85:105–125.

Mani, S., Hoddinott, J., and Strauss, J. (2013). Determinants of schooling: Empirical evidence from rural ethiopia. *Journal of African Economies*, 22(5):693–731.

Moshoeshoe, R. (2016). Birth order effects on educational attainment and child labour: Evidence from lesotho. Technical report, Economic Research Southern Africa (ERSA), Working paper 621.

Nichols, A. and Schaffer, M. (2007). Clustered errors in stata. In *United Kingdom Stata Users' Group Meeting.*

Piper, B. and Mugenda, O. (2010). The primary math and reading (primr) initiative: Baseline report. Technical report, Kenya National ExaminationsCouncil, Nairobi.

Piper, B., Schroeder, L., and Trudell, B. (2015). Oral reading fluency and comprehension in kenya: reading acquisition in a multilingual environment. *Journal of Research in Reading.*

Piper, B., Zuilkowski, S. S., and Mugenda, A. (2014). Improving reading outcomes in kenya: First-year effects of the primr initiative. *International Journal of Educational Development*, 37:11–21.

Price, J. (2008). Parent-child quality time does birth order matter? *Journal of Human Resources*, 43(1):240–265.

Sahoo, S. (2016). Intra-household gender disparity in school choice: Evidence from private schooling in india. *Journal of Development Studies (forthcoming).*

Tenikue, M. and Verheyden, B. (2007). Birth order, child labor and schooling: Evidence from cameroon.

Tenikue, M. and Verheyden, B. (2010). Birth order and schooling: Theory and evidence from twelve sub-saharan countries. *Journal of African Economies*, 19(4):459–495.

Uwezo (2012). Are our children learning? annual learning assessment report 2012. Technical report, Dar es Salaam: Uwezo. National Assessment System for Monitoring.

Uwezo (2014). Are our children learning? annual learning assessment report- the state of education in kenya in 2015 and beyond. Technical report, Nairobi: Uwezo. National Assessment System for Monitoring.

Vyas, S. and Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health policy and planning*, 21(6):459–468.

Wittenberg, M. and Leibbrandt, M. (2015). *Measuring inequality by asset indices: A general approach with application to South Africa*.

Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *The American Economic Review*, 93(2):133–138.

Zajonc, R. B. (1976). Family configuration and intelligence: Variations in scholastic aptitude scores parallel trends in family size and the spacing of children. *Science*.

Zajonc, R. B. (1983). Validating the confluence model. *Psychological Bulletin*, 93(3):457.

Zajonc, R. B. and Markus, G. B. (1976). Birth order and intellectual development. In *Social psychology in transition*, pages 157–177. Springer.

southern africa labour and development research unit

The Southern Africa Labour and Development Research Unit (SALDRU) conducts research directed at improving the well-being of South Africa's poor. It was established in 1975. Over the next two decades the unit's research played a central role in documenting the human costs of apartheid. Key projects from this period included the Farm Labour Conference (1976), the Economics of Health Care Conference (1978), and the Second Carnegie Enquiry into Poverty and Development in South Africa (1983-86). At the urging of the African National Congress, from 1992-1994 SALDRU and the World Bank coordinated the Project for Statistics on Living Standards and Development (PSLSD). This project provide baseline data for the implementation of post-apartheid socio-economic policies through South Africa's first non-racial national sample survey.

In the post-apartheid period, SALDRU has continued to gather data and conduct research directed at informing and assessing anti-poverty policy. In line with its historical contribution, SALDRU's researchers continue to conduct research detailing changing patterns of well-being in South Africa and assessing the impact of government policy on the poor. Current research work falls into the following research themes: post-apartheid poverty; employment and migration dynamics; family support structures in an era of rapid social change; public works and public infrastructure programmes, financial strategies of the poor; common property resources and the poor. Key survey projects include the Langeberg Integrated Family Survey (1999), the Khayelitsha/Mitchell's Plain Survey (2000), the ongoing Cape Area Panel Study (2001-) and the Financial Diaries Project.

SALDRU

www.saldru.uct.ac.za
Level 3, School of Economics Building, Middle Campus, University of Cape Town
Private Bag, Rondebosch 7701, Cape Town, South Africa
Tel: +27 (0)21 650 5696
Fax: +27 (0) 21 650 5797
Web: www.saldru.uct.ac.za

University of Cape Town · iYunivesithi yaseKapa · Universiteit van Kaapstad